慧智微
SMARTER MICRO

慧智微电子公众号
最新文章持续更新

听慧智微
SmarterMicro on RF
聊射频

从半导体到射频系统

2025

本材料为内部资料，版权归慧智微所有，禁止任何形式商用

序言

射频技术，作为现代信息社会的基石，已深深嵌入人类文明的脉络之中。从无线电波的初次探索，到如今无处不在的无线通信网络，射频技术不仅革新了我们与世界互动的方式，更引领了人类文明的飞跃式发展。它如同空气与水一般，虽无形却至关重要，支撑着信息时代的每一次跨越。

过去百年，射频技术从最初的实验性尝试，逐步成长为全球通信网络的核心支柱。过去的半个世纪里，射频技术更是突飞猛进，从模拟通信迈向数字通信，从 1G 跨越至 5G，从有线网络到 Wi-Fi，每一次技术革新都深刻地改变了人类的生活方式。近十年来，射频技术迎来了前所未有的蓬勃发展。随着 5G、Wi-Fi7 物联网、人工智能等技术的迅猛崛起，射频技术的应用场景日益丰富多样，从智能手机到自动驾驶汽车，从智慧城市到低空通信网络，射频技术正在重塑世界的每一个角落。展望未来，射频技术将继续不断演进。未来十年乃至百年，6G、卫星通信、太赫兹通信等新兴技术将逐渐走向成熟，射频技术将继续成为推动人类文明发展的核心引擎。未来的世界将更加智能、互联和高效，而射频技术将是实现这一美好愿景的关键基石。

慧智微有幸置身于这个充满机遇与挑战的射频行业之中。过去的 14 年里，我们伴随着射频技术的蓬勃发展而成长。在成长的过程中，我们得到了业界专家们的无数次深入指导与交流，也得到了行业同仁的无私帮助与支持。正是有了这些宝贵的支持和讨论，慧智微才能在射频领域不断前行，持续创新，为射频技术的发展贡献自己的力量。

为了更好地传递和表达我们的产品与技术理念，方便产品应用，从 2021 年起，我们开始总结并发布射频相关的技术文章。我们希望通过这些文章，深入浅出地解释复杂的射频原理，并结合实际应用场景，梳理清晰的产品方案与演进逻辑。截至目前，我们已经发布了 50 余篇原创文章。

2024 年，经过电子工业出版社的精心编纂，其中部分文章以《无线重构世界：射频技术的过去、现在与未来》为题正式出版。这本书荣幸地得到了邬贺铨院士的作序，以及联发科技董事长蔡明介先生、龙旗科技集团董事长杜军红博士、慧智微董事长李阳博士，以及来自中国移动、华为、vivo、OPPO 等众多行业专家的推荐。这本书的出版，不仅是对我们过去努力的认可，更是对我们未来工作的激励。

尽管出版社的图书更加系统全面，但我们也了解到，许多工程师更喜欢阅读公众号上的原版文章。这些文章更加原汁原味，每个主题也更加深入透彻。因此，在 2023 年推出第一版《听慧智微聊射频》的基础上，我们推出了第二版。相比第一版，**第二版不仅加入了近两年来的最新文章，还首次提供了电子版**，方便读者随时随地阅读。我们希望通过这种方式，让更多人领略射频技术的魅力，共同探索未来的无限可能。欢迎关注慧智微公众号下载查阅。

射频技术是信息时代的空气与水，它将伴随着人类文明和科技的进步持续演进。我们所做的工作虽然微不足道，但希望它能成为大家在射频探索之路上的小小伴侣，一起走向更加多姿多彩的未来世界。让我们携手并进，共同书写射频技术的辉煌篇章！

彭洋洋
2025 年 2 月 10 日
慧智微

衷心感谢参与慧智微射频技术讨论的每一位专家和射频从业者，

是你们的热情参与以及智慧分享，让本册充满了活力和价值



目录

第一篇 射频芯片中的半导体 **5**

1.1 不只需要光刻机：芯片制造的五大关键工艺 6
1.2 5G 射频芯片中的半导体 17
1.3 收音机为什么叫“半导体”? 34
1.4 胡正明：晶体管的未来，是我们的未来 42
1.5 一片晶圆有多大？可以切出多少芯片？ 47

第二篇 射频芯片设计 **50**

2.1 射频收发机：从无线电到 5G 51
2.2 射频 PA 的工艺选择 59
2.3 5G 射频 PA 的 Load-line 与 Load-pull 62
2.4 高效率 PA 设计：从 Class A 到 Class J 74
2.5 5G 射频 PA 架构 90
2.6 高效率 PA：系统级设计 105
2.7 5G PA“记忆效应”的现象、形成与消除 117
2.8 如何才能不烧射频 PA？ 126
2.9 PA“自激”问题的分析与解决 131
2.10 射频贴片电感的选型与使用 140
2.11 射频开关，简约不简单 149
2.12 高可靠芯片的设计与评估 159

第三篇 射频前端模组 **165**

3.1 从 4G 到 5G，手机射频芯片十年之路 166
3.2 什么是 Phase8？它将给 5G 带来什么变化？ 179
3.3 5G 射频前端模组中的滤波器 192
3.4 5G L-PAMiD 的前世与今生 204
3.5 尺寸缩小 40%：国产工规级 5G n77/n79 L-PAMiF 领先量产 215

第四篇 Wi-Fi 技术与市场 **223**

4.1 Wi-Fi：连接未来，缔造美好生活 224
4.2 Wi-Fi 市场探析：今日应用与未来趋势 234
4.3 Wi-Fi 系统：技术方案与实现路径 241
4.4 Wi-Fi PA 和手机 PA，有什么不同？ 250

第五篇 射频系统设计 **261**

5.1 EVM 详解 262
5.2 除了调匹配，射频人还要掌握系统知识 279
5.3 射频中的“Bypass”与“Decoupling” 290
5.4 协议再演进：从 5G-Advanced 到 6G 297
5.5 手机“射频增强”的技术方案 304
5.6 如何评价一部手机的通信能力？ 317
5.7 今年，是时候了解“毫米波相控阵”了 327
5.8 “iPhone 13 支持卫星通信”，支持的究竟是什么？ 339

5.9 为什么有的手机能连卫星？ 348
5.10 宽带卫星通信市场综述 357
5.11 “有线”与“无线”通信，有什么异同？ 364
5.12 MIPI RFFE 协议解读 387
5.13 为了更流畅地刷抖音，射频工程师都做了什么 400
5.14 蜂窝移动终端的 Cat 指的是什么？ 407
5.15 5G 物联网技术之“5G RedCap” 415
5.16 车联网，需要什么样的射频前端芯片？ 422
5.17 车，怎么联网？ 431

第六篇 射频工具及学习 **441**
6.1 AI 会取代射频工程师吗？ 442
6.2 射频最推荐书目榜-慧智微射频问答群 463
6.3 慧智微射频技术百问百答发布 468
6.4 《无线重构世界》正式全面开售 472
6.5 射频人专属鼠标垫更新 475
6.6 慧智微射频工具箱小程序上线 478

第一篇 射频芯片中的半导体

本篇内容

- 不只需要光刻机：芯片制造的五大关键工艺
- 5G 射频芯片中的半导体工艺
- 收音机为什么叫“半导体”?
- 胡正明：晶体管的未来，是我们的未来
- 一片晶圆有多大？可以切出多少芯片？

在当今这个信息爆炸的时代，射频芯片作为无线通信系统的核心元件，扮演着举足轻重的角色。它们不仅是信号传输与接收的桥梁，更是决定通信质量、速度和效率的关键因素。而射频芯片之所以能够在如此微小的尺寸内实现如此复杂的功能，很大程度上得益于半导体技术的巧妙运用。

射频芯片与其他射频电路的不同之处在于，它采用了半导体技术来设计并制造芯片。半导体，这一介于导体和绝缘体之间的神奇材料，具有独特的电学性质，使得我们可以在其基础上构建出各种复杂的电子器件和电路。因此，对于想要深入了解射频芯片技术的读者来说，掌握一些基本的半导体知识是非常有益的。

在这一章节中，我们将从半导体的角度出发，深入探讨射频芯片的技术奥秘。首先，我们将介绍芯片工艺的主要步骤。接着，我们将重点介绍 5G 射频芯片中所使用到的半导体技术。此外，我们还将简要介绍晶圆的基本知识。另外，也会分享半导体发展过去和未来的一些趣事。

让我们一起走进芯片的半导体世界，探索这个微小而神奇的领域。

1.1 不只需要光刻机：芯片制造的五大关键工艺

作者：彭洋洋 博士

2023 年 3 月 21 日，英伟达创始人兼 CEO 黄仁勋再次穿着标志性皮衣亮相，发布了多款针对 AI 的最新技术。不同于以往 AI 只是针对于软件产品，这次英伟达还发布了用于先进制程芯片设计和制造，可以让计算光刻变的更“聪明”的加速技术。一时间让光刻机再次成为聚焦。

近年来，“芯片”成了热词，而代表芯片最底层最基础的“芯片制造工艺”也成了全民关注的焦点，这其中最夺目的，莫过于舞台中央的“光刻机”。

“光刻机”是芯片制造中的关键设备，并且随着芯片技术演进，晶体管特征尺寸越来越小，需要用到的光刻机就越尖端。因为光刻机技术的差距，以及光刻机故事的不断被渲染，“光刻技术”仿佛成为了半导体技术的唯一重要技术。

其实在芯片制造工艺中，除了光刻工艺外，还有其他多个重要工艺步骤，这些步骤同样不可或缺。这些关键步骤与光刻一起，才能共同实现芯片的“点沙成金”。这些关键步骤主要包括：

1. 晶圆制备：画布始开
2. 氧化工艺：制作铠甲
3. 光刻刻蚀：图形绘制
4. 掺杂工艺：灵魂注入
5. 薄膜工艺：阡陌交通

本文中，就对这些工艺实现做一个简单探讨，一同探究半导体工艺中的五大关键步骤。

1.1.1 晶圆制备：画布始开

半导体工艺的第一步，就是制造晶圆。晶圆是一种很薄而且非常光滑的半导体材料圆片，是集成电路的“画布”。一切后续的半导体工艺都是在这个“画布”上展开。图1.1为半导体晶圆片照片。



图 1.1: 半导体晶圆片

以硅基晶圆为例，半导体晶圆的主要制备步骤有 [1]：

1. 硅提炼及提纯：大多数晶圆是由从沙子中提取的硅制成的。将沙石原料放入电弧熔炉中，还原成冶金级硅，再与氯化氢反应，生成硅烷，经过蒸馏和化学还原工艺，得到高纯度的多晶硅。

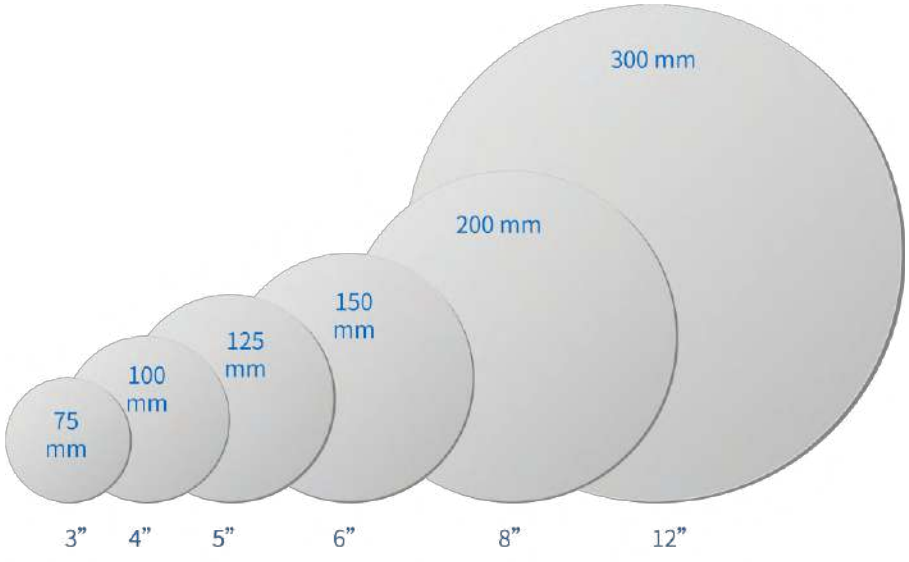


图 1.2: 晶圆尺寸的演进

2. 单晶硅生长：将高纯度的多晶硅放在石英坩埚中，并用外面围绕着的石墨加热器不断加热，使多晶硅熔化。然后把一颗籽晶浸入其中，并且由拉制棒带着籽晶作反方向旋转，同时慢慢地、垂直地由硅熔化物中向上拉出。这样就形成了圆柱状的单晶硅晶棒。
3. 晶圆成型：将单晶硅棒经过切段、滚磨、切片、倒角、抛光、激光刻等工序，制成一片片薄薄的半导体衬底，即晶圆。

半导体晶圆的尺寸在这一步骤中确定。晶圆的尺寸一般以“英寸”为单位。在半导体行业的早期，由于工艺能力的限制，硅棒直径只有 3 英寸，约合 7.62 厘米。此后，随着技术进步和生产效率提高，晶圆尺寸不断增大。目前，在半导体制造中使用的最大直径为 12 英寸（又称 300 毫米）。图1.2为不同直径的晶圆尺寸。

在芯片晶圆上，有一些特殊的部分和特定的名称，比如如图1.3所示:

1. Wafer：指整张晶圆
2. Chip、Die：是指一小片带有电路的硅片
3. 划片道（Scribe line）：指 Die 与 Die 之间无功能的空隙，可以在这些安全的切割晶圆，而不会损坏到电路
4. 测试单元：一些用于表征 Wafer 工艺性能的测试电路单元，规律分布于 Wafer 各位置
5. 边缘 Die（Edge Die）：Wafer 边缘的一部分电路，通常这部分因为工艺一致性或切割损坏，会被损失。这部分损失在大的晶圆片中占比会减少
6. 切割面（Flat Zone）：被切成一个平面的晶圆的一条边，可以帮助识别晶圆方向

晶圆制备完成后，半导体的画布就形成了。后续半导体工艺由此开始。

1.1.2 氧化工艺：制作铠甲

在半导体电路中，除了用于可控导电的各种二极管、三极管外，还必须要用绝缘物质将不同的电路隔离开来。对于硅基元素来说，形成这种绝缘物质最方便的方法就是将硅进行氧化，形成二氧化硅（SiO_2）了。

SiO_2 是自然界中常见的一种材料，也是玻璃的主要元素。SiO_2 材料的主要特点有：

1. 具有高熔点和高沸点（分别为 1713 ºC 和 2950º C）
2. 不溶于水和部分酸，溶于氢氟酸
3. 具有良好的绝缘性、保护性和化学稳定性

由于上特性，SiO_2 在芯片制备的多个步骤工艺中被反复使用。芯片工艺中的氧化工艺是在半导体制造过程中，在硅晶圆表面形成一层薄薄的 SiO_2 层的过程。如图1.4所示，这层氧化层有以下作用：

1. 作为绝缘层，阻止电路之间的漏电

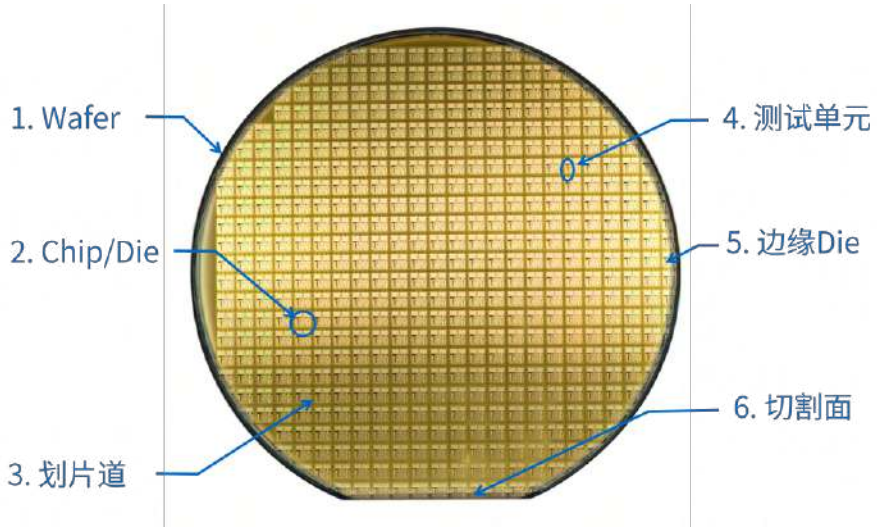


图 1.3: 芯片晶圆的构成

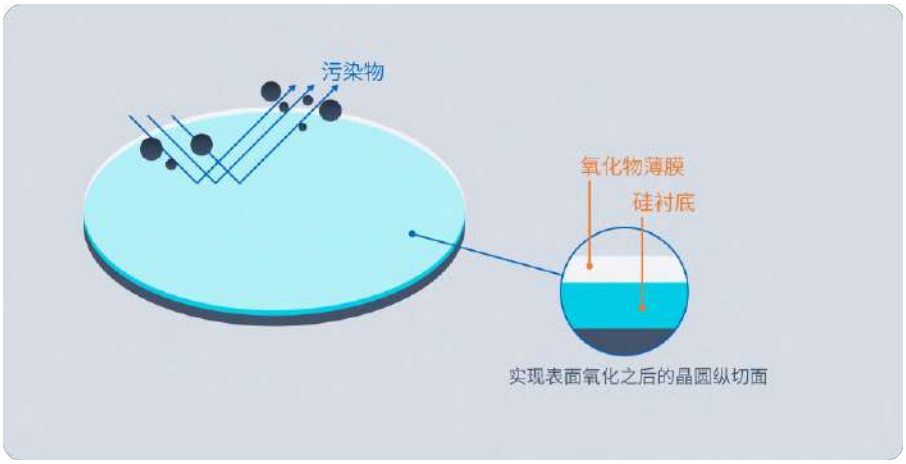


图 1.4: 氧化物在晶圆表面的保护作用

2. 作为保护层，防止后续的离子注入和刻蚀过程中对硅晶圆造成损伤
3. 作为掩膜层，定义电路图案

这些氧化层在半导体器件中也有举足轻重的作用。比如说 CMOS 器件中的重要结构：MOS（金属-氧化物-半导体）结构中用于金属和半导体之间绝缘的“氧化物”层（或称栅氧），就是采用氧化工艺制备的。另外，用于隔离不同 CMOS 器件的厚层氧化物场氧（Field Oxide）、SOI 器件中用于隔离衬底与器件的绝缘隔离层，都是采用氧化工艺实现的 SiO_2 材料。

氧化工艺的实现方法有多种，如热氧化、电化学阳极氧化等。其中最常用的是热氧化法，即在高温（800～1200℃）下，利用纯氧或水蒸汽与硅反应生成 SiO_2 层。如图1.5所示，热氧化法又分为干法和湿法：

1. 干法只使用纯氧，形成较薄、质量较好的氧化层，但生长速度较慢。
2. 湿法使用纯氧和水蒸汽，形成较厚、密度较低的氧化层，但生长速度较快。

不同类型和厚度的 SiO_2 可以满足不同功能和要求。

在半导体工艺中，氧化工艺非常重要，它为后续的制造步骤提供了基础和保障。氧化层不仅可以隔离和保护硅晶圆，还可以作为掩膜层来定义电路图案。没有氧化层，半导体器件就无法实现高性能、高可靠性和高集成度。

SiO_2 和部分氧化物有透光特性，由于这些材料的厚度不同，就会对特定波长的光线产生衍射或反射，也就使芯片表面看上去五彩斑斓，如图1.6。所以芯片表面的颜色并不是真正的彩色，而是这些薄膜结构对光的反射或干涉。

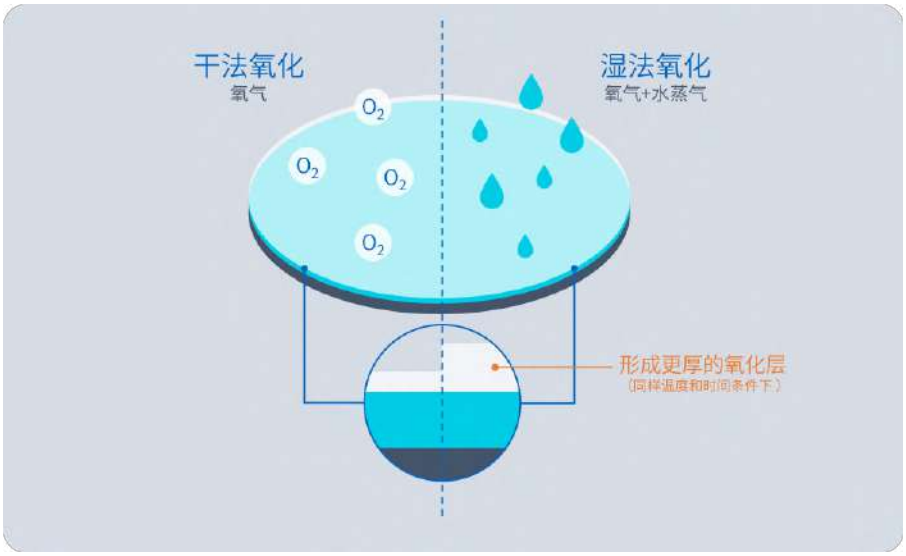


图 1.5: 干法氧化和湿法氧化

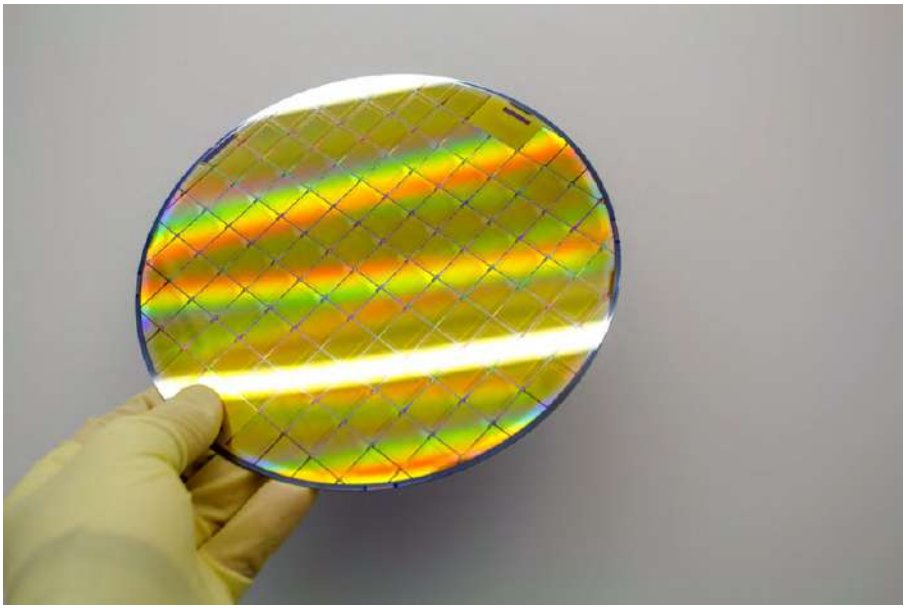


图 1.6: 五彩斑斓的晶圆表面

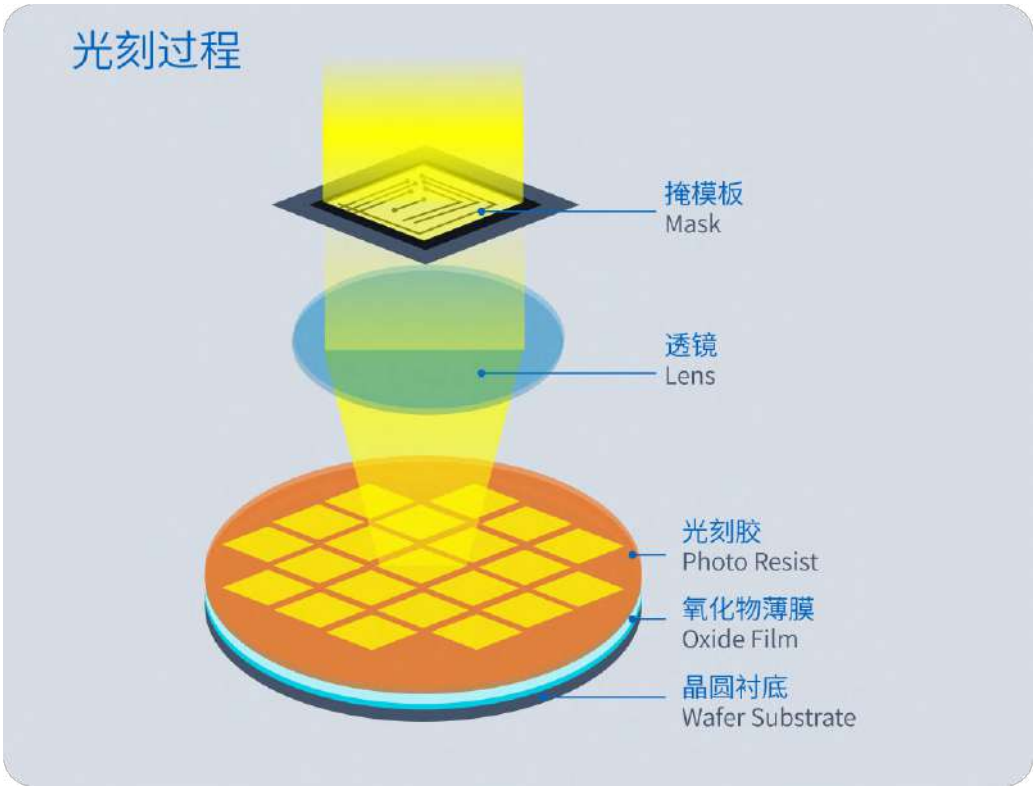


图 1.7: 半导体的光刻工艺

通过氧化工艺，脆弱的硅基晶圆就像穿上了一层“铠甲”。

1.1.3 光刻/刻蚀：图案绘制

有了画布素材，终于可以任由芯片设计师挥毫泼墨，自由创作了。光刻和刻蚀步骤就是将芯片设计师所设计的图案，转移到晶圆片上的过程。

1.1.3.1 光刻：完成图案转移

光刻技术是一种将掩模板（Mask）上的图形转移到涂有光刻胶的晶圆片上的技术。光刻技术可以将半导体表面上特定的区域去除或者保留，从而构建半导体器件，如图1.7所示。

如图1.8所示，光刻步骤主要包括：

1. 设计电路并制作掩模板。这一步一般是通用计算机辅助设计（CAD）软件完成的，在完成电路设计正确性检查（LVS）和设计规则检查（DRC）后，设计图形被转移到掩模板上。掩模板一般是由透明的超纯石英玻璃基片制成，在基片上，需要透光的地方保持透明，需要遮光的地方用金属遮挡。
2. 涂光刻胶：使晶圆对光敏感。执行这一步骤时，会在晶圆表面均匀涂抹一层对光敏感的物质，光刻胶。光刻胶对光敏感，光照射后会产生化学变化，于是根据光照射与否，光刻胶也形成溶解和不可溶解的部分。
3. 曝光。将光源发出的光线经过掩模板照射到晶圆片上时，掩模板上的图形也就被转移到了晶圆片上。根据掩模板上图形的不同，光刻胶会溶解形成对应图形。
4. 显影与坚膜。用化学显影液溶解掉光刻胶中可溶解的区域，使可见的图形出现在晶圆片上。显影后再进行高温烘培，使剩余的光刻胶变硬并提高粘附力。

光刻之所以得名，就是因为它通过利用光线，把带有图案的掩模板上的图形转移到晶圆片上。由于半导体技术的主要目标是尽可能的缩小电路尺寸，所以对光刻的精度要求也越来越高。高精度的光刻机是光刻步骤的基础，这就是为什么“光刻”成为备受关注的工艺步骤。

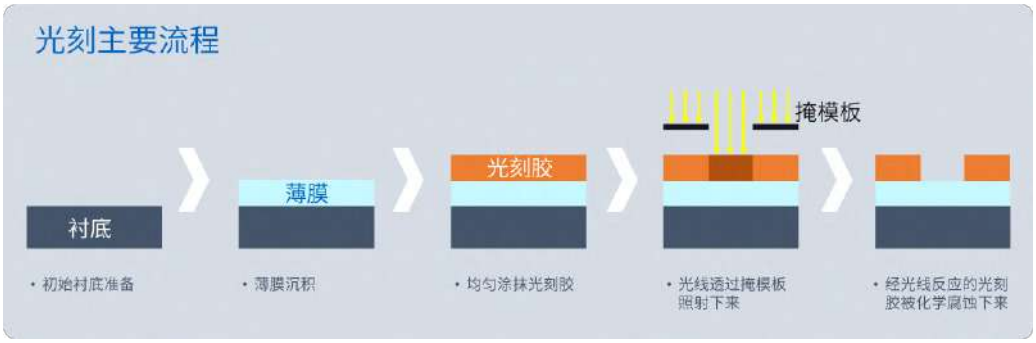


图 1.8: 光刻的主要流程

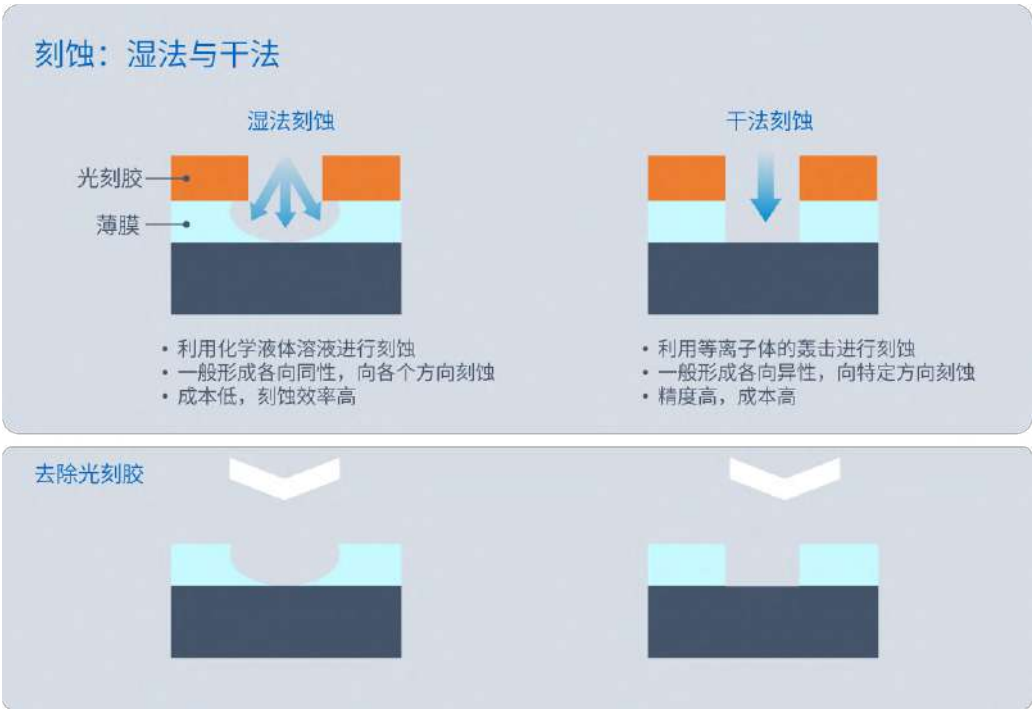


图 1.9: 湿法刻蚀和干法刻蚀

为了支持更高精度的光刻，也有先进的光刻机被制造出来。目前最先进的光刻机技术是极紫外光刻技术（EUV，Extreme Ultra-violet），它使用波长为 13.5 纳米的极紫外线作为光源进行电路光刻，可以制造出 7 纳米及以下工艺节点的芯片。ASML 是 EUV 光刻机的领导厂商，其最新型号的光刻机号称可实现 0.3 纳米的精度。

光刻技术是半导体芯片工艺中最昂贵的工艺，在先进工艺中，光刻步骤的成本可以占整个芯片加工成本的三分之一甚至更多。

1.1.3.2 刻蚀：形成图案

经过了光刻步骤之后，所需要的图案已经被印在了晶圆表面的光刻胶上。但要实现半导体器件的制作，还需要把半导体器件按照光刻胶的图形复刻出来。这个复刻的过程就叫刻蚀（Etching）。

半导体刻蚀方法分为两类，分别是湿法刻蚀和干法刻蚀，如图1.9所示。

- **湿法刻蚀**

  湿法刻蚀是将晶圆片浸入到含有特定化学剂的液体溶液中，利用化学反应来溶解掉未被光刻胶保护的半导体结构。由于液体化学品不能很好的控制方向性，所以可能会导致刻蚀不均匀，造成刻蚀的不足或过度；另外，由于液体化学品会残留在晶圆上，所以需要额外的清洗步骤来去除污染物。

- **干法刻蚀**

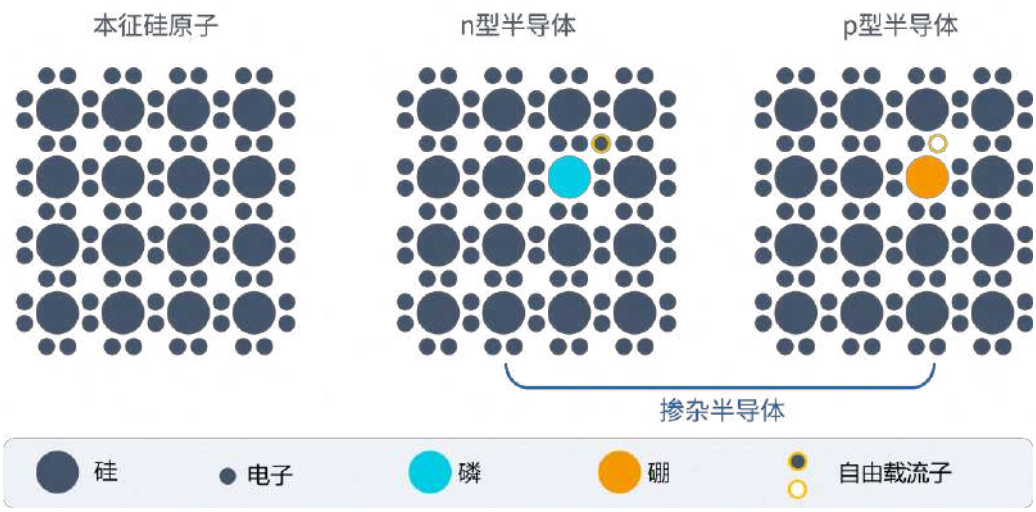


图 1.10: 掺杂实现半导体特性的控制

干法刻蚀是用等离子体或者离子束等来对晶圆片进行轰击将未被保护的半导体结构进行去除的方法。相比于湿法刻蚀，干法刻蚀精度高、选择性和方向性好，并且不会产生残留物，适用于制造高集成度的芯片。然而干法刻蚀也有缺点，例如成本高，设备复杂，处理时间长。

在半导体工艺步骤中，会根据不同的目标和需求，灵活选择最适合的工艺。甚至在同一个器件制作的不同步骤中，会混合使用湿法刻蚀和干法刻蚀。

通过光刻及刻蚀步骤，就可以将希望的图形，在晶圆片上真正实现。需要说明的是，光刻/蚀刻步骤一次只能实现一层半导体结构，由于半导体器件是多层器件，通过需要迭代多次才能将半导体器件完整蚀刻出来。并且随着工艺复杂度的不同，需要的层数也不同。例如在 0.18 微米的 CMOS 工艺中，需要的光罩层数约为 20 层，而对于 7nm 左右的 CMOS 工艺来说，则需要 55-60 层 [3]。

1.1.3.3 掺杂工艺：注入灵魂

如果说以上工艺步骤是任何微机械器件都必须要考虑的工艺的话，那么“掺杂工艺”（Doping）就是半导体工艺中的灵魂工艺了。

说掺杂工艺是半导体工艺中的灵魂工艺，是因为电路中各半导体器件的电学性能在此步骤形成。在此步骤之前，整片晶圆不过是一片冷冰冰的材料一片，过了此步骤，才有了各种各样的二极管、三极管、CMOS 以及电阻等。

掺杂工艺在半导体中如此重要，是因为它可以改变半导体的电导率、载流子类型和浓度、能带结构等电学性质，从而实现不同的功能和性能。半导体的导电性能可控，就是通过掺杂来实现，如图1.10所示。

有了掺杂工艺，就可以制造出 PN 结、双极型晶体管、场效应晶体管等基本的半导体元器件，也可以用来调节 MOS 晶体管的阈值电压、改善接触电阻、增强辐射耐受性等。掺杂工艺是半导体工艺中最核心和最基础的技术之一，对于半导体器件的设计和制造具有决定性的影响。

半导体工艺中实现掺杂的主要方法有两种，即热扩散和离子注入，两种方法对比如图1.11所示。

1. 热扩散是在高温下（约 1000℃）将半导体暴露在一定掺杂元素的气态下，利用化学反应和热运动使杂质原子扩散到半导体表层的过程
2. 离子注入是将杂质原子电离成离子，用高能量的电场加速，然后直接轰击半导体表面，使杂质原子“挤”进到晶体内部的过程

经过了掺杂工艺，从沙子而来的晶圆片终于有了灵魂。

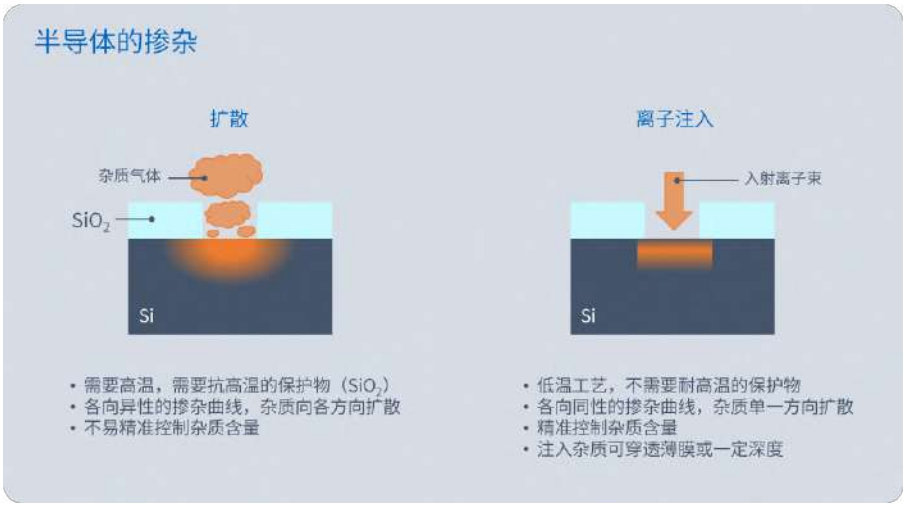


图 1.11: 两种掺杂工艺对比

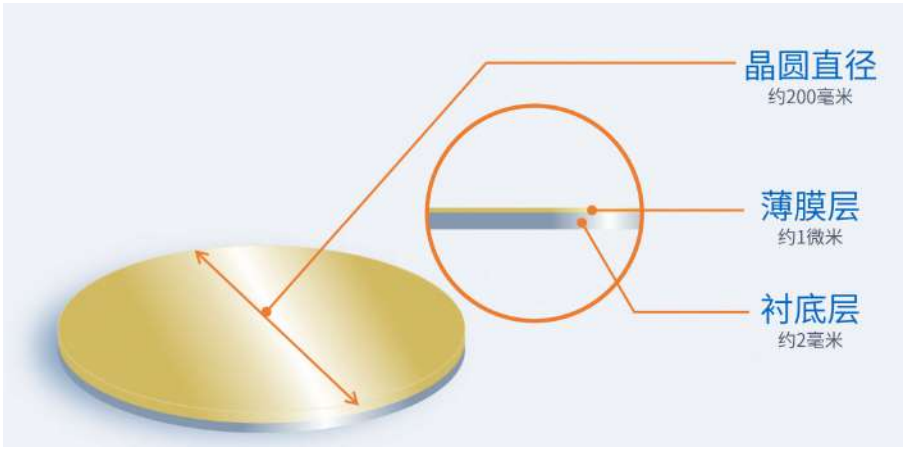


图 1.12: 半导体器件中大部分结构在 1 微米左右厚度的薄膜中实现

1.1.4 薄膜沉积：阡陌交通

薄膜工艺就是指在半导体晶晶圆片上沉积各种材料，以实现不同电路功能或特性的过程。和前面几个工艺步骤是按工艺过程来命名不同，这一步骤命名方式是按材料的“状态”来命名的。这个命名甚至让人看起来有些费解，从示意图看，半导体器件看上去都是厚厚的一个个器件，哪来的薄膜？

其实为了理解方便，半导体器件结构在图示时经过了抽象和不等比例的放大，实际中半导体器件在晶圆片上非常薄的一层内实现。通常厚度在 1 微米以内。对于一个 8 英寸（200 毫米）的晶圆片来说，1 微米的厚度薄膜的制作，相当于在 200 米直径的操场上，均匀的堆积 1 毫米厚的沙子，如图1.12所示。

这层薄膜非常的薄，但功能却非常强大。各个半导体器件之间或是金属连接，或是电场联系，均需要用这些薄膜层实现。薄膜层实现了各器件间复杂贯穿的“阡陌交通”，如图1.13所示。

正是因为这层结构非常的薄，于是就被称作“薄膜”工艺。这么薄的膜层不能通过机械方式来制造，于是，掺杂“沉积”（Deposition）工艺被发明出来。

1.1.4.1 沉积：晶圆画布上的喷涂刷

在半导体工艺里，沉积是指在原子或分子水平上，将材料沉积在晶圆表面作为一个薄层的过程。沉积工艺就像是喷涂刷，将涂料均匀的薄薄喷洒在晶圆表面上。

根据实现方法的不同，沉积主要分为物理气相沉积（PVD）和化学气相沉积（CVD），两种工艺比较如图1.14所示。

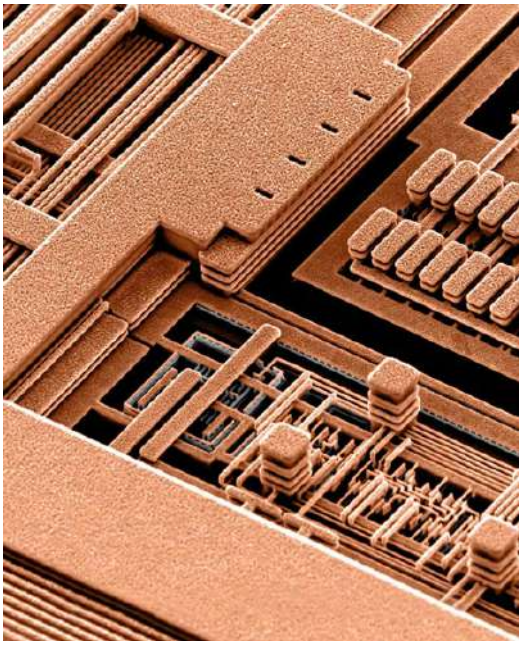


图 1.13: 显微镜下的芯片一角

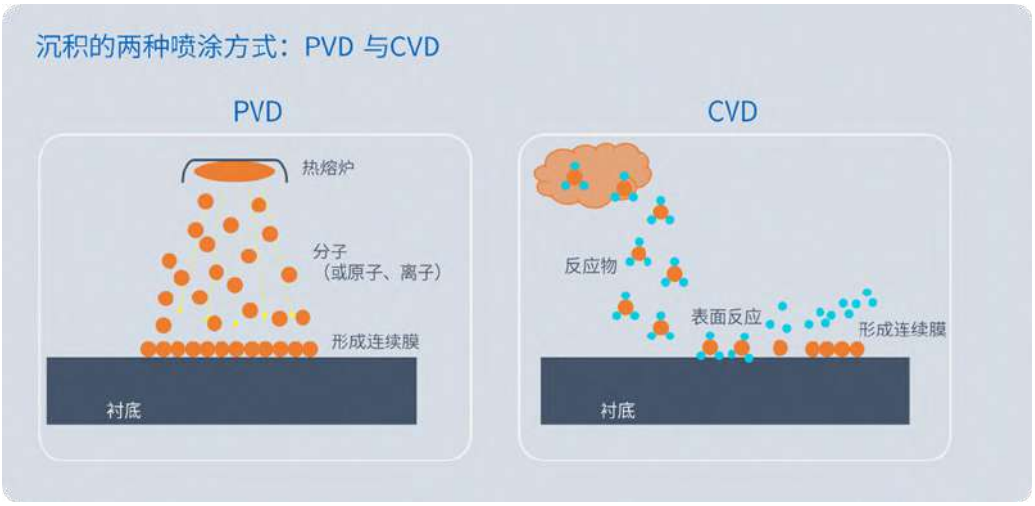


图 1.14: 薄膜沉积中的两种工艺：PVD 与 CVD

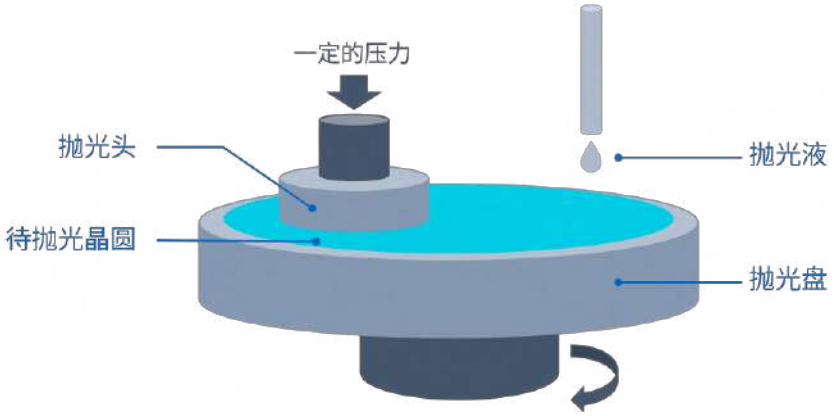


图 1.15: CMP（化学机械抛光）工艺

PVD 是利用物理方法，将材料源气化成气态原子、分子，或电离成离子，并通过低压气体，在基体表现沉积成薄膜的过程。一般用来沉积金属薄膜。

CVD 是利用含有薄膜元素的一种或几种气相化合物，在衬底表面进行化学反应形成薄膜的方法。一般用于沉积半导体或绝缘体，以及金属合金等。

为了增强化学反应，CVD 也可以与其他方法相结合。如 PECVD（等离子增强 CVD，Plasma Enhanced CVD）就是利用等离子体来激活化学反应，改善 CVD 的方法。

根据不同目标和需求，PVD 和 CVD 在实际工艺流程中也可以自由选择

1.1.4.2 抛光：喷涂后的找平

随着半导体技术的进步，对各薄膜层精度的要求也越来越高，于是，需要对晶圆表面进行平坦化，消除不同材料层之间的起伏和缺陷，提高光刻的精度和质量。

抛光（Polishing）就是用于晶圆表面薄膜层平整化的技术。抛光工艺中，最主要的工艺是 CMP（化学机械抛光，Chemical Mechanical Polishing），CMP 是一种利用化学腐蚀和机械摩擦的结合来实现晶圆表面平坦化的技术，研磨对象主要是浅沟槽隔离 (STI)，层间膜和铜互连层等，如图1.15所示。

1.1.5 半导体的工艺整合

以上就是芯片制造中的主要工艺，以上工艺以硅基半导体为主要参考，其他工艺（如 GaAs、SiGe 等化合物半导体）会略有不同，但基本思路一致。

在具体半导体工艺实现上，通过将以上关键工艺的有机整合，形成一个完整的工艺流程，就可以完成半导体工艺的开发。

虽然看起来半导体工艺步骤并不是太过复杂，但在实际工艺中，需要考虑各个工艺步骤之间的影响和优化；需要确保各工艺步骤的稳定性和一致性；需要达到最佳的器件特性、产量和成本，使得半导体工艺整合变得非常重要，是推动半导体行业发展和创新的关键因素，也成为半导体器件代工厂核心竞争力之一。

1.1.6 总结

芯片的制造工艺是一个复杂的过程，关键工艺也并不只有光刻，还包括晶圆加工、氧化、刻蚀、掺杂、薄膜沉积等多个步骤，每个步骤都对半导体性能和功能有重要影响。

了解芯片的制造步骤同样对设计芯片、应用芯片有帮助。在芯片制造产业链越来越向国内靠拢的大背景下，希望大家有机会可以深入了解芯片的制造工艺，通过整个供应链的深度理解与整合，提升芯片产品竞争力。

本文撰写过程中得到慧智微工艺团队大力支持，在此表示感谢！

参考文献

1. 晶圆制备工艺，https: / /zhuanlan.zhihu.com /p /366924703
2. TSMC 0.18 μm CMOS Process Technology –CMC Microsystems; https: / /www.cmc.ca /tsmc-0-18-um-cmos
3. 7nm Technology - Taiwan Semiconductor Manufacturing Company Limited (tsmc.com)
4. The Evolution of Microprocessors，https: / /www.ibm.com /ib m/history /ibm100 /us /en /icons /copperchip
5. 温德通，集成电路制造工艺与工程应用，机械工业出版社，2019
6. 张汝京，纳米集成电路制造工艺，清华大学出版社, 2014
7. MichaelQuirk 等，半导体制造技术，电子工业出版社, 2015.

在线阅读

本文摘自慧智微微信公众号，扫描以下二维码，在线阅读

1.2 5G 射频芯片中的半导体

作者：彭洋洋 博士

最近在与手机终端客户交流过程中，发现客户对于射频前端芯片的理解已经越来越深入：从最早的只把射频芯片当成一个黑盒子，到对 PA 等模块架构特点熟悉，再到对放大原理深入了解，国内手机终端公司逐渐把射频芯片从“用得了”，到“用的好”。

在对射频芯片的工作原理了解清楚之后，一些手机终端公司逐渐向芯片的更深层次：半导体器件层级研究。考虑研究深入的器件物理机理，对应用中看到的问题加以解释。

芯片之所以是芯片，就是因为半导体技术；芯片系统与板级系统最大的区别，也就是因为使用到了半导体物理器件。半导体器件是芯片公司的根基所在，手机终端公司对于半导体器件的研究，让我们这些芯片公司的“硅农”们感受到同道相益，备受鼓舞。

不过我们也发现，即使对于芯片公司或芯片工程师来说，“半导体物理”都是一个让人望而生畏的词。由于半导体物理太过晦涩难懂，很多的集成电路设计类专业甚至不再教授此课程；即使对于半导体专业的学生来说，学校里的“半导体物理”、“半导体器件”课程也常常成为学生时代的心理阴影，以至于工作之后再也不想谈及。

于是，我们就想梳理出一份材料，一起就射频芯片中常见的半导体器件做一个讨论。

1.2.1 半导体工艺简介

1.2.1.1 什么是半导体

在半导体被发现之前，人们认为世界上的材料根据导电性分类只可以被分为“导电”和“不导电（绝缘）”两种。按分类方法中的“相互独立、完全穷尽”的原则看，这是对世界上材料非常完美的分类方法，那为什么还会出现“半导体”这一分类呢？

一些文献中将“半导体”定义为“导电性能介于导体和绝缘体之间的材料”，准确来讲，半导体材料并不是导电的性能处于二者之中，而是导电特性可以在二者之间可控切换。**这种导电特性可以在导体与绝缘体之间可控切换的材料，被称为半导体材料。**

半导体现象的首次被发现要追溯到近 200 年前的 1833 年，电子之父法拉第发现硫化银的电阻随温度变化特性不同于一般金属，温度可以实现对硫化银材料导电性的可控，这是人类首次观察到半导体现象。

在 1833 至 1945 年这 100 多年时间里，物理学家对这一现象进行了深入研究。20 世纪初的物理学革命为半导体科技奠定了坚实的理论基础，而材料生长技术为半导体科技奠定了实现中的物质基础。

1.2.1.2 掺杂：半导体材料的实现基础

半导体导电可控特性的实现是通过**“掺杂（Doping）”**来实现。

以硅原子为例，每个硅原子最外层有 4 个电子，在本征硅材料中，每个硅原子与周围的 4 个硅原子形成共价结合的稳定结构，从而没有可自由流动的自由载流子，不能形成电流。

这时，如果对本征硅材料进行掺杂，如果加入最外层有 5 个电子的磷元素，这时在形成 4 个共价键之外，还会多出一个自由移动的电子，这个电子就是一个**自由载流子**，当上电压之后，掺杂材料就可以导电。自由移动的电子是半导体材料中的第一种载流子。

同理，如果掺杂材料为最外层只有 3 个电子的硼元素，这时会出现一个电子的空缺，电子在不断填满这个空缺的过程中，也可以使材料导电。由于电子不断填充这个空缺的过程不易描述，人们就发明了一个新的表征空缺的方式，即定义一种新的载流子来表示这个空缺，这种新的载流子就是半导体中的第二种自由载流子：空穴。

半导体的掺杂如图1.17所示，半导体的掺杂工艺实现了对半导体内自由载流子的控制。

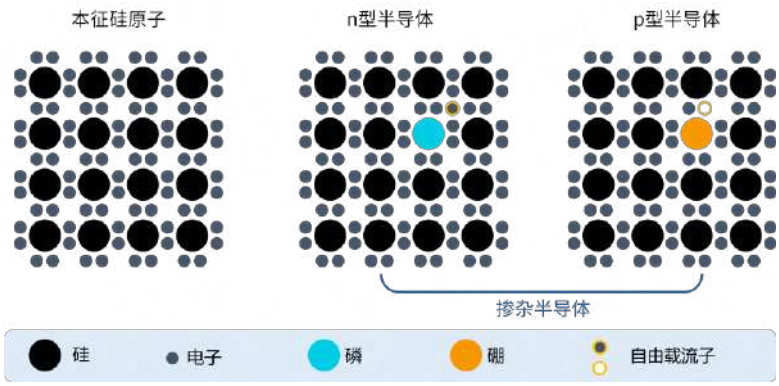


图 1.17: 半导体掺杂

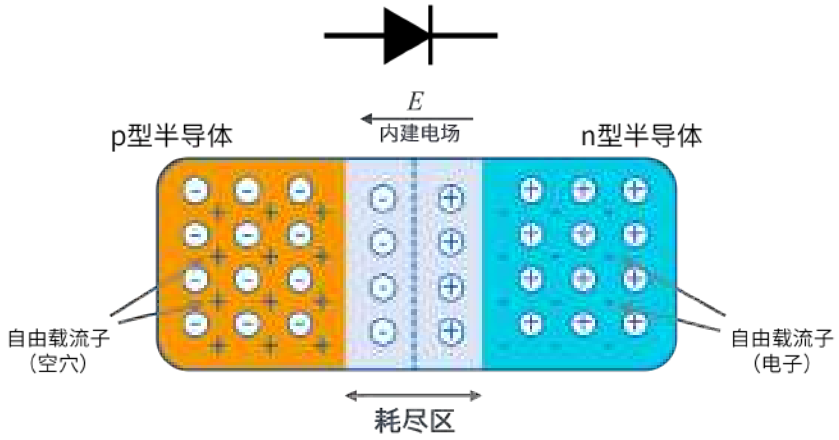


图 1.18: p-n 结的基本结构

1.2.1.3 p-n 结：最简单的半导体器件

利用半导体的掺杂特性，就可以设计出简单的半导体器件：p-n 结。p-n 结英文名称是 p-n Junction，因为有正负两个端口，所以又称为 p-n 结二极管。

p-n 结是在同一衬底上同时进行 p 型和 n 型掺杂，并使之交界，这样在二者交界处就形成耗尽区（也叫空间电荷区），从而形成 p-n 结。p-n 结的结构如图1.18所示。

p-n 结的一个重要特性就是单向导电特性。当 p 型半导体侧加入正电压时，p 型半导体中的空穴在外加电场作用下向右侧移动。当外加电场大于 p-n 结的内建电场时，空穴就会跨越过耗尽区，从而进入 n 型半导体区，之后在电场作用下进入电源负极，形成电流。

当 n 型半导体侧加入正电压时，n 型半导体中的自由载流子电子向右侧移动，拉大内建电场宽度，使自由载流子更难跨越耗尽区，无法形成电流。

p-n 结的导电特性如图1.19所示。

1.2.1.4 从“晶体管”开始，直到改变世界

只是用 p-n 结二极管，还不足以设计集成电路。

在半导体物理的基本原理被人类掌握之后，人类就开始用半导体材料设计制造一些特殊器件。比如：

- 利用半导体导电性能与温度之间的关系，可以设计出热敏电阻，来感知温度变化；
- 利用有些半导体导电性能与光照之间的关系，可以设计出光敏电阻，来感知光的变化；
- 利用有些半导体光电转换特性，可以实现电能和光能的相互转化，设计出发光器件，或者设计出光伏发电器件；

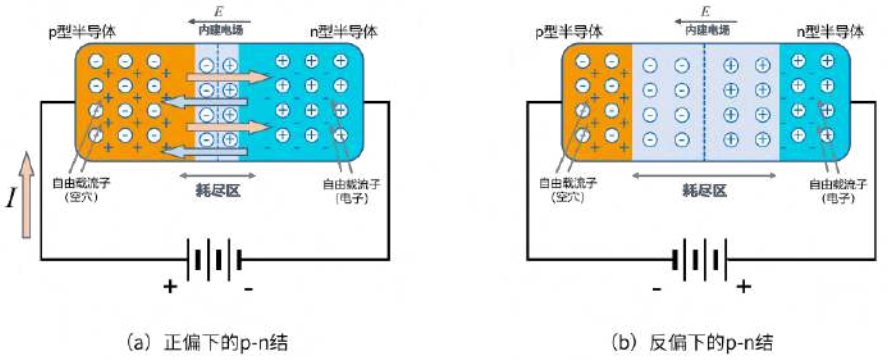


图 1.19: 正偏及反偏下的 p-n 结



图 1.20: 20 世纪初，RCA 公司采用电子管设计的放大器

虽然利用半导体材料“导电性能可控”这一特性，可以在很多领域设计出重要应用的器件。不过这个时期的半导体还是很难和之后改变世界的“集成电路”联系起来。真正建立起“半导体材料”与“集成电路”之间联系的，是 1945 年被发明的“晶体管”。

早在 1899 年，人类就实现了无线电信号的跨英吉利海峡穿越。但在晶体管被发明之前，电子电路系统一般由真空电子管设计。真空电子管体积大、功耗大、发热厉害、寿命短，并且需要高压电源，所以真空电子管所设计的电路一般只用在政府、军事部门中。真空管极大地限制了电路系统的大规模应用。图1.20为 20 世纪初 RCA 公司采用电子管设计的放大器。

为了克服真空电子管的局限性，第二次世界大战后，贝尔实验室加紧研究，探讨用半导体材料制作放大器件的可能性。

1945 年秋天，贝尔实验室成立了以肖克莱为首的半导体研究小组，成员有布拉顿、巴丁等人。他们发现，在锗片的底面接上电极，在另一面插上细针并通上电流，然后让另一根细针靠近及接触它，并通上微弱的电流，这样就会使原来的电流产生很大的变化。

微弱电流少量的变化会对另外的电流产生很大的影响，这就是“放大”作用。利用这种特性，半导体器件也就可以被用于制作放大器。

在为这种器件命名时，布拉顿想到它的电阻变换特性，**于是取名为 Trans-resistor(转换电阻)，后来缩写为 Transistor，中文译名就是晶体管。** 1956 年，肖克莱、巴丁、布拉顿三人，因发明半导体晶体管同时荣获诺贝尔物理学奖。图1.21为三人照片。

在直观理解上，可以将晶体管理解成一个水龙头，如图1.22所示：

- 晶体管一共有三个极，对于 FET 器件，一般称为源、漏、栅（S、D、G），对应水龙头的进水口、出水口、龙头把手；
- 晶体管源极（Source）流入电子，对应水龙头进水中流入水源；



图 1.21: 半导体晶体管的发明者，左起：布拉顿，肖克莱，巴丁

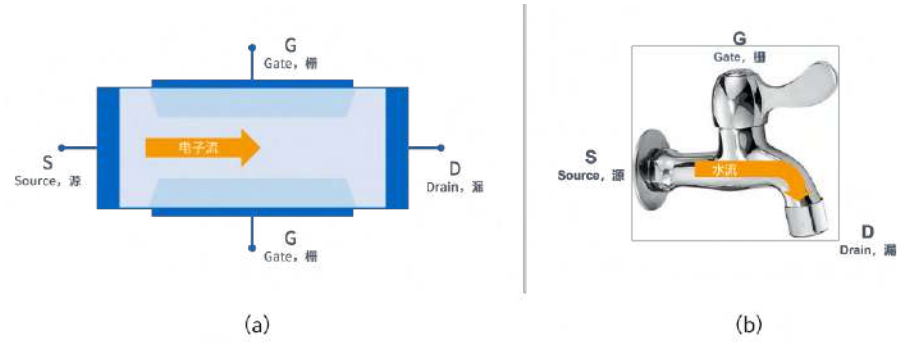


图 1.22: 典型的晶体管构造（JFET），及与水龙头的等效

- 栅极（Gate）是晶体管的核心，控制电流的大小，对应于龙头把手是水龙头的核心，决定水流的强弱；

基于以上原理，晶体管就可以实现信号的放大。即在栅极加一个微小信号，只要能控制住晶体管漏极到源极这个通道的通断，就可以控制漏源之间电流的大小。

晶体管还可以实现数字电路里“0”、“1”基本信号的表征。比如可以定义“水龙头”打开状态为“1”，关闭状态为“0”。成千上万个“水龙头”放在一起，就可以进行数字逻辑运算。

晶体管（Transistor）是集成电路领域中最为重要的基本器件，没有之一。有了这个基本器件，才有了现在改变世界的集成电路。**我们经常看到的 FET、HBT、pHEMT 等器件名称中的最后一个字母“T”，均是“Transistor”的缩写。**

在半导体器件里，晶体管也有多种变形，材料也多种多样，一些缩写也让人眼花缭乱。不过万变不离其宗，只要找到晶体管中水龙头的“进水口”和“出水口”，理解清楚“龙头把手”的工作原理，就可以将这种类型晶体管分析清楚。

1.2.1.5 晶体管的分类

过去近 100 年，半导体晶体管的发展过程中，先驱科学家和工程师们不断尝试各种器件结构与器件材料，来满足不同领域集成电路特性的需求。产生出各种各样的半导体器件。

半导体本来就抽象、难以捉摸，纷杂的名称和简称也使得半导体更加高深莫测，给非半导体出身的电子行业从业者带来不少的困扰。

晶体管的分类主要可以从两个方面下手理解，分别为器件结构，和器件材料。图1.23为常见的结构和材料类型。

晶体管的材料和结构互为独立，理论上可自由组合。比如可以设计在硅基的 BJT 器件，也可以设计在砷化镓基的 BJT 器件。

器件	结构分类		元素分类	材料	
BJT	BJT 双极型晶体管	晶体管 Transistor	四族元素	Si	硅
HBT				Ge	锗
JFET	FET 场效应晶体管		四四族化合物	SiGe	锗硅
MESFET				SiC	碳化硅
HEMT			三五族化合物	GaAs	砷化镓
pHEMT				GaN	氮化镓
MOSFET				InP	磷化铟
CMOS			二六族化合物	ZnO	氧化锌
SOI				ZnS	硫化锌
Fin-FET				CdS	硫化镉

图 1.23: 晶体管的分类

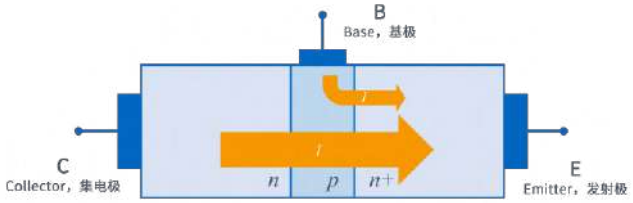


图 1.24: BJT 器件的基本结构（npn 管为例）

对一个器件的准确描述，应该将材料与器件种类同时说明，比如对于 5G 手机射频 PA 中使用的 HBT 器件，准确名称应该是 GaAs HBT 器件。不过由于大家平时工作中的约定俗成，叫法经常加以简化。

比如在手机射频领域，大家一般用 HBT，或者 GaAs 来简称 GaAs HBT；在电源控制领域，大家用 SiC 来简称 SiC MOSFET，用 GaN 来简称 GaN FET 类器件。这种简称在某个细分行业领域是有效的，但在跨出本行业交流时可能会引起误解，必要时需要加以注意。

1.2.2 以结构分类半导体

从结构上区分，半导体器件主要分为 BJT 和 FET 两种类型。

BJT 的全称是是 Bipolar Junction Transistor，双极型晶体管。FET 的全称是 Field Effect Transistor，场效应晶体管。二者都可以实现晶体管的放大特性。

1.2.2.1 BJT 器件

基本结构

BJT 器件由两个背靠背的 p-n 结构成。由于用于电流传输的 p-n 结包含电子与空穴两种载流子，所以 BJT 器件取名为“Bipolar（双向/双极）”Junction Transistor。其基本结构如图1.24所示。

虽然 BJT 由两个 p-n 结构成，但不是任意两个背靠背的 p-n 结二极管都可以构成 BJT。BJT 器件对于各区的掺杂浓度以及厚度有着精确的要求。BJT 器件中发射极需要重掺杂，基极需要较重掺杂，并且宽度极薄，以使大量的电子与空穴可以穿越。

BJT 器件制备

在集成电路制备中，BJT 器件先通过在 n 型外延中形成 p 型扩散，形成基极；再进行 n 型扩散，形成发射极，从而实现紧邻的两个 p-n 结。

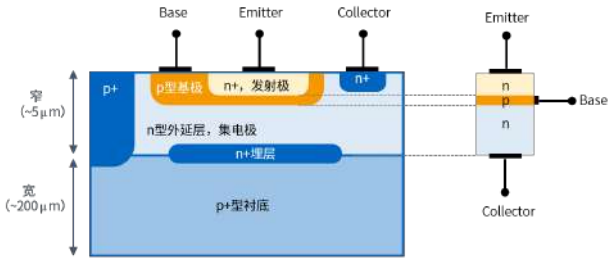


图 1.25: BJT 器件截面和简化模型

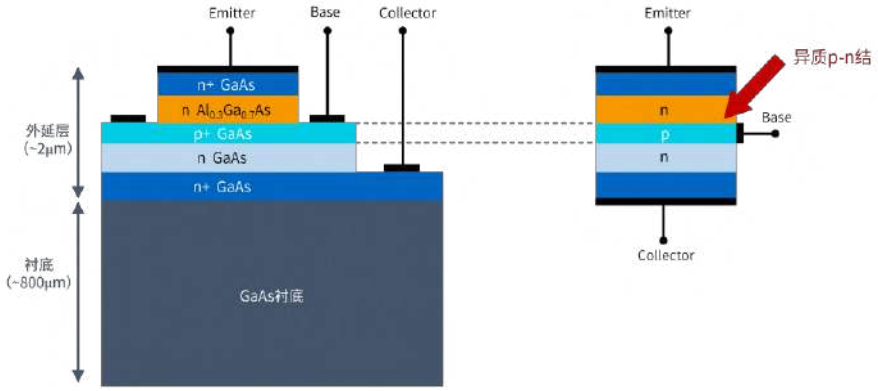


图 1.26: GaAs HBT 器件截面和简化模型

通过图1.25可以看到，在 BJT 器件中，电流在垂直方向进行流动，所以集成电路中的 BJT 器件是垂直器件。

HBT：一种特殊的 BJT 器件

HBT 的全称是 Hetero Junction Bipolar Transistor，中文名为异质结双极型晶体管。HBT 是一种特殊的 BJT 器件。

HBT 对普通 BJT 的改进是在发射极和基极之间采用不同的半导体材料，形成异质 p-n 结，来抵挡住基极载流子向发射极的注入，这样就可以使发射极中更多的载流子流入集电区，从而提高集电极到发射极之间的电流。外围观察到的现象是基极的电流变小了，集电极电流变大了，基极电流对集电极电流有了更强的控制能力。

常见的 HBT 器件有 GaAs HBT 器件、SiGe HBT 器件。图1.26为典型的 GaAs HBT 器件横截面图。

AlGaAs/GaAs 是研究最广泛，应用最广泛的异质结系统。5G 射频 PA 中所使用的 GaAs HBT，就是此种类型的异质结 HBT。在 AlGaAs/GaAs 异质结 HBT 中，基极 p+ GaAs 层与发射极 n+ GaAs 中间，注入了 n 型 Al0.3Ga0.7As 发射区薄膜，从而形成了基极与发射极之间的异质 p-n 结。

外延层：HBT 器件的关键材料

在 HBT 器件的结构中，可以看到 HBT 器件截面由外延层、衬底两部分构成。

在集成电路中，晶圆制备包含衬底制备和外延工艺两大环节。衬底（Substrate）是由半导体单晶材料制造的晶圆原片，衬底可以直接进入晶圆制造环节生产半导体器件，也可以进入外延工艺，生长完外延层后，再进行半导体器件制造。

外延（Epitaxy，简称 Epi）是指在单晶衬底上生长一层新单晶的过程，由于是在衬底上延伸生长，所以被称为“外延”。外延层一般只有几微米厚，外延出来的新单晶可以和衬底是同一材料，也可以是不同材料。

外延工艺解决了器件只能使用一种衬底材料的问题。使器件不同区域使用不同材料成为可能，极大地增加了器件设计的灵活性。对于 HBT 器件，由于需要设计 AlGaAs/GaAs 的异质半导体 p-n 结，就需要使用外延层工艺将不同材料及掺杂的半导体材料层设计出来。

通过不同外延层的设计，还可以对 HBT 器件的特性进行调整。HBT 器件的性能依赖于器件中发射极、基极以及集电极的厚度及掺杂浓度曲线，这些数据都是在外延层的设计中进行调整。

在 HBT 生产产业链运行中，一般由衬底厂商生产出 GaAs 衬底，再交由外延厂商生长外延层，最后交由代

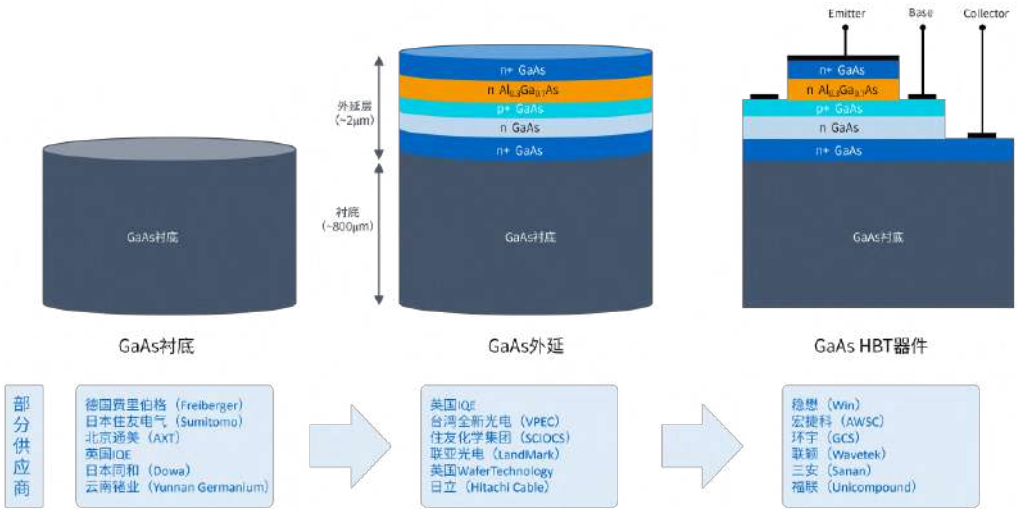


图 1.27: GaAs HBT 器件产业链流程

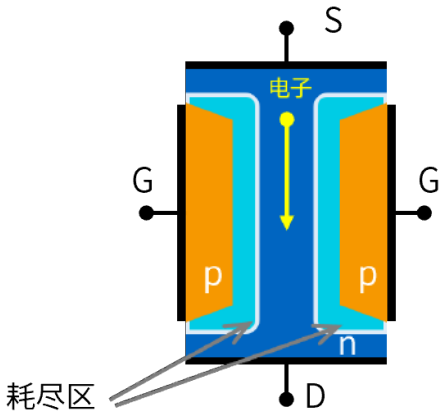


图 1.28: FET 器件的原理示意

工厂生产出 HBT 器件。由于外延层材料中半导体器件的材料参数已确定，所以外延层的生长是 HBT 器件生产的关键步骤。图1.27为 GaAs HBT 器件产业链流程。

1.2.2.2 FET 器件

FET 器件的全称是 Filed Effect Transistor，中文名是场效应晶体管。晶体管名称前面的 Filed Effect（场效应）指的是用电场（Electric Filed）来控制半导体内电流流动的器件。

FET 的思路非常简单直接，在一个经过掺杂的半导体材料（如 n 型）两边，加上另一种掺杂的半导体材料（p 型），通过控制 p 型掺杂上的电压，就可以控制 n 型沟道的夹断或者导通，就达到了控制电流的目的。FET 的工作原理，更加像水龙头了。

在 FET 器件中，由于参与导电的只有一种载流子（如 n 型沟道 FET 器件中的电子），这与 BJT 器件有电子与空穴均参与导电有很大的不同。所以最早的 FET 器件又叫 Unipolar Transistor（单极晶体管），以强调这种单一载流子导电特性。

在 n 沟道 FET 器件中，有三个极，分别是提供电子的源极（Source，S 极），流出电子的漏极（Drain，D 极），以及控制沟道的栅极（Gate，G 极）。图1.28为 FET 器件的原理示意。

相比于 BJT 器件复杂的 p-n 结运行理论，FET 器件的概念简单直接，“沟道”和“夹断”，非常符合人们的直觉，所以 FET 的理念在 BJT 商用之前就被提出来。

在 1926 年 J. E. Lilenfeld 申报的专利文件中，最早提出了 FET 的理念 [1]。比 1947 年贝尔实验室的肖克利团

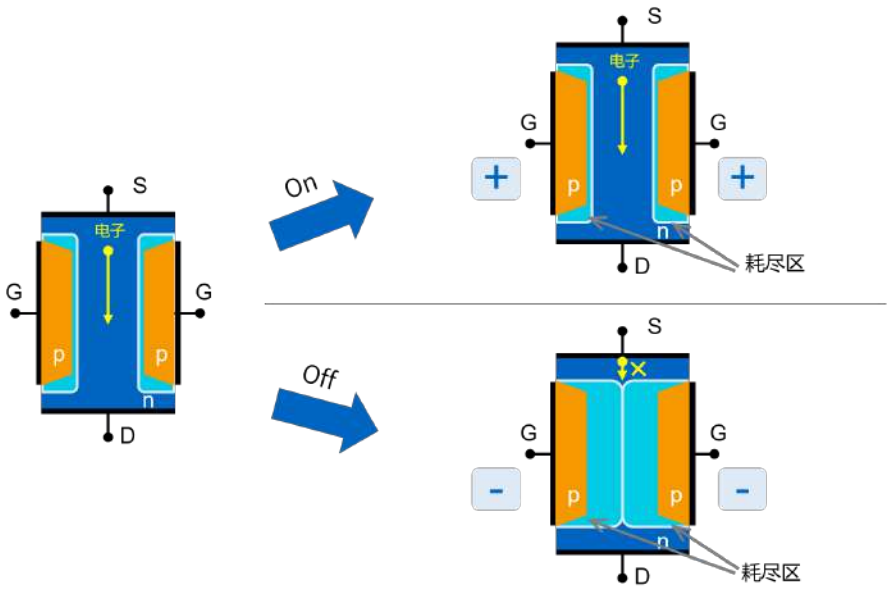


图 1.29: JFET 的构造及工作原理

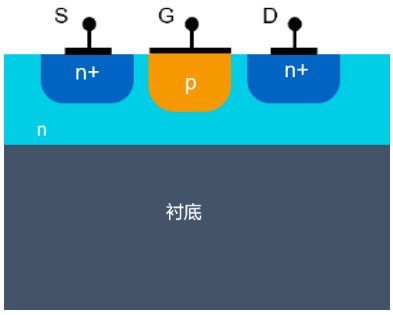


图 1.30: 现代的外延层 JFET 器件结构

队发明 BJT 器件还要早 21 年。但由于当时工艺所限，FET 器件只存在了前期科学家们脑中的概念之中。一直到 1953 年，得益于工艺进步，FET 器件才被真正生产出来。第一个被生产出来的 FET 器件是 JFET（Junction FET）器件。

- JFET

  JFET 是首个被正式生产出来的 FET 器件。因为名称中有“Junction，结”，所以 JFET 的工作中也利用到了 p-n 结特性。JFET 的构造及工作原理如图1.29所示。

  JFET 由肖克利团队于 1952 年首次提出并加以分析。在 JFET 中，所加的栅电压改变了 p-n 结耗尽层宽度，进而改变了源、漏极之间的电导。

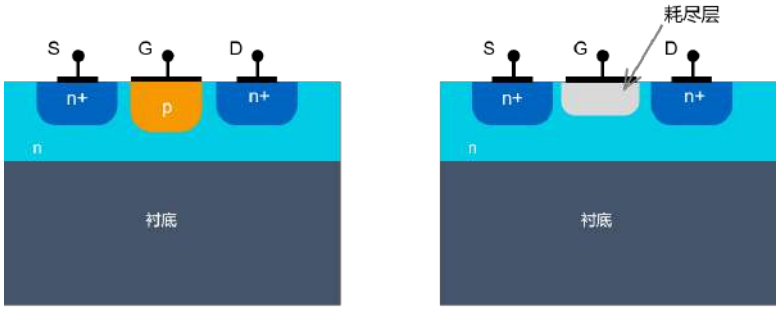
  经过多年工艺进步，JFET 的结构也有了变化，图1.30为现代的 JFET 结构，虽然在物理外观与最早的 JFET 有些不同，但仍然是利用电场控制栅极 p-n 结，等效于最早的肖克利结构。

- MESFET

  MESFET 的是 Metal-Semiconductor FET 的缩写，中文名是金属-半导体接触场效应晶体管。MESFET 是利用金属与半导体接触的特性开发的晶体管。JFET 与 MESFET 的结构比较如图1.31所示。

  MESFET 的工作原理与 JFET 类似，唯一不同点是用于控制沟道夹断与否的并不再是 p-n 结，而是金属-半导体结（简称金-半结）。

  金属-半导体接触是在半导体理论研究中非常重要的研究，在许多半导体器件中都有广泛的应用。金属-半导体接触最早始于 1874 年半导体物理理论的建立时期，1938 年，Schottky（肖特基）完成了金属-半导体势垒理论，完成了理论体系。于是，形成肖特基势垒的金属-半导体接触又叫肖特基接触，由此开发的二极



(a) JFET器件　　(b) MESFET

图 1.31: JFET 与 MESFET 的结构对比

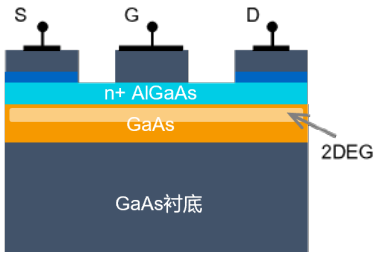


图 1.32: HEMT 器件的典型结构

管叫肖特基二极管。除了肖特基接触外，金属与重掺杂的半导体接触还可以形成欧姆接触，这是所有半导体器件流入和流出所必须的。

商用产品中被广泛使用的 MESFET 器件是 GaAs 基 MESFET，得益于 GaAs 的电子输运特性，GaAs MESFET 有良好的射频性能，是现在单片微波集成电路（MMIC）的核心。

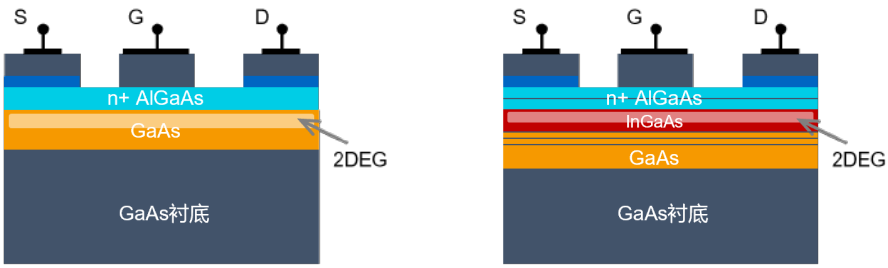
- HEMT 及 pHEMT

HEMT 的全称是 High Electron Mobility Transistor（高电子迁移率晶体管）。HEMT 最早由日本 Fujitsu 公司于 1979 年发明。HEMT 器件的想法是，利用不同半导体材料异质结的特性，在接合面上聚集大量的电子，形成一种名叫“二维电子气（2DEG）”的高移动层，达到更好的器件性能。HEMT 器件在高频毫米波领域、低噪声领域有着不可替代的应用。

因为利用到了异质结，所以 HEMT 又被称为异质结 FET （Heterostructure FET，HFET）。图1.32为 HEMT 器件的典型架构。

通常情况下，异质结接触表面会存在晶格失配，这个失配会影响到器件性能，也影响到更大带隙电压材料的选取。于是就有了一种改进型的 HEMT 器件：pHEMT。

pHEMT 器件的全称是 Pseudomorphic HEMT。Psedomorphic 的意思是假的、赝配的，pHEMT 器件在异质结转换时加入薄的“赝晶层”，用于将两边的晶格拉向匹配。图1.33为 HEMT 与 pHEMT 器件结构示意图。



(a) HEMT器件　　(b) pHEMT器件

图 1.33: HEMT 与 pHEMT 器件结构示意图

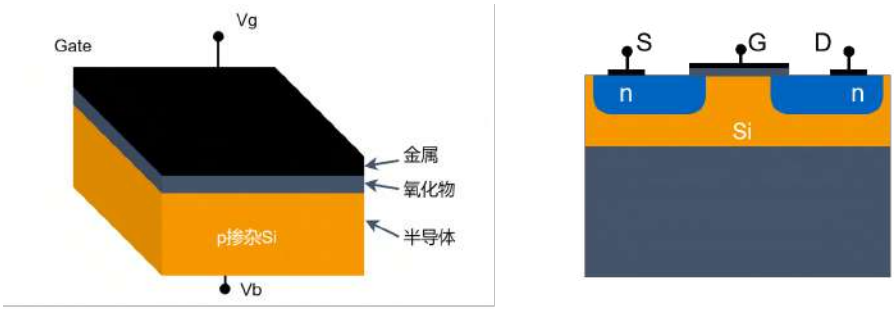


图 1.34: MOS 结构及 MOSFET 器件结构

由于出色的射频性能，pHEMT 器件在高性能射频微波领域应用广泛。尤其是 GaAs pHEMT 器件，是微波低噪声放大器、微波毫米波电路的重要半导体工艺。

与 HBT 器件相同，HEMT 及 pHEMT 器件的特性强烈依赖于材料特性，在 HEMT/pHEMT 产业链中，同样需要外延层厂商生产出相应的材料外延，再由代工厂进行器件加工。

- MOSFET

  MOSFET 结构毫无疑问是当今集成电路领域最为核心的结构。

  MOSFET 缩写自 Metal-Oxide-Semiconductor FET（金属-氧化物-半导体 FET），MOSFET 是 MIS（Metal-Insulator-Semiconductor，金属-绝缘层-半导体）器件的一种特殊结构。图1.34为典型 MOS 结构，以及 MOSFET 的示意图。

  在 MOSFET 之前，已经有了平面结构的 MESFET 器件与垂直结构的 BJT 器件，但这两种电路都无法适应于大规模电路的设计：

  - 垂直结构的 BJT 器件无法做到有效集成；BJT 器件的工作机理使得器件无法有效关断或打开；Base 端的电流也使 BJT 器件有较大的工作电流
  - MESFET 虽然是平面器件，解决了集成的问题，但 MESFET 器件所使用的金属-半导体结会钳位住过大 Gate 电压，造成 Vg 只能在一定范围内使用

  于是，MOSFET 器件被发明了出来。

  现代 MOSFET 器件的提出在 1960 年，Ligenza 等人提出了基于 Si-SiO2 结构的 MOSFET 器件，改变了 MESFET 的一系列缺点，适合大规模集成，适合尺寸等比例缩小。自此，半导体器件再也不是一个个单独的晶体管，大规模集成电路的序幕就此拉开。

  基于 MOSFET 器件的半导体结构多种多样，主要有 CMOS、SOI、Fin-FET 等。

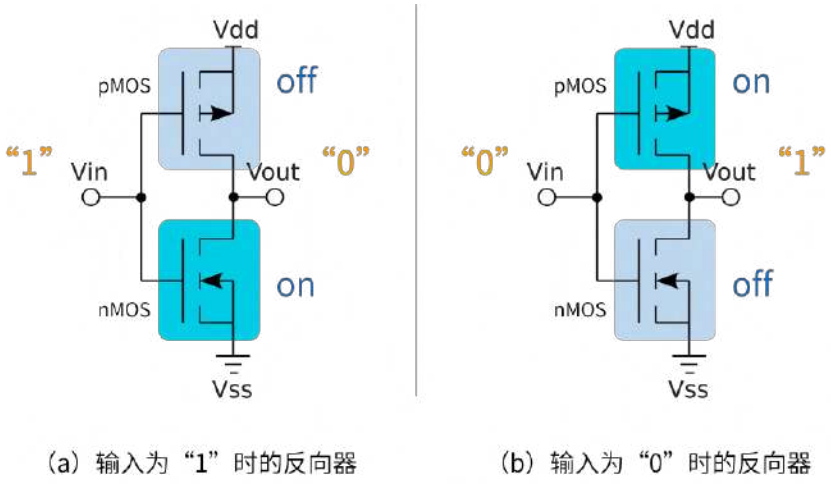
- CMOS

  CMOS 电路的发明，使集成电路进入到了“等比缩小”的快车道，是人类集成电路史上的一大飞跃。

  CMOS 的全称是 Complementary metal-oxide-semiconductor（互补型 MOS），“互补”的意思是电路中不止由 nMOS，还有 pMOS。两种 MOS 器件结合进行设计，就可以完美完成数字集成电路中的逻辑电路设计。虽然也可以用 BJT 器件设计类似 TTL（Transistor-Transistor Logic），但 CMOS 逻辑的优点在于静态功耗低。如图1.35简单反向器为例，在同一时刻，通路中的 nMOS 与 pMOS 只有一个器件导通，理论电路静态功耗为 0。

  CMOS 器件在 1963 年于仙童公司被发明出来 [3]，随后被应用于存储、处理器等数字电路设计中。图1.36为 CMOS 器件的基本结构。

  CMOS 器件被发明之后，人们发现这项技术与数字电路技术相得益彰、珠联璧合。数字电路需要的是小尺寸、低功耗的基本逻辑单元，而 CMOS 器件刚好可以满足这个需求。于是，人类不断改进半导体工艺，使 CMOS 器件尺寸不断缩小，以使数字电路的能力越来越强大。



（a）输入为“1”时的反向器　　（b）输入为“0”时的反向器

图 1.35: CMOS 反向器工作原理说明

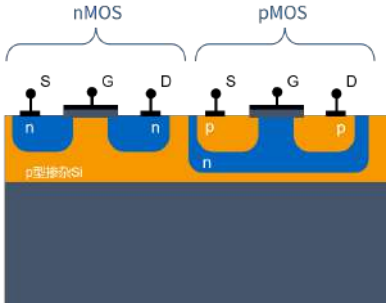


图 1.36: CMOS 器件的基本结构

1965 年，仙童公司的 Gordon Moore（戈登·摩尔）在 Electronic Magazine 杂志中发表了一篇简短的文章，预测 1965 年到 1975 年期间：集成电路上可以容纳的晶体管数目在大约每经过 18 个月到 24 个月便会增加一倍，如图1.37所示。换言之，处理器的性能大约每两年翻一倍，同时价格下降为之前的一半 [4]。

让摩尔有信以做出这种预测的，正是因为对 CMOS 技术的期待。

不过，让戈登·摩尔没想到的是，摩尔定律有效性绝不是在他预期的 1965-1975 年的 10 年之内，而是支撑整个集成电路行业发展了近 60 年。在这 60 年里，处理器上的晶体管数目从几百个增加到了几十亿甚至几百亿个，CMOS 器件特征尺寸由几毫米缩小到了几纳米。如图1.38，在高通发布的 8 系列智能手机平台芯片里，采用 4nm 工艺，集成了近 200 亿个晶体管。小小的集成电路，有了无与伦比的强大功能。

- SOI

  SOI 的全称是 Silicon on Insulator，绝缘体上硅。是指在带有绝缘层的硅衬底上生产半导体器件的技术。SOI 技术可以减小衬底损耗对器件的影响，因此有较好的器件特性。SOI 技术生产的半导体器件如图1.39所示。

- Fin-FET

  Fin-FET 中文名称是鳍式场效应晶体管，根据其形状像鱼鳍来命名。是在摩尔定律的推动下，为使器件尺寸不断减小，所发明的一种器件结构。

  在 Fin-FET 发明之前，随着摩尔定律使器件尺寸不断减小，原来平面结构的栅极已经没办法控制住器件的沟道，器件漏电流明显。于是，加州大学伯克利分校的胡正明教授于 1999 年提出 Fin-FET 结构，这种结构将栅极由平面结构升级成鱼鳍状的三维结构，于是形成了更好的对沟道的控制，其结构如图1.40所示。

  目前，TSMC、Intel、三星等均采用 Fin-FET 进行先进 CMOS 工艺的设计。

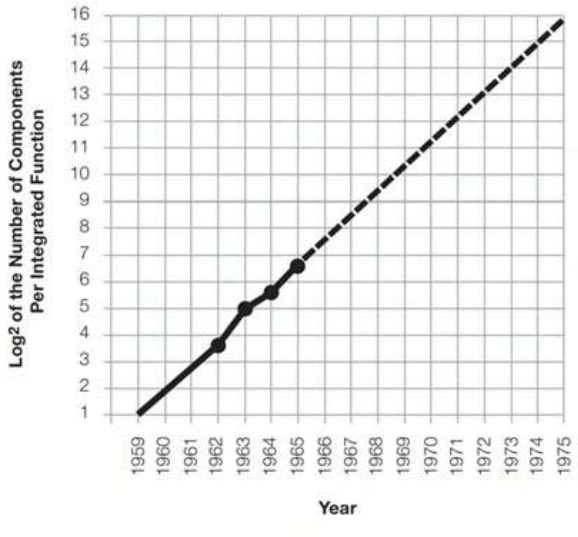


图 1.37: 戈登 • 摩尔在 1965 年提出的“摩尔定律“



图 1.38: 高通骁龙 8 系列手机平台芯片，采用 TSMC 4nm 工艺设计

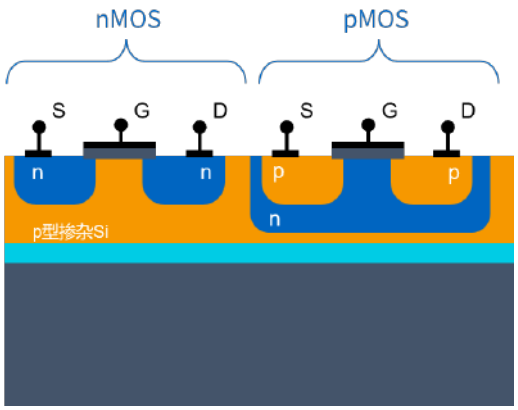


图 1.39: SOI 器件基本结构

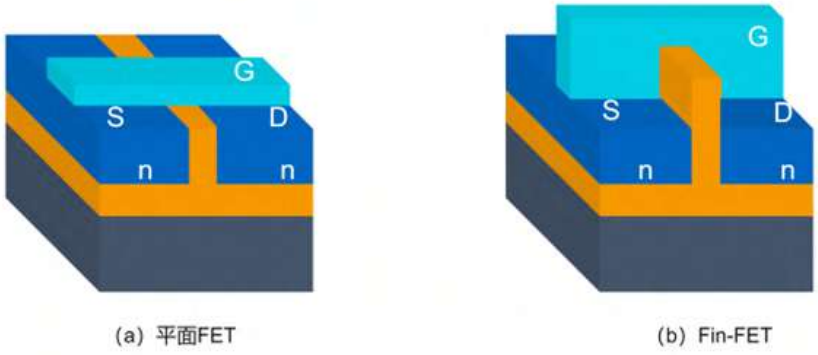


图 1.40: 从平面 CMOS FET，至 Fin-FET

II	III	IV	V	VI
4 铍 Be	5 硼 B	6 碳 C	7 氮 N	8 氧 O
12 镁 Mg	13 铝 Al	14 硅 Si	15 磷 P	16 硫 S
30 锌 Zn	31 镓 Ga	32 锗 Ge	33 砷 As	34 硒 Se
48 镉 Cd	49 铟 In	50 锡 Sn	51 锑 Sb	52 碲 Te
80 汞 Hg	81 铊 Tl	82 铅 Pb	83 铋 Bi	84 钋 Po

图 1.41: 用于半导体的常见材料元素

1.2.3 以材料分类半导体

除了根据器件的不同分类，半导体还可以根据材料的不同来分类。

半导体材料种类非常多，用于制作半导体的材料不仅有我们熟悉的硅材料，还有锗。另外，一些化合物材料如 GaAs、GaN、SiC 等，也都是常见的半导体材料。下表列出了在元素周期表中处于 II-VI 族的部分元素，这些元素或者这些元素的集合可以构成典型的半导体材料。图1.41为用于半导体的常见材料元素。

由以上元素构成的常见半导体材料如图1.42所示。

需要说明的是，用于半导体的材料很多，甚至在一个半导体器件中，就会有不同的半导体材料。在对一个器件命名时，一般以构成该半导体器件重要部分（如沟道）的材料对其命名。

比如对于图1.44所示 GaN HEMT 器件，虽然器件可以生长在 Si 或者 SiC 衬底上，但是器件的沟道是由 GaN，以及 GaN 和 AlGaN 的异质结构成，所以器件称为 GaN HEMT 器件，或者 AlGaN/GaN HEMT 器件。

一般元素类	IV族化合物	III-V族化合物	II-VI族化合物
Si Ge	SiC SiGe	GaAs GaN InP InAs AlSb GaSb InSb	ZnS ZnSe ZnTe CdS CdSe CdTe

图 1.42: 常见半导体材料

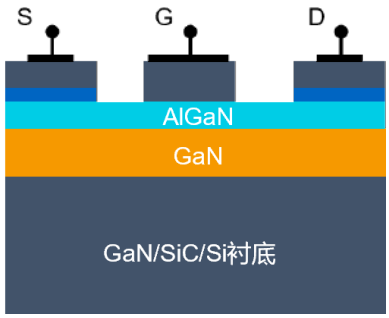


图 1.43: GaN/SiC/Si 基 GaN HEMT 器件示意图

1.2.3.1 材料与结构的结合

如文章开头所述，材料与器件结构结合，才可以对器件进行准确描述。完整的器件名称应为“材料 + 器件结构”。

理论上，不同材料和器件结构之间可以任意两两结合。不过，由于材料特性和特殊需求的关系，并不是所有材料上都可以生长出任意类型的器件。在使用中，主要的材料与结构结合器件如图1.44所示。

1.2.4 5G 射频中的半导体器件

在射频前端芯片设计中，也需要用到多种半导体工艺。射频前端各模块常用的半导体工艺及工艺特点总结如图1.45所示。

1.2.5 总结

66 年前，世界上第一个半导体晶体管在贝尔实验室诞生，此后，随着半导体工艺发展、集成电路的发明，小小的半导体器件改造了人类生活的方方面面。

射频技术也深受半导体技术的影响。无论是手机终端中所使用的模组芯片，还是基站中所使用的半导体单管，半导体器件作为核心元器件，成为当今射频技术的基础。

不过半导体技术抽象、半导体器件纷杂，很多时候没有办法直接和日常使用结合起来，这给大家平时的工作造成非常多的困扰。

本文将射频各模块中常见到的半导体器件类型做了整理，希望可以在需要的时候给大家工作带来帮助。

本文整理过程中，得到众多专家的指导，在此表示感谢

参考文献

1. Lilienfeld, J.E. ”Method and apparatus for controlling electric current” US Patent no. 1,745,175 (filed: 8 October 1926 ; issued: 28 January 1930).
2. Ligenza J R, Spitzer W G. The mechanisms for silicon oxidation in steam and oxygen[J]. Journal of Physics and Chemistry of Solids, 1960,14:131-136.
3. ”1963: Complementary MOS Circuit Configuration is Invented”. Computer History Museum. Retrieved 6 July 2019.
4. Moore, Gordon E. (1965-04-19). ”Cramming more components onto integrated circuits” (PDF). intel.com. Electronics Magazine. Retrieved April 1, 2020.
5. S. M. Sze, Physics of Semiconductor Devices.

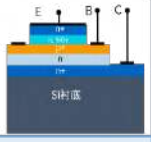
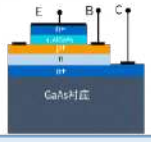
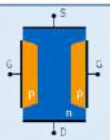
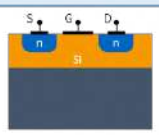
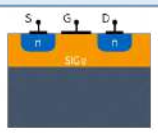
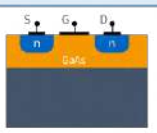
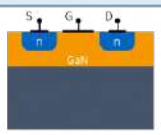
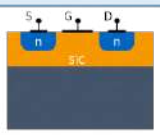
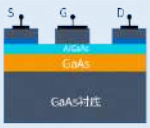
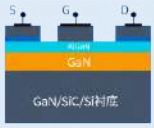
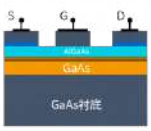
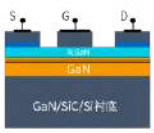
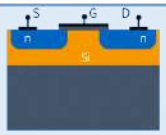
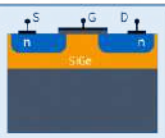
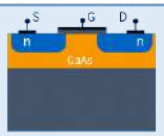
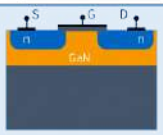
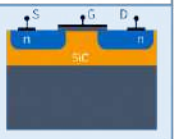
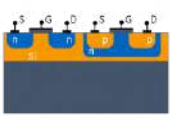
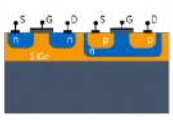
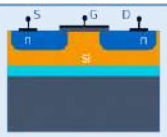
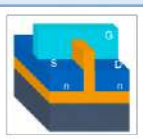
		Si	SiGe	GaAs	GaN	SiC
BJT	BJT		-	-	-	-
	HBT	-			-	-
FET	JFET		-	-	-	-
	MESFET					
	HEMT	-	-			-
	pHEMT	-	-			-
	MOSFET					
	CMOS			-	-	-
	SOI		-	-	-	-
	Fin-FET		-	-	-	-

图 1.44: 不同材料和结构的半导体器件

	选用工艺	工艺特点	应用场景
功率放大器	GaAs HBT	• 功率能力强 • 价格适中	终端射频PA
	SOI	• 射频性能较好 • 可做控制电路 • 价格较低	中功率线性PA（<25dBm）
	SiGe HBT	• 射频性能良好 • 可做LNA，实现PA/LNA单芯片 • 价格中低	部分终端射频PA
	Si CMOS	• 价格最低 • 射频性能较差	低成本PA
	GaN FET	• 功率能力极强 • 需要高压供电 • 价格高	基站大功率PA
	GaAs pHEMT	• 高频性能好 • 价格高	微波毫米波PA
低噪声放大器	SOI	• 较低噪声系数 • 可做控制电路 • 价格较低	广泛的LNA应用
	SiGe HBT	• 较低噪声系数 • 可做PA，实现LNA/PA芯片 • 价格中低	部分终端射频LNA
	GaAs pHEMT	• 低噪声系统 • 可应用于微波毫米波频段 • 价格高	微波毫米波LNA
开关	SOI	• 较好射频性能 • 可做控制电路 • 可用叠管设计高功率耐受开关 • 成本较低	终端射频开关
	GaAs pHEMT	• 射频性能好 • 可应用于微波毫米波频段 • 单管功率能力强 • 价格高	微波毫米波开关

图 **1.45:** 5G 射频中的半导体器件

在线阅读

本文摘自慧智微微信公众号，扫描以下二维码，在线阅读



1.3 收音机为什么叫“半导体”?

作者：彭洋洋 博士

在我国不少地方，“半导体”和收音机，说的是一个东西。

半导体（Semi-conductor）不是集成电路（IC，Integrated Circuits）所用的材料吗？为什么收音机会叫“半导体”呢？

1.3.1 收音机的发明

1.3.1.1 电报机的发明

早在 1835 年，摩尔斯在美国发明了最早的电报机，摩尔斯用电流的“通”“断”和“长短”来代替文字进行文字的电传送，这种编码也是鼎鼎大名的摩尔斯电码。虽然电报机可以实现远地相互通信，但还是必须依赖“导线”来连接，非常不方便。

1.3.1.2 无线电的发明

1887 年，德国物理学家赫兹，用实验证明了电磁波的存在。电磁波的发现为无线传输找到了传播途径，也开启了无线通信的新纪元。为了纪念赫兹的伟大贡献，电磁波频率的国际制单位也以赫兹的名字来命名。

在随后的几十年时间里，电磁波被应用到无线通信中：

- 1894 年，意大利电气工程师马可尼开始进行实验，并于 1899 年实现跨越英吉利海峡传输。
- 1906 年，美国 33 岁发明家德福雷斯特设计了历史上第一个收音机。这个收音机将真空电子管放大器与马可尼的无线电发明相结合，利用无线电传送人的声音。
- 1919 年，被誉为美国广播通讯业之父的大卫·萨尔诺夫，参与创立美国无线电公司（Radio Corporation of America, RCA），RCA 是商用无线电和广播电视的起源。而 RCA 公司的收音机产品也成为当时最为流行的无线接收设备。

1.3.2 真空电子管收音机

在早期的收音机中，采用真空电子管进行收音机的设计。

1.3.2.1 真空电子管及工作原理

电子管（Electron Tube），又被称为真空管（Vacuum Tube ），是一种被广泛用于信号放大器设计的重要器件。真空电子管由外部的玻璃壳体、内部的电极及连接电极的管脚构成。由于为了有利于游离电子的流动，玻璃壳体内部被抽为真空，所以电子管也被称为真空管。图1.47为一些真空电子管的图片。

二极管是最简单的电子管器件。电子管二极管由阴极 Cathode 和阳极 Anode 构成。电子管工作时，位于阴极的灯丝上有电流通过就会发热，当阴极到达一定温度后，就会有电子获得足够的能量而从上面发射出来，这些电子将被在阳极的屏极（Plate）吸收。而由于在阴极的灯丝只能发射电子而不吸收电子，所以这个方向不能被反过来，这个单向导电特性使此电子管形成了二极管。电子管二极管原理如图1.48所示。

三极管是在二极管的阴极和阳极之间增加了一个栅极（Control grid），以控制电子的运动，而正是栅极的控制作用，使得电子管拥有了放大电压信号的能力。同时，还有四极管、五极管等，是在三极管的基础上增加第二栅极（帘栅极）以及第三栅极（抑制栅极）。如图1.49。

利用电子管形成的不同器件，就可以设计成不同的电路。采用电子管设计的放大器如图1.50所示。



图 1.47: 一些典型真空电子管图片

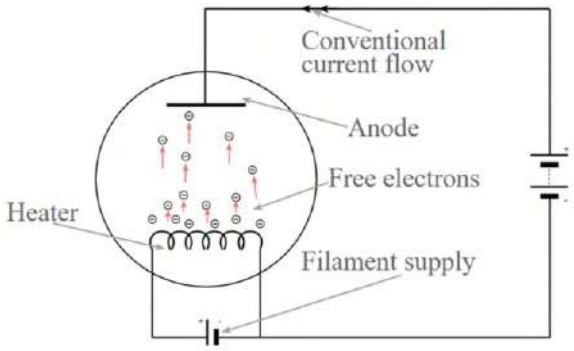


图 1.48: 电子管二极管原理

1.3.2.2 真空电子管收音机

利用真空电子管，可以设计实现检波、变频、放大等电路，通过以上电路的搭配，就可以实现收音机功能。采用电子管实现的收音机电路如图1.51所示。

20 世纪 20 到 30 年代，RCA 公司正是利用电子管，设计实现了如图1.52所示的电子管收音机，利用无线广播的方式将音乐、信息等无线传输到全球各个地方。

采用电子管可以实现无线通信和放大功能，并且有良好的大功率及线性度特性。但电子管的缺点也显而易见：体积大、功耗大、发热厉害、寿命短，并且需要高压电源。所以，电子管收音机只应用在政府、军事部门以及无线电爱好者中。

1.3.3 “半导体”的横空出世

1.3.3.1 半导体晶体管的发明

为了克服电子管的局限性，第二次世界大战结束后，贝尔实验室加紧了对固体电子器件的基础研究。肖克莱等人决定集中研究硅、锗等半导体材料，探讨用半导体材料制作放大器件的可能性。

1945 年秋天，贝尔实验室成立了以肖克莱为首的半导体研究小组，成员有布拉顿、巴丁等人。他们经过一系列的实验和观察，逐步认识到半导体中电流放大效应产生的原因。布拉顿发现，在锗片的底面接上电极，在另一面插上细针并通上电流，然后让另一根细针靠近及接触它，并通上微弱的电流，这样就会使原来的电流产生很大的变化。微弱电流少量的变化会对另外的电流产生很大的影响，这就是“放大”作用。这种“放大”作用，也使得半导体器件可以被用于制作放大器。

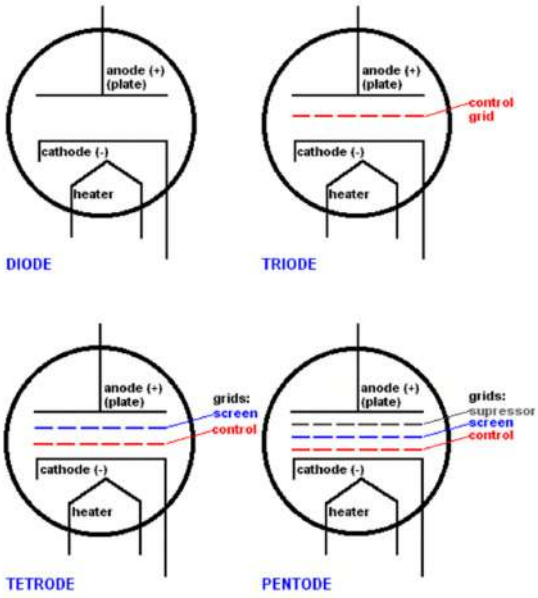


图 1.49: 电子管二极管、三极管、四极管、五极管结构



图 1.50: RCA 公司采用电子管设计的放大器



图 1.51: 采用电子管实现的收音机电路



图 1.52: RCA 公司设计的收音机型号 Model T10-1 Tombstone Radio (1935)

在为这种器件命名时，布拉顿想到它的电阻变换特性，**于是取名为 Trans-resistor(转换电阻)，后来缩写为 Transistor，中文译名就是晶体管**。这个小小的半导体晶体管，也成为日后集成电路发展的基础。

1956 年，肖克莱、巴丁、布拉顿三人，因发明半导体晶体管同时荣获诺贝尔物理学奖。如图1.53。

1.3.3.2 半导体晶体管对真空电子管的取代

由于晶体管尺寸小、所需电压低，所以晶体管出现之后，就开始对电子管进行取代。晶体管与电子管的对比如图1.54所示。

而半导体晶体管第一个实现商品化应用的产品，就是收音机。

1954 年，Regency 公司发布 TR-1，成为首个商用的半导体晶体管收音机（TR 即 Transistor，晶体管缩写）。半导体晶体管收音机的发明，使原来只能摆放在家中的笨重真空电子管收音机可以装进口袋，一时间“Pocket Radio”（可装进口袋的收音机）成为当时时尚之选。如图1.55及图1.56。

不过半导体晶体管收音机在刚刚推出来的前几年，依然只是一个小众产品，主流的收音机还是采用真空电子管实现。为了区分采用半导体晶体管制作的收音机和主流电子管制作的收音机之间的区别，**半导体晶体管收音机被直接称为“晶体管”或者“半导体”。这也是现代收音机被称之为“半导体”的原因。**

一个小小的半导体收音机，也改变了很多公司的命运。1957 年，东京通信工业公司发布 Sony 系列半导体收音机。之后的 Sony TR-63 是 20 世纪 50-60 年代最受欢迎的消费类电子产品之一。Sony TR-63 在当时实在太成功，也使得东京通信工业公司在 1958 年直接将公司更名为 Sony，Sony 也成为日后称霸 50 年的日本消费类电子产品供应商，公司也变得家喻户晓。图1.57为 Sony 公司推出的晶体管收音机。

以 Sony、Regency 公司产品为代表的便携晶体管收音机，改变了人们听取音乐的习惯，人们开始可以在任何时间和地点，都能听到音乐。人们因为便携收音机养成的随时随地欣赏音乐的习惯，也促使了 80 年代 Walkman、CD，乃至 21 世纪 iPod 的发明。

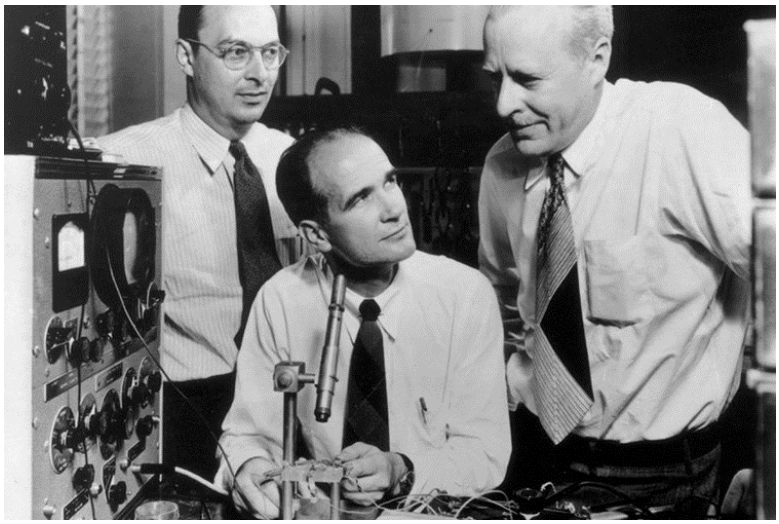


图 1.53: 半导体晶体管的发明者左起：布拉顿，肖克莱，巴丁

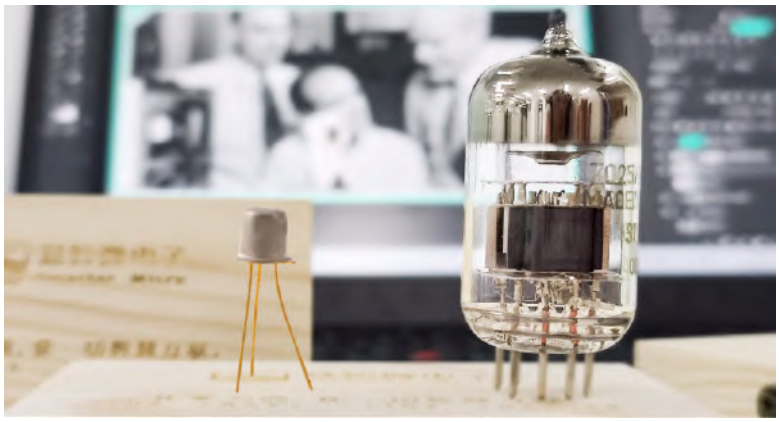


图 1.54: 采用半导体技术实现的晶体管（左）与采用真空管技术实现的电子管（右）对比

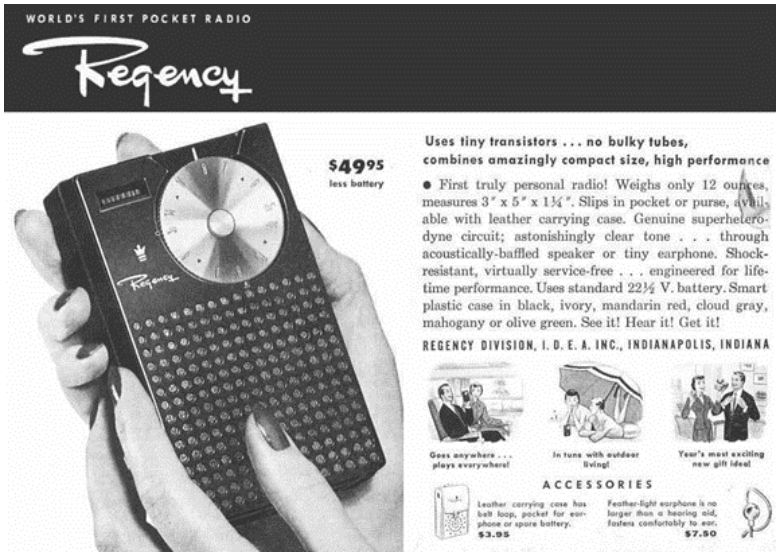


图 1.55: 1954 年，Regency 推出 TR-1，主打为：首个可装进口袋的收音机

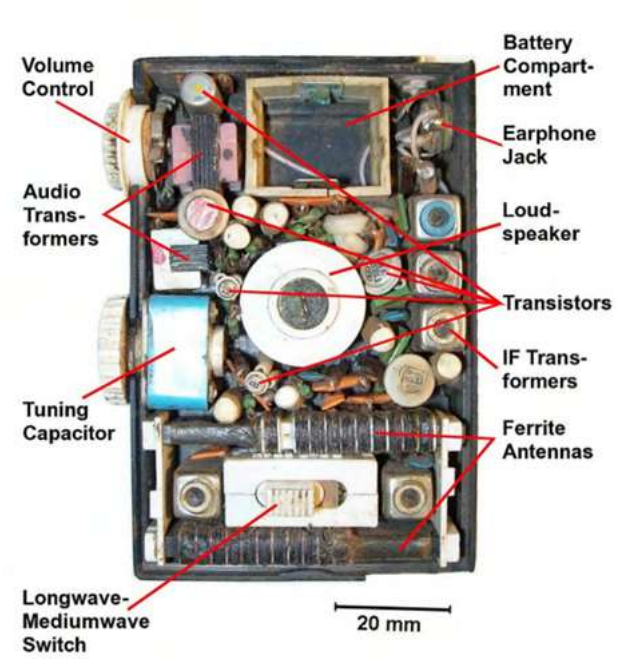


图 1.56: 典型的晶体管收音机内部构造主要核心器件为晶体管（Transistor）



图 1.57: Sony 在 20 世纪 50-60 年代推出的晶体管收音机晶体管（Transistor）被直接印在了壳体上



图 1.58: 上海无线电三厂在 1962 年生产的美多 28A-1 型半导体收音机（展出于上海无线电博物馆）

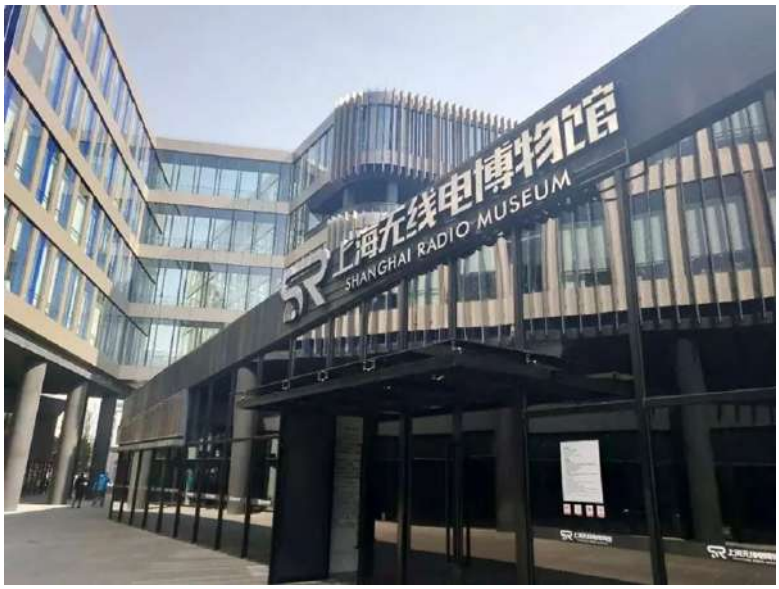


图 1.59: 上海无线电博物馆，位于上海市徐汇区田林路 200 号

1.3.4 中国自己的“半导体”收音机

虽然近年来中国在集成电路领域不断在追赶世界先进水平，但在上个世纪中期，中国无线通信的发展并不落后。

新中国成立以来，电子工业就是被高度重视的行业。1958 年，上海宏音无线电器材厂等 9 个工厂及上海无线电子技术研究所联合研制成功了我国第一台基于晶体管的收音机。此后，上海、北京、南京等地的一些无线电工厂，先后生产出“上海牌”、“春蕾”、“飞乐”、“红灯”等晶体管收音机。

在上海无线电博物馆，展出了大量中国人自己的收音机。包括经典的电子管收音系列：中国成立 10 周年献礼之作的上海牌 132 型电子管收音机，70 年代中国爆款的 711 型六灯两波段电子管收音机；也有珍贵的早期半导体晶体管收音机：中国第一台简单半导体收音机卫星牌 636 半导体收音机，生产于 1962 年的第一台全部采用国产元器件收音机–美多 28A 型便携式半导体晶体管收音机等，如图1.58。

上海无线电博物馆位于上海市徐汇区田林路 200 号 A5 栋，如图1.59及图1.60，是一家以无线电为主要展示内容的博物馆，由上海仪电（集团）有限公司投建，藏品约 2300 件。主要展品由沪上知名收藏家张明律先生捐赠，免费向公众开放。欢迎大家前往，一起感受无线通信发展给生活带来的改变。

人类对品质生活的需求为技术指明方向，技术的发展改变了人类的生活方式。集成电路、半导体的发明到现在还不到一个世纪，在这几十年中，世界发生了本质的变化，我们的生活方式也在不断受到影响。

你认为还有哪些技术的出现，改变了我们的生活方式呢？欢迎扫描二维码，在线留言区讨论。

注：部分图片来自网络。



图 1.60: 上海无线电博物馆部分展品

在线阅读

本文摘自慧智微微信公众号，扫描以下二维码，在线阅读

1.4 胡正明：晶体管的未来，是我们的未来

转自 IEEE

1945 年的贝尔实验室，肖克利、巴丁和布拉顿组成的研究小组，开始研究采用半导体工艺实现放大器件的可能性。

1947 年 12 月，三位科学家在贝尔实验室向世人展示了人类第一代的半导体放大器件，他们将这种器件取名为晶体管（Transistor）。至 2022 年的 75 年里，基于晶体管的集成电路不断发展，改变了整个人类社会。



图 1.62: 肖克利、巴丁和布拉顿团队；和他们发明的晶体管

在 2022 年，为了纪念晶体管被发明 75 周年，IEEE（电气与电子工程师协会）电子器件分会（Electron Device Society）组织活动，纪念晶体管被发明的 75 周年。



图 1.63: 2022 年，IEEE EDS 组织晶体管发明 75 周年纪念活动

在此活动中，Fin-FET 的发明者，著名半导体专家胡正明（Chenming Hu）教授发表题为《晶体管的未来，是我们的未来》文章，回顾了半导体晶体管的过去，展望了晶体管的未来。IEEE Spectrum 于 2022 年 11 月 30 日刊登了此篇文章。

下文为翻译的中文版本，英文原文请用以下链接阅读（https://spectrum.ieee.org/the-future-of-transistors）。

1.4.1 The Future of the Transistor Is Our Future，晶体管的未来，是我们的未来

作者：胡正明

在我为此奉献一生的晶体管发明 75 周年之际，我想回答两个问题：世界还需要更好的晶体管吗？如果需要的话，那是什么样的晶体管？

我的回答是：是的，我们需要新的晶体管。而且我认为我们今天已经有了一些关于它们会是什么样子的暗示。问题在于，我们是否有意愿和经济能力去制造它们。

我相信晶体管现在是，并将继续是应对全球变暖影响的关键。气候变化可能会给社会、经济和个人带来巨变，因此人类需要使我们更强大的工具，来应对这种变化。

和其他技术一样，半导体毫无疑问可以提高人类的能力。不过和其他技术不一样的是，大多技术受到自然资源和能源资源的限制，都无法获得使人类能力获得数量级的增长。但基于晶体管的技术是个独特的例外。原因是：

1. 随着晶体管的改进，使人类掌握了从未想象到的新能力，例如计算和高速通信、互联网、智能手机、内存和存储、机器人技术、人工智能。可以想象，未来还会有其他新技术涌现出来；
2. 这些能力的广泛的应用，改变了所有技术、工业和科学。同时半导体技术的增长不像其他技术那样受到其材料和能源使用的限制。IC 使用相对少量的材料就可以生产。并且，它们还在变得越来越小，使用的材料也越来越少，它们变得更快、更节能、更有能力；
3. 从理论上讲，信息处理所需的能量仍然可以减少到今天所需能量的千分之一以下。虽然我们可能还不知道如何达到这种理论效率，但我们知道理论上这是可行的。相比之下，大多数其他技术（例如电机和照明）的能源效率已经达到其理论极限。

1.4.1.1 晶体管：过去、现在和未来

我们清楚在短期内如何继续改进晶体管，但离今天越远的未来，晶体管技术就会变得越模糊。不过，我们可以通过查看晶体管的过去，来洞见晶体管的未来。

从 1960 年到 2010 年左右，基本的平面 (2D) MOSFET 结构一直保持不变，直到进一步增加晶体管密度以及降低器件功耗变得不可能。我在加州大学伯克利分校的实验室早在十多年前就看到了这一点。**我们在 1999 年报道了平面晶体管的后继者：Fin-FET 的发明**。Fin-FET 作为第一个 3D MOSFET，将扁平而宽的晶体管结构变为高而窄的晶体管结构。这样做的好处是在更小的占地面积内获得更好的性能，就像在拥挤的城市中多层建筑相对于单层建筑的优势一样。

Fin-FET 也被称为薄体（thin-body）MOSFET，Fin-FET 的设计概念继续指导新器件的开发。Fin-FET 源于这样一种认识，即电流不会通过硅表面几纳米内的晶体管泄漏，因为那里的表面电势受到栅极电压的良好控制。Fin-FET 牢记这种薄体概念。该器件的主体是垂直的硅鳍片，被氧化物绝缘体和栅极金属覆盖，在强栅极控制范围之外没有留下任何硅。Fin-FET 将漏电流降低了几个数量级，并降低了晶体管工作电压。它还指出了进一步改进的路径：进一步降低 body 厚度。

Fin-FET 的鳍片随着每个新的技术节点变得横向越来越薄、且高度越来越高。但这种进步现在变得难以维持。因此业界正在采用一种新的 3D 薄体 CMOS 结构，称为环栅 (GAA)。

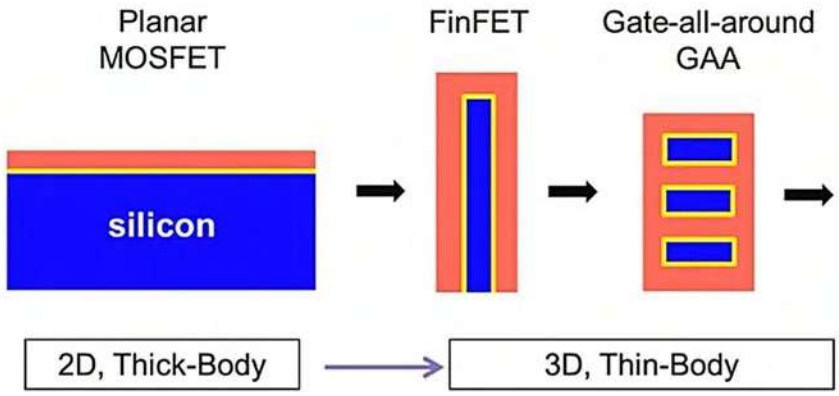


图 1.64: MOSFET 结构的每一次演变都旨在通过栅极 [粉红色] 更好地控制硅中的电荷。电介质 [黄色] 防止电荷从栅极移动到硅体 [蓝色]

3D 薄体趋势将从这些 3D 晶体管延续到 3D 堆叠晶体管、3D 单芯片（Monolithic）电路和多芯片封装。在某些情况下，这种 3D 趋势已经达到了顶峰。例如，电荷陷阱（charge-trap）存储器晶体管阵列的规律性使 NAND 闪存成为第一个从 2D 电路过渡到 3D 电路的 IC。自 2007 年东芝首次报告 3D NAND 以来，堆叠层数已从 4 层增长到超过 200 层。

单芯片 3D 逻辑 IC 可能会适度起步，堆叠 CMOS 反相器的两个晶体管以减少所有逻辑门的“占地面积”。但是堆栈的数量可能会增加。通往 3D IC 的其他途径可能采用将额外的半导体薄膜层（例如硅、硅锗或砷化铟镓）转移或沉积到硅晶片上。

薄体趋势可能会在二维半导体中达到其最终终点，其厚度以原子为单位。例如，二硫化钼分子（Molybdenum disulfide molecules）既天然又薄又相对较大，形成可能不超过三个原子宽但具有非常好的半导体特性的二维半导体。2016 年，加利福尼亚州和德克萨斯州的工程师使用二维半导体分子二硫化钼和碳纳米管薄膜展示了具有关键尺寸的 MOSFET：栅极长度仅为 1 纳米。即使栅极短至 1 nm，晶体管漏电流也仅为每毫米 10 毫微安，可与当今最好的生产晶体管相媲美。

可以想象，在遥远的未来，整个晶体管可能被预制为一个分子（molecule）。这些预制构件可能会通过称为定向自组装 (DSA：directed-self-assembly) 的制程被带到它们在 IC 中的精确位置。要理解 DSA，回顾一下 COVID 病毒使用其尖峰来寻找并化学停靠在特定人类细胞表面的确切位置可能会有所帮助。在 DSA 中，对接点（docking spots）、“尖峰”（spikes）和晶体管 cargo 都是经过精心设计和制造的。最初的对接点（docking spots）可以通过在基板上进行光刻来创建，但是额外的对接点（docking spots）可能会在后续步骤中作为 cargo 引入。如果仅在制造过程中需要而在最终产品中不需要，则可以通过加热或其他方式去除一些 cargo 。

除了使晶体管更小之外，我们还必须不断降低它们的功耗。在这里，我们可以看到通过使用所谓的负电容场效应晶体管 (NCFET：negative-capacitance field-effect transistors) 实现了数量级的减少。这些需要在 MOSFET 的栅极堆叠中插入纳米薄的铁电材料层，例如氧化铪锆（hafnium zirconium oxide）。由于铁电体包含自己的内部电场，因此打开或关闭器件所需的能量更少。薄铁电体的另一个优点是可以利用铁电体的容量来存储其电场状态，从而将存储和计算集成在同一器件中。

在某种程度上，我所描述的半导体器件是从现有趋势中产生的。但未来的晶体管可能与今天的晶体管具有截然不同的材料、结构和工作机制。例如，纳米机电开关（nanoelectromechanical switch）是过去几十年机械继电器的技术回归，而不是晶体管的技术延伸。这种机电开关不依赖于半导体物理学，而是仅使用金属、电介质以及施加不同电压的紧密间隔的导体之间的力。

虽然这些例子都在多年前用实验证明了可行性。然而，与以往的半导体技术突破相比，这些技术在批量生产中有更大的阻力。

1.4.1.2 走向未来

“晶体管技术的进步其实并不平坦和顺利。”

未来，我们能否继续实现壮举？过去的一些经验表明，我们可以。

第一个经验是：晶体管技术的进步并不平坦或顺利。

1980 年前后，每颗芯片的功耗上升到了令人痛苦的地步。采用 CMOS、取代 NMOS 和双极技术后，工作电压从 5 伏逐渐降低到 1 伏，由此带来了 30 年直截了当的进步。

2000 年至 2010 年间，功耗再次成为一个问题。按当时的趋势，研究人员预测每平方厘米 IC 产生的热量很快就会达到核反应堆堆芯的热量。此时，3D 薄体 Fin-FET 和多核处理器架构的采用避免了危机，迎来了又一个相对平稳的发展时期。

晶体管技术的发展史可谓一座座山峰攀登。只有当我们到达一座山顶时，我们才能看到远处的景色，并绘制出一条路线来攀登下一座更高更陡的山峰。

第二个经验是，半导体产业的核心力量：纳米加工技术，非常强大。历史证明，只要有足够的时间和经济激励，只要该想法不违反科学规律，该行业就能够将任何想法变为现实。

但该行业是否有足够的时间和经济激励，来让我们继续攀登更高更陡峭的山峰，不断提高人类的能力？

这是一个非常合理的问题：随着晶圆厂行业资源的增长，技术发展的山峰增长得更快，终有一天，没有一家晶圆厂公司能够登上山顶，看清前方的道路。那么会发生什么？

所有半导体晶圆厂（包括独立的和像英特尔这样的综合性公司）的收入约占半导体行业收入的三分之一。但晶圆厂仅占半导体技术所带来的 IT、电信和消费电子行业总收入的 2%。然而，晶圆厂行业承担着发现、生产和营销新晶体管和纳米制造技术的大部分重担。这种情况需要做出改变了。

在晶圆厂行业，为了让生存，晶圆厂中相对贫乏的资源必须优先考虑晶圆厂建设和股东需求，而不是科学探索。虽然晶圆厂正在延长其研究视线范围，但这也需要得到行业中的其他厂商支持。人类长期解决问题的能力，

值得所有公众有针对性的支持。

对于半导体行业，该行业需要长期探索性研究的帮助。比如由公共资助，或在类似贝尔实验室的环境中进行，或者由大学提供对研究人员支持，由此使得这些研究人员具有长期的职业生涯，并且在物理、化学、生物学和算法方面的知识比企业研究更广泛和更深入。

这样，人类将会继续寻找到新的晶体管，并获得新的能力。这些新的能力是人类应对未来几个世纪挑战时所必须具备的。

胡正明教授简介：



图 1.65: 胡正明教授

胡正明教授，1947 年 7 月出生于中国北京，微电子学家，美国国家工程院院士、中国科学院外籍院士，美国加州大学伯克利分校杰出讲座教授。

胡正明教授是微电子微型化物理及可靠性物理研究的一位重要开拓者，对半导体器件的开发及未来的微型化做出了重大贡献。

胡正明教授的主要科技成就为：

- 领导研究出 BSIM 模型，从实际 MOSFET 晶体管的复杂物理推演出数学模型，该数学模型于 1997 年被国际上 38 家大公司参与的晶体管模型理事会选为设计芯片的第一个且唯一的国际标准；
- 发明了在国际上极受注目的 FinFET 等多种新结构器件；对微电子器件可靠性物理研究贡献突出。

胡正明教授主要经历为：

- 1968 年，台湾大学电机系毕业；
- 1969 年，赴美国加州大学伯克利分校留学，先后获得硕士、博士学位；
- 1973 年，博士毕业后前往麻省理工学院任助理教授；
- 1976 年，回到加州大学伯克利分校电子工程与计算机科学系任教；
- 1996 年, 创办思略微电子有限公司兼任董事长；
- 1997 年，当选为美国国家工程院院士；
- 1999 年，开发出了 Fin-FET（鳍式场效应晶体管）；
- 2001 年至 2004 年，担任台积电首任技术执行长；
- 2004 年，回到加州大学伯克利分校任教；
- 2007 年，当选中国科学院外籍院士；
- 2008 年，担任闪迪公司董事；
- 2015 年获得美国国家技术创新奖；2016 年获得美国国家科学奖章；
- 2017 年加入台湾交通大学半导体学院

参考文献

1. https://www.ieee.org/ns/periodicals/EDS/EDS-JULY-2022-HTML/index.html
2. https://spectrum.ieee.org/the-future-of-transistors

在线阅读

本文摘自慧智微微信公众号，扫描以下二维码，在线阅读

1.5 一片晶圆有多大？可以切出多少芯片？

作者：小慧

芯片的制造过程可以分为前道工艺和后道工艺。前道是指晶圆制造厂的加工过程，在空白的硅片完成电路加工，出厂后依然是完整的圆形硅片。后道是指封装和测试过程，在封测厂中将圆形的硅片切割成单独的晶粒，完成外壳的封装，最后完成终端测试，出厂为芯片成品。

芯片产出数量受晶圆大小、晶粒尺寸以及良率三个因素影响。接下来我们将从这三个方面简单聊聊。

1.5.1 一片 Wafer 有多大？

晶圆，又称 wafer, 其大小通常以直径来表示，常见的尺寸有 1 英寸 (25 毫米)、2 英寸 (51 毫米)、3 英寸 (76 毫米)、4 英寸 (100 毫米)、4.9 英寸 (125 毫米)、150 毫米 (5.9 英寸，通常称为“6 英寸”)、200 毫米 (7.9 英寸，通常称为“8 英寸”)、300 毫米 (11.8 英寸，通常称为“12 英寸”) 和 450 毫米 (17.7 英寸) 。

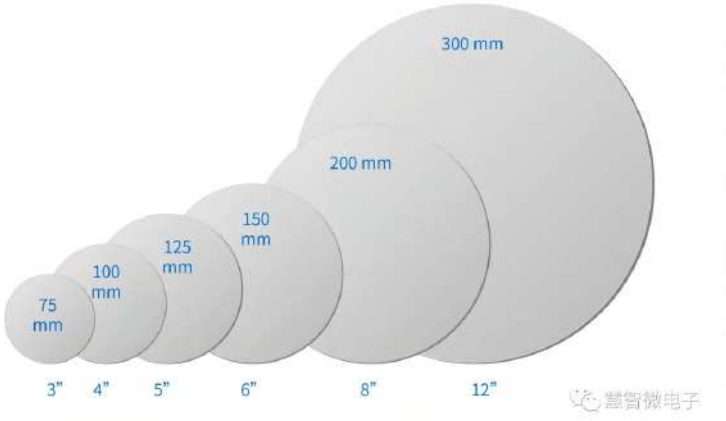


图 1.67: 晶圆尺寸

1.5.2 什么是“Die”？

Die 指的是芯片未封装前的晶粒，是从 wafer 上用激光切割而成的一个单独的晶圆区域，它包含了芯片的一个完整功能单元或一组相关功能单元。每个 Die 最终都切割成一个小方块并封装起来，成为我们常见的芯片。

1.5.3 晶圆可切割晶粒计算器 DPW

一片晶圆可以切出多少芯片？抛去良率因素，这是一个简单的图形面积计算问题，它的专有的表征名词是“DPW”，DPW 是 Die Per Wafer 的缩略词。晶圆可切割晶粒数 (DPW) 的计算是非常简单的。它的计算实际上是与圆周率 π 有密切的关联。

晶圆上的晶粒其实可以看作是圆形所能容纳下的所有方形的集合。所以，可切割晶粒数的计算就是利用圆周率和晶圆尺寸作为已知参数，确定出整体圆形区域能容量下的方形数量。

$$DPW = \frac{\pi \times (\frac{d}{2})^2}{W \times H} - \frac{\pi \times d}{\sqrt{2 \times W \times H}} \tag{1.1}$$

晶圆尺寸和晶粒尺寸虽然是已知的，但是，由于晶粒相互之间是有空隙（如预留的划道）的，晶圆的边缘去除区也不可用。这些因素使得计算变得稍微有点复杂和棘手。因此，把 DPW 工具的计算结果作为可切割晶粒估算值而非精确的计算值可能更准确一点。

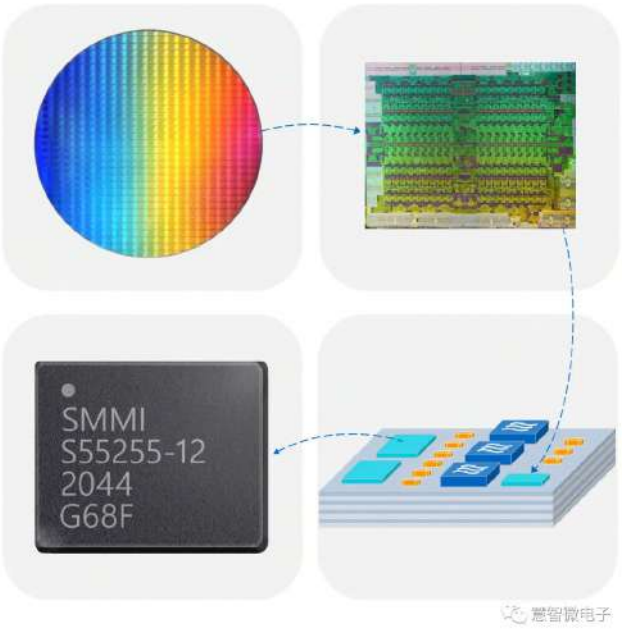


图 1.68: 芯片中“Die”的示意图

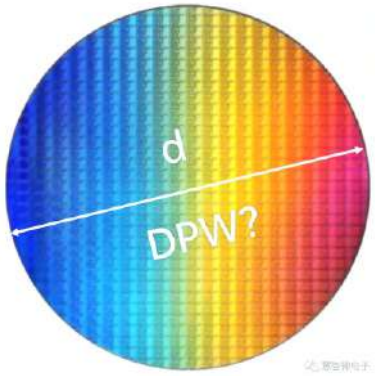


图 1.69: DPW 的计算

除了前面提到的无效区域外，晶圆厂还会额外占用部分区域做测试 (PCM 结构)，相对而言也会占用晶圆一小部分面积。另外还有划道、晶圆裕量，以及因为各工序之间或晶圆厂之间要求不同而导致的测试结构大小不一致而浪费的区域。因此，如需精确的最终 DPW 数值应直接向晶圆厂问询，以得到更专业准确的数据。

1.5.4 小程序计算工具

为方便大家计算芯片 DPW，慧智微开发了一款小程序，方便在线计算。有需要可以扫描正文小程序码使用。



图 1.70: 扫码添加小程序



图 1.71: DPW 计算页面

在线阅读

本文摘自慧智微微信公众号，扫描以下二维码，在线阅读



第二篇 射频芯片设计

本篇内容

- 射频收发机：从无线电到 5G
- 射频 PA 工艺选择
- 5G 射频 PA 的 Load-line 与 Load-pull
- 高效率 PA 设计：从 Class A 到 Class J
- 5G 射频 PA 的架构
- 高效率 PA：系统级设计
- 5G PA“记忆效应”的现象、形成与消除
- 如何才能不烧射频 PA
- PA“自激”的分析与解决
- 射频贴片电感的选型与使用
- 射频开关，简约不简单
- 高可靠性芯片的设计与评估

在上一章节中，我们领略了射频芯片中半导体技术的魅力。而射频芯片之所以能够在无线通信系统中发挥如此重要的作用，离不开其精心设计的各个子模块。这一章，我们将对射频系统中各芯片子模块的设计展开详细讨论，揭开射频芯片的神秘面纱。

射频芯片作为无线通信系统的核心，其内部包含了多个子模块，每个子模块都承担着不同的功能，共同协作以实现高效的信号传输与接收。

射频收发机是射频芯片的核心部分，它负责将基带信号转换为射频信号进行发送，同时将接收到的射频信号转换为基带信号进行处理。射频功率放大器（PA）也是射频芯片中不可或缺的一部分。PA 的主要作用是将射频信号放大，以便其能够足够强大地传输到天线并辐射出去。此外，射频开关也是射频芯片中的一个重要子模块。它负责在不同的通信路径之间进行切换，以确保信号能够准确地传输到目标设备。在这一章节中，我们将对以上子模块展开进行讨论。

在讨论了射频芯片的各个子模块之后，我们不得不提及芯片的可靠性设计。可靠性是射频芯片设计中的重要考量因素之一，它直接关系到芯片的使用寿命和稳定性。

通过这一章的讨论，我们希望射频芯片不再是一个黑盒子，希望大家能够了解射频芯片的各个子模块及其功能，掌握其设计原理和方法，从而更好地应用射频芯片，推动无线通信技术的不断发展。

2.1 射频收发机：从无线电到 5G

作者：彭洋洋 博士

你知道吗？你手中的智能手机，其实是一个强大的无线电收发器。它可以让你用手机打电话、发短信、上网、看视频、玩游戏；也可以帮助你用无人机拍摄美景、控制智能家居……这些都是无线通信的奇妙应用，给人类生活带来了无限的便利和乐趣。

无线通信是采用射频信号作为媒介进行信号传输的技术，要完成无线通信，就首先需要把信号先转换为射频信号。

“射频收发机（RF Transceiver）”就是这样一种装置，它负责把基带信号、模拟信号转化为射频信号，交给放大器、天线输出；还负责把接收到的射频信号还原为基带信号、模拟信号，让这些信息变成可看到的视频、可听到的声音。射频收发机是无线通信中的基础模块，是手机、卫星通信、雷达等无线通信设备中必要的构成部件。

射频收发机的历史可以追溯到 19 世纪末，当时人们利用电磁波进行无线电报通信。随着科技的进步，射频收发机也不断地演变和创新。从最早的晶体管收发机，到后来的集成电路收发机，再到现在的多频多模收发机，射频收发机的性能、功能和规模都有了巨大的提升。

随着 5G 的到来，手机系统变的越来越复杂，对射频收发机也提出更高的要求。5G 手机中的 RF Transceiver 芯片是如何一步步演进而来的，未来又有什么演进趋势？接下来，我们将详细介绍射频收发机的基本原理和结构，并探讨它在 2G 到 5G 通信中的不断演进。

2.1.1 射频收发机简介

“射频收发机”一词翻译自 Radio Frequency Transceiver。Transceiver 一词是发射机（Transmitter）与接收机（Receiver）的合成词，通过这个词的构成就可以看出，Transceiver 的功能是完成信号的发射与接收。

在行业应用中，为了与收发机设备整机区分，射频收发机一般被直接称为 Transceiver，有时也被简写为 XCVR。在一些 SoC 芯片厂商中，Transceiver 芯片因为是一颗射频芯片，所以也被称为 RFIC。

虽然 Transceiver 也是负责信号的发射与接收，但其功能与同样具有发射与接收功能的“射频前端（RF Front-end）”不同。射频前端一般指天线之后，用于处理射频信号的通路、信号强度的部分，包含四大基本模块：功率放大器、低噪声放大器、开关、滤波器。而 Transceiver 负责将模拟信号和射频信号进行相互转换：在发射时将模拟信号转换为射频信号，在接收时将射频信号转换为模拟信号。

Transceiver 与射频前端的关系就像是一对兄弟。哥哥 Transceiver 头脑清晰，是一家之主，不断地将信息放置在合适的射频通道上；再把合适的射频通道的信号取下来，转换成有用的信息。弟弟射频前端身体条件好，可以把哥哥转换好的射频信号用强大的力气发射出去；同时还可以把非常微小的射频信号小心放大，交给哥哥处理。

哥哥和弟弟一定要配合紧密，才能完成信号的完美收发。如果需要设计和使用好射频前端芯片，就一定要对 Transceiver 芯片的工作原理有所了解。

Transceiver 芯片与基带芯片、调制解调芯片、射频前端芯片在系统中的构成如图2.1所示。

2.1.2 为什么要进行射频传输

Transceiver 的主要功能是完成模拟信号到射频信号的传输，在了解 Transceiver 之前，首先要回答的第一个问题就是：为什么要进行射频传输？

射频传输是利用电磁波在空气或其他介质传播的特性，将信息以射频信号形式进行无线传输的过程。射频信号的频率范围一般为 3kHz 到 300GHz。

射频传输有以下几个优点：

- 射频传输可以克服有线传输的物理局限性，实现远距离、无障碍、移动和灵活的通信
- 射频传输可以实现多用户、多场景通信，利用多种调制、复用和编码技术，提高通信效率和质量

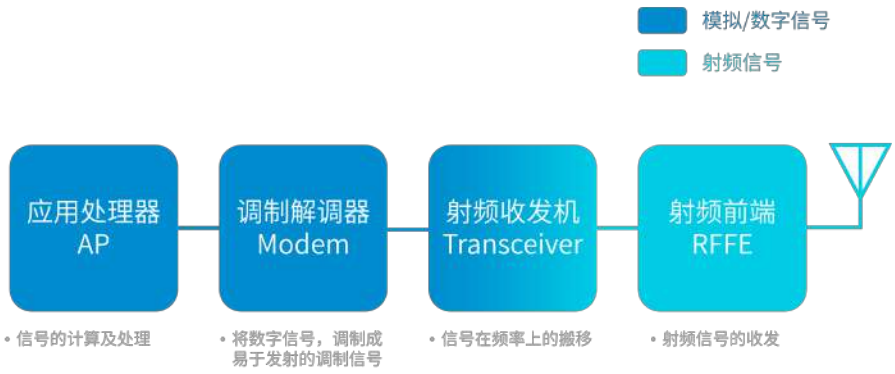


图 2.1: 射频系统构成

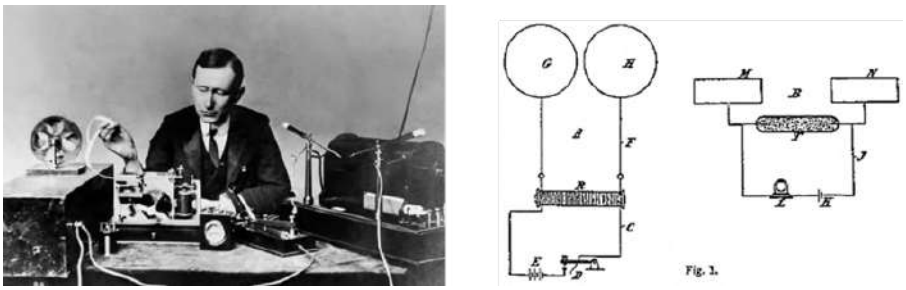


图 2.2: 马可尼和他设计的射频收发电路

- 射频传输可以实现更复杂的组网，利用多种天线技术，如定向天线、智能天线、相控阵天线等，实现信号发射与接收方向的控制和优化
- 射频传输可以实现比有线传输更强的保密性，利用加密、扩频和跳频等技术，提高通信安全性

射频传输是无线通信技术的基础和核心，它为人类社会带来了巨大的便利和价值。为了实现射频传输，重要的一个步骤就是将日常中的图像、声音、视频等信号，转化为射频信号。模拟信号转换为射频信号的过程，就是射频收发机的主要功能。

2.1.3 历史上重要的“Transceiver”电路

19 世纪末：雏形显现

自 1864 年麦克斯韦提出电磁波理论以来，人类一直想象着这个看不见、摸不着的神奇物体在哪里才能用起来。1895 年前后，马可尼、波波夫、特斯拉等人均意识到电磁波可以用来实现有线通信的无线化，并设计了无线电发射、接收器原型。1896 年，意大利天才无线电工程师马可尼获得了世界上第一个无线电专利，由此也打开了无线电通信快速发展的大门。

在马可尼设计的射频转换电路中，发射机使用摩斯电码键作为输入，产生间歇性电流脉冲，脉冲信号连接至高频振荡器，由此摩斯电码就可以完成对高频信号的调制，调制后的信号通过天线发射到空间中。在接收机中，马可尼使用了金属粉末检波器，通过检波，可以将无线信号转换为可以听见的声音信号，并通过耳机输出。由此，马可尼完成了人类历史上首个“Transceiver”电路。这篇专利也成为马可尼获得诺贝尔物理学奖的重要依据之一。马可尼和他设计的射频收发电路如图2.2所示。

随后，马可尼对此架构做了改进，加入了调谐电路，可以改变电路的震荡频率，更加方便了无线电信号的发射与接收。马可尼使用他发明的无线电系统，分别实现了跨英吉利海峡及跨大西洋的通信。

如果说马可尼的发明只是带领人类简单领教无线电的功能的话，1907 年，美国工程师德 • 福雷斯特（De Forest）发明的真空三极管就使得全球范围内的广播、电话、通信成为了可能。德 • 福雷斯特发现在真空二极管的基础上，增加一个栅极可以进行对二极管电流的控制，根据这个特性，德 • 福雷斯特发明了放大器、振荡器等电路，使无线信号的放大和振荡成为可能，进而帮助了无线电广播和远程电话的实现。

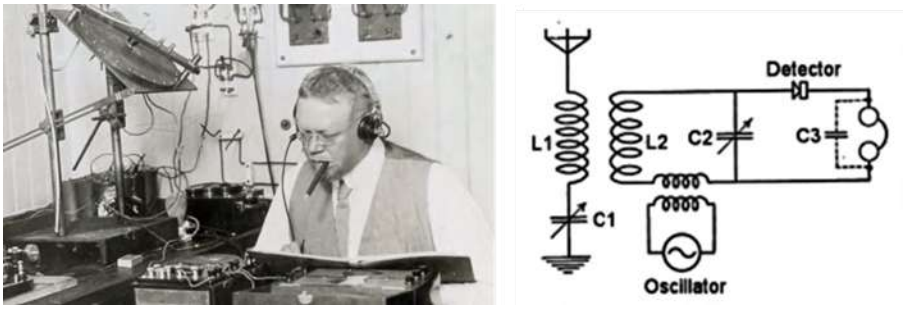


图 2.3: 费森登和其发明的外差接收电路

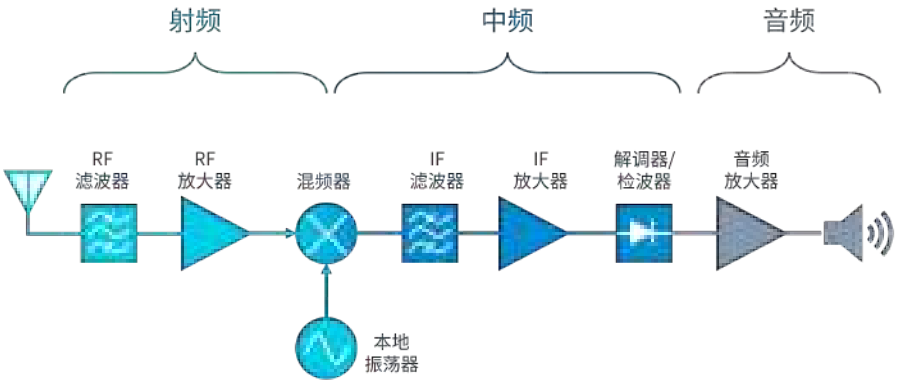


图 2.4: 超外差架构

外差与超外差：走向现代

外差收发机的英文名是 Heterodyne，是人类 Transceiver 历史上的伟大发明。

外差是通过混频技术，将两个频率的信号混合而创建新的频率信号的技术。两个输入信号通过一个非线性器件（如真空管、晶体管或二极管）进行混频，如两个频率为 f1 与 f2 的信号，混频之后会产生 f1+f2 与 f1-f2 两个新频率的信号。这种现象叫混频处理，用于实现混频的非线性器件被称为混频器。

通过混频，就可以将原来在空间传播的电磁波，变换为人耳朵可听到的较低的频率范围信号，再通过简单的检波器，就完成了信息的接收。

1901 年，雷金纳德 • 费森登（Reginald Fessenden）展示了这种架构的收发机，虽然此时三极管还未发明，振荡器的工作频率还无法稳定，但这种架构为现代 Transceiver 奠定了坚实的基础。费森登和其发明的外差接收电路如图2.3所示。

在发明这个架构后，费森登从希腊单词“Hetero-”（不同的，差异的）和“dyn-”（动力、能力）得到灵感，将此种架构取名为 Heterodyne，中文翻译为“外差”。

发现了“外差”现象后，工程师们继续探索。工程师们发现采用更高频率的电磁波传输时，对于某些应用场景有帮助，但工作于高频率的放大器极难设计。工程师就想，既然原来外差的思路是把声音频率的信号，通过混频搬移到高频电磁波频率，那对高频率信号放大的时候，是不是也可以先在比较低的射频频率进行放大，然后再通过频率搬移的方式，将放大后的信号搬移到高频率呢？这样不就可以省去对高频高线性的放大器需求了吗？

以上这个设计理念就是美国工程师爱德华 • 阿姆斯特朗（Edwin Howard Armstrong）等人在 1918 年提出的想法。频率搬移过程中，中间预先设定的固定射频频率被称为“中频”（中间频率，IF，Intermediate Frequency），由于这个频率超出了声音可听的范围，是“超声波”（supersonic），所以被命名为“超外差”（Super-Heterodyne）。其架构如图2.4所示。

相比于高频放大式收发机，超外差架构具有高灵敏度、高选择性和稳定性，能适应远程通信对高频率、弱小信号的接收需要。在过去 100 年时间里，超外差结构在无线通信系统中得到了广泛的应用。

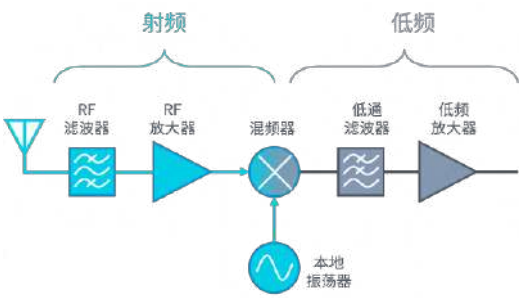


图 2.5: 零中频方案

零中频：方案简化，难度增加

零中频的思路是不再经过 IF 频率，而是直接将射频信号转化为 0Hz 频率范围的基带信号。由于相当于在超外差结构中将 IF 频率设为了 0，所以称为零中频方案（Zero IF），又称为直接变频方案（Direct Conversion），以及零差方案（Homodyne）。其架构如图2.5所示。

零中频方案有其独特的优点，比如：

- 零中频方案可以简化设计，不需要先将信号搬移至中频
- 零中频方案可以解决超外差方案中的镜像抑制问题
- 零中频方案不需要中频滤波器等电路，便于做单芯片集成

零中频方案再也不需要经过一次中频转换了，看起来非常简洁，但这会给实际设计带来诸多问题：

- 没有中频的预处理后，基带输出电平会因为接收信号强度的不同出现很大范围的波动
- 本振频率与射频频率相同，可能造成信号泄露干扰
- 混频后的信号在 0 频率附近，可能会发生 DC 偏移
- 本地振荡器需要有精确的锁相，才能使射频频率准确搬移至 0 频率附近

正因为以上挑战，零中频架构在 1924 年被提出后，并没有得到大规模的推广。在 1932 年，工程师们采用本地振荡器与射频频率二者相比较的方式，修正本地振荡器的频率，可以让本地振荡器频率与射频频率锁定，这个电路也就成为了当今锁相环（Phase Locked Loop，PLL）的雏形。

零中频的其他一些问题也随着 1958 年世界上第一款集成电路被发明之后而被逐渐解决。集成电路的发展使得锁相环电路得以实现更复杂的功能，高动态范围、高补偿特性的电路使得电路可以应对空间大范围波动的射频信号。同时，零中频方案便于单芯片集成的特性使得其与集成电路的快速发展相得益彰。目前，零中频方案在手机、航空电子设备及软件定义无线电系统中得到广泛应用。

2.1.4 从 2G 到 5G：Transceiver 不断演进

经过百年发展，射频 Transceiver 从原来只能发射/接收一个电火花，发展到如今支持全球频段、多功能、多模式的复杂芯片系统。进入 21 世纪后，通信协议仍在不断发展，也促使 Transceiver 技术不断演进。

2.1.4.1 2G：CMOS 涌现，单片集成

2G 蜂窝标准（以 GSM 为例）的主要应用是语音通信，2G 于 1990 年后，在全球开始大规模商用。

2G 手机的普及伴随着集成电路的快速发展而来。随着摩尔定律的演进，1995 年前后 CMOS 工艺的特征尺寸已经缩小至 1μm 量级。而 0.6μm 特征长度的 CMOS 器件已经可以用于设计 2.4GHz 的射频电路，0.35μm 的器件甚至可以使 5GHz 的电路实现成为可能 [2]。

仅仅具备单个射频模块的构建能力还不足以展示出 CMOS 工艺在射频应用中的优势，吸引大家关注 CMOS 工艺的还是其提供的大规模单片集成可能性。CMOS 工艺原来是为数字工艺准备的，并且也可以做一部分模拟电路，如果连射频也能攻克下来，就可以实现复杂的模数、射频混合电路，同时做到单芯片的集成。因为这一特性，CMOS 工艺实现的 2G Transceiver 成为当时的研究热点 [2][3] 。

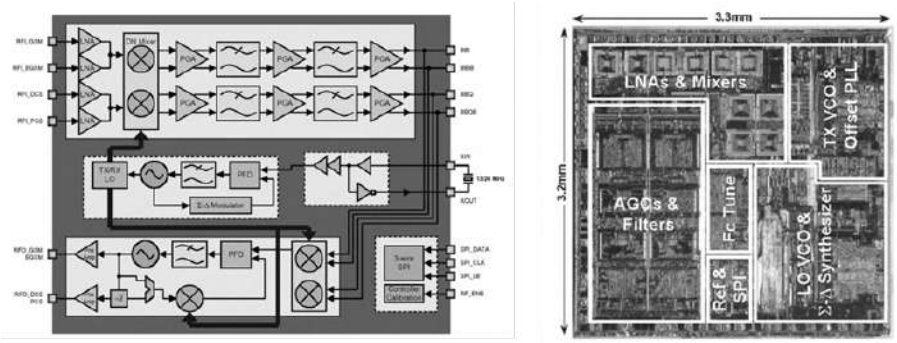


图 2.6: 采用 0.25μm CMOS 工艺设计的四频 GSM 收发机

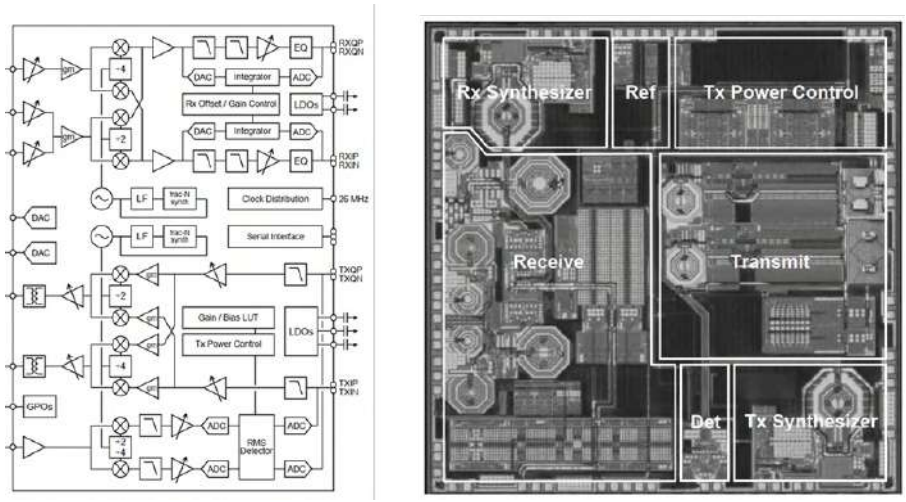


图 2.7: 采用 0.18μm CMOS 工艺设计的 WCDMA/HSDPA 收发机

CMOS 实现全集成的 GSM Transceiver 并不顺利，早期 GSM Transceiver 采用 BJT 技术，并且需要大量的外部器件 [4]。随后，一些 CMOS 工艺设计的单频段的 GSM Transceiver 被设计出来 [5][6]，随后才逐渐开始设计出现多频段全集成的 CMOS Transceiver 芯片。文章 [7] 中展示了采用 0.25μm CMOS 设计的全集成 4 频段 GSM Transceiver 设计，设计采用直接转换的接收机架构，和偏移本振的发射机结构，集成 PLL、VCO、混频器、中频滤波器及放大器，芯片面积 3.2x3.3mm，如图2.6所示。

2.1.4.2 3G：FDD 收发同时，单片集成更大挑战

3G 时代代表的通信制式是 WCDMA，WCDMA 是一种 FDD 频分利用系统，发射机与接收机同时工作在不同频率上，这对单片集成的 Transceiver 设计提出了更大的挑战。

在 FDD 系统中，接收机的接收灵敏度受以下四种情况影响，分别是：接收机的噪声系数；Rx 接收带内的 Tx 噪声；Tx 大信号的混频噪声；Tx 信号的 IM2 产物。以上几种情况的影响中，有三项与发射机与接收机之间的隔离直接相关。

在 3G Transceiver 的设计中，可以采用增强 LNA IIP2、增加陷波网络的方法解决阻塞问题，提升收发机的接收性能。文章 [8] 展示了采用 0.18μm 设计的单片集成 WCDMA/HSDPA Transceiver，文章利用数字信号处理和可调谐滤波器来消除外部元件，从而实现了高度集成和高收发抑制度的 WCDMA 收发器，如图2.7所示。

2.1.4.3 4G：频段碎片，CA 增加

4G 与智能手机几乎在同一时代出现，为了满足智能手机对高数据速率蜂窝通信的需求，越来越多的频段被开辟出来。运营商也在频率资源上展开激烈竞争，结果使得每个运营商掌握的都是非连续和碎片化的多个窄频段。在 4G 手机中，需要支持的频段可能多达 40 个。

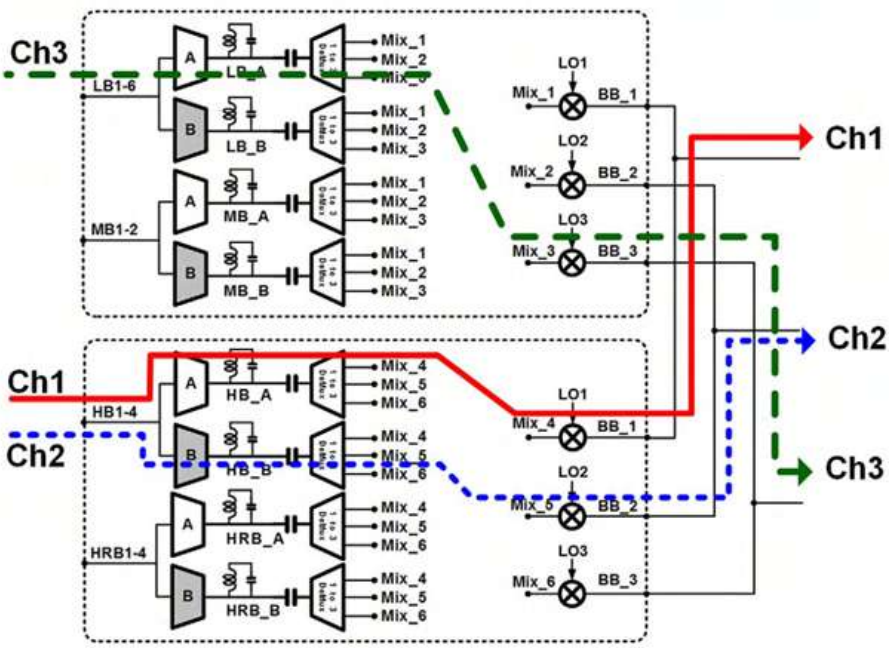


图 2.8: 支持 3CC CA 的 4G LTE Transceiver 设计

频段的增加给 Transceiver 设计带来极大的挑战，在设计中，必须考虑充分的复用，来使子模块的数目维持在合理范围内。

4G Transceiver 另一个更大的挑战是 CA（Carrier Aggregation，载波聚合）的支持。CA 要求多个射频通路同时工作，而这些同时工作的信号之间不可避免的会产生耦合。在设计中，需要将射频通路有效分组。如图2.8所示支持 3CC 的 LTE 接收机系统，Ch1 为 2.1GHz 信号，Ch2 为 2.3GHz 信号，Ch3 为 700MHz 信号。由于 Ch1 为 Ch3 信号的三倍频，所以二者之间需要分配至不同的混频器组 [9]。

2.1.4.4 5G：MIMO/EN-DC，更多通路

5G 到来使得无线通信的速率再次提升，Transceiver 需要实现 Gb/s 吞吐量的收发功能。为此 5G NR 系统中引入了大规模的 MIMO、高达 200MHz 的 CA 来实现。另外，加上 LTE+NR 双连接（EN-DC）的需求，5G NR Transceiver 的设计难度大大增加。

文章 [10] 中，MediaTek 提供了一种采用 12nm CMOS 工艺设计的 5G Transceiver 系统，该系统最多支持 2 个带间上行 CA，6 个带间下行 CA，支持 4x4 MIMO，并且支持 NR 200MHz 的 CA。为了实现以上功能，该 Transceiver 集成了 20 个 Rx 路径，频率覆盖 600MHz 至 6GHz 频段。即使经过了内部的 LNA 复用技术，内部 LNA 还是达到了 28 个。Transceiver 还使用了大量数字电路，来达到 200MHz 的带宽支持。在 NR 200MHz/4x4 MIMO/256QAM 下，可达到 5Gb/s 的吞吐量。其架构如图2.9所示。

2.1.5 总结

随着协议演进，对终端的通信能力要求越来越高，射频 Transceiver 的功能从最早只需要完成一个电火花的接收与发射，到现在必须要完成每秒数 Gb 数据量的传输。这些需求给 Transceiver 设计带来极大挑战。

CMOS 射频集成电路的发展使得这一切成为可能，基于不断演进的 CMOS 工艺，可以在单颗芯片上实现不同电路模块的集成，进而实现多个频段、多个模式，甚至多个通道的电路集成，使无比复杂的射频收发功能，都集成于方寸之间。

Transceiver 的复杂实现，也使得 Transceiver 设计的门槛不断拉高。在早期 3G 时代，还存在多家设计射频 Transceiver 的第三方公司，到 5G 之后，移动终端的射频 Transceiver 已被头部 SoC 平台公司所垄断，如高通、联

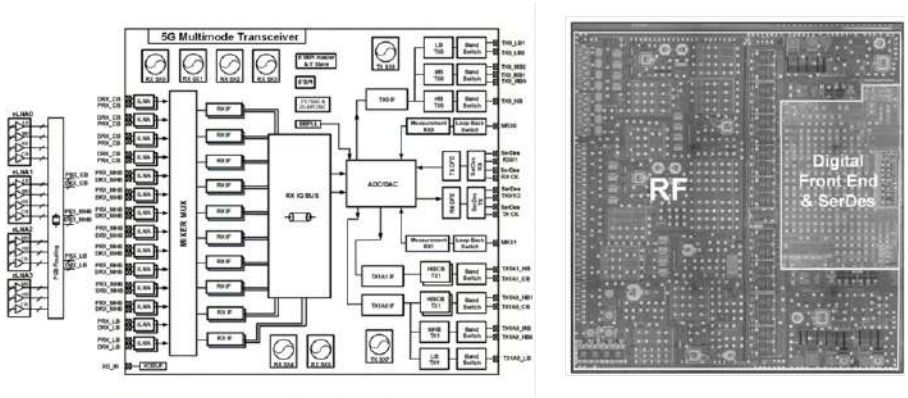


图 2.9: 采用 12nm CMOS 工艺设计的 5G Transceiver

发科等。

随着 Transceiver 功能的复杂，也对 Transceiver 的使用及射频前端其他电路的配合提出了更高的要求。在高通及联发科最新的 5G Transceiver 使用手册中，除了介绍基本的性能之外，大量篇幅用于介绍 Transceiver 在 CA、EN-DC 下使用的注意事项，射频前端配合时需要的注意事项。

在未来，射频 Transceiver 将面临更多频段、更高带宽、更低功耗、更高集成度等要求。为了满足这些要求，射频 Transceiver 可能会采用更先进的工艺、更灵活的架构、更智能的控制等技术创新，同时也对 Transceiver 的应用和使用提出了较高的要求。

期待和您一起，更深入的了解射频 Transceiver 系统，更好的实现 5G 射频系统。

参考文献

1. GUGLIELMO MARCONI, Patent 5028, March 5, 1896
2. A. A. Abidi, “The future of CMOS wireless transceivers,” in IEEE International Solid-State Circuits Conference. Digest of Technical Papers (ISSCC), San Francisco, CA, USA, 1997, pp. 118-119.
3. B. Razavi, “Challenges in portable RF transceiver design,” in IEEE Circuits and Devices Magazine, vol. 12,1996
4. T. D. Stetzler, I. G. Post, J. H. Havens, and M. Koyama, “A 2.7–4.5 V single chip GSM transceiver RF integrated circuit,” IEEE J. Solid-State Circuits, vol. 30, no. 12, pp. 1421–1429, Dec. 1995
5. P. Orsatti, F. Piazza, and Q. Huang, “A 20-mA-receive, 55-mA-transmit, single-chip GSM transceiver in 0.25-µm CMOS,” IEEE J. Solid-State Circuits, vol. 34, no. 12, pp. 1869–1880, Dec. 1999.
6. M. S. J. Steyaert, J. Janssens, B. De Muer, M. Borremans, and N. Itoh, “A 2-V CMOS cellular transceiver front-end,” IEEE J. Solid-State Circuits, vol. 35, no. 12, pp. 1895–1907, Dec. 2000.
7. E. Song, Y. Koo, Y.-J. Jung, D.-H. Lee, S. Chu and S.-I. Chae, “A 0.25-µm CMOS quad-band GSM RF transceiver using an efficient LO frequency plan,” IEEE Journal of Solid-State Circuits, vol. 40, no. 5, pp. 1094-1106, May 2005.
8. B. Tenbroek, J. Strange, D. Nalbantis, M. Poulton and A. Burrows, “Single-chip tri-band WCDMA/ HSDPA transceiver without external SAW filters and with integrated TX power control,” in IEEE International Solid-State Circuits Conference - Digest of Technical Papers (ISSCC), San Francisco, CA, USA, 2008, pp. 480-481.
9. Menzel, W., &Lee, K. (2015). A Rel-12 2G/3G/LTE-Advanced 3CC Cellular Receiver. IEEE Journal of Solid-State Circuits, 50(1), 54-68.
10. Ming-Da Tsai et al., ”A 12nm CMOS RF Transceiver Supporting 4G/ 5G UL MIMO,” 2020 IEEE International Solid-State Circuits Conference (ISSCC), San Francisco, CA, USA, 2020, pp. 180-182, doi: 10.1109/ ISSCC19947. 2020. 9062939.

在线阅读

本文摘自慧智微微信公众号，扫描以下二维码，在线阅读



2.2 射频 PA 的工艺选择

作者：小慧

与数字、模拟芯片只选择 CMOS 工艺所不同，射频 PA 可选的工艺众多，比如：

1. GaAs
2. GaN
3. CMOS
4. SiGe
5. LDMOS
6. InP

这些工艺选择已经让设计师们犯了难，再加上“频率”、“功率”维度，工艺选择就更加困难。

所以，“某个应用下 PA 应该选什么工艺？”，成了 PA 设计中旷日持久的讨论热点。虽然有一些定性的参考可以帮助选择，比如说：

1. CMOS 工艺更适合低功率、低成本应用；
2. GaAs 工艺更适合较大功率应用；
3. GaN 可应用于高频极大功率等。

但这些参考太过定性。有没有清晰直观的数据参考，来指导大家设计中进行工艺选择呢？

乔治亚理工的 Hua Wang 教授团队就将 2000 年至今发表在知名期刊的 3,813 篇文章做了统计，得出了 PA 工艺与频率、饱和功率间的关系图。

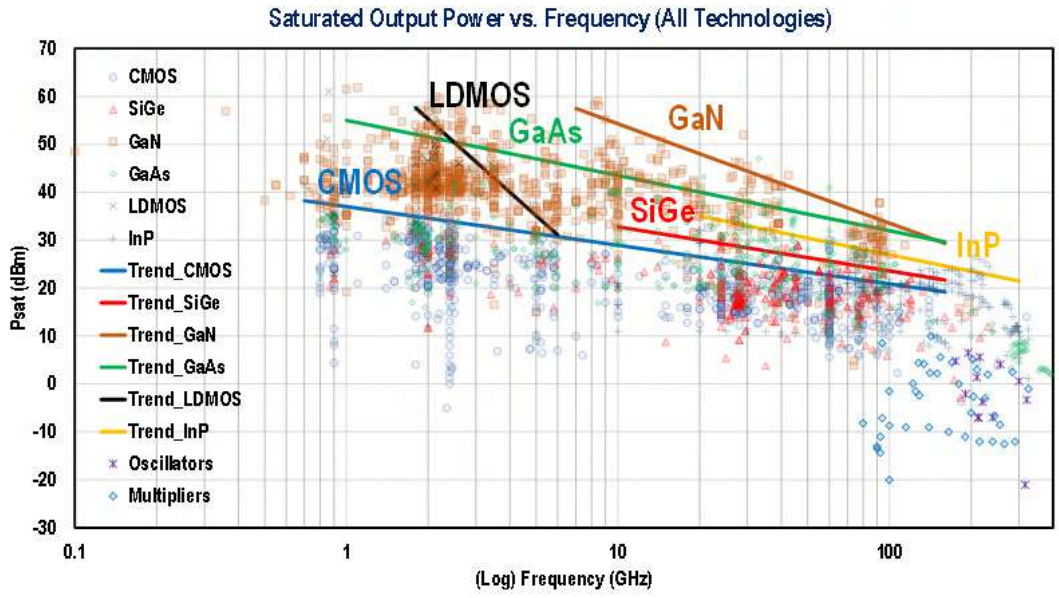


图 2.11: 不同 PA 工艺与频率及功率间的关系图

Dr. Hua Wang 教授还给出了不同统计维度下的数据统计结果，如分不同工艺 PA 的饱和功率、频率统计。

Hua Wang 教授团队不仅将统计结果给出来，而且还给出了原始数据，扫描文章末尾二维码，关注公众号点击页末阅读原文，可下载统计结果的原始 excel 文件。

对 Hua Wang 教授的整理和分享表示由衷的感谢！

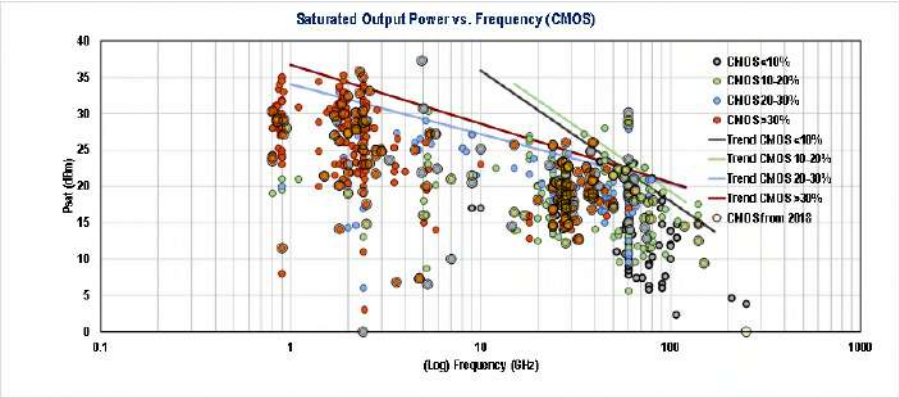


图 2.12

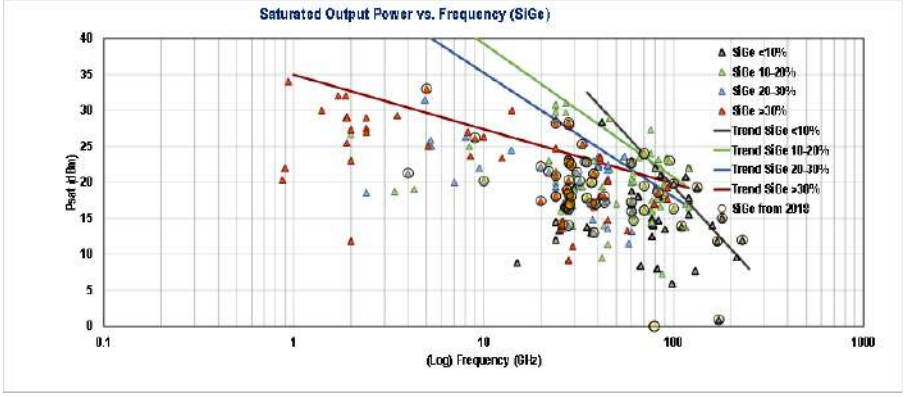


图 2.13

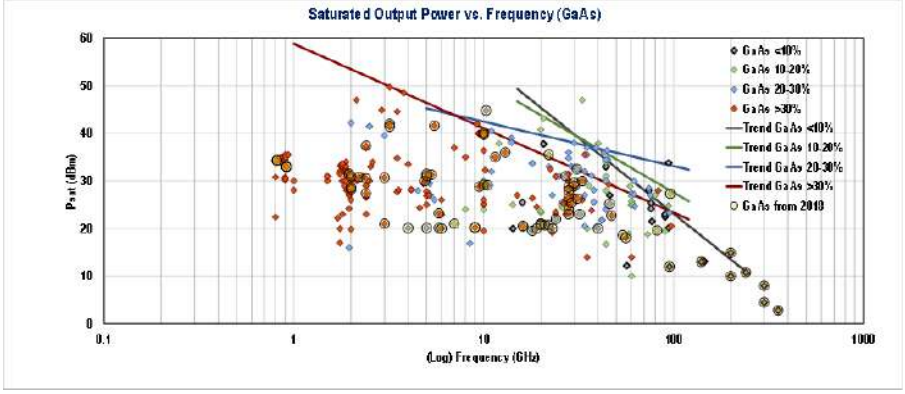


图 2.14

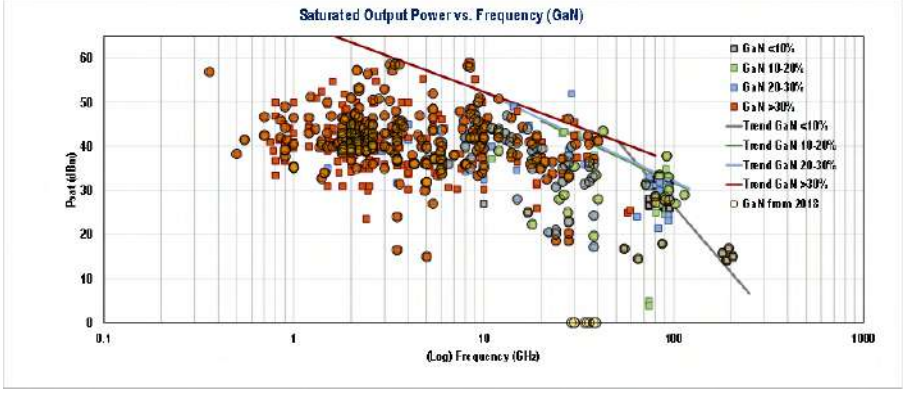


图 2.15

参考文献

1. Hua Wang, Tzu-Yuan Huang, Naga Sasikanth Mannem, Jeongseok Lee, Edgar Garay, David Munzer, Edward Liu, Yuqi Liu, Bryan Lin, Mohamed Eleraky, Hossein Jalili, Jeongsoo Park, Sensen Li, Fei Wang, Amr S. Ahmed, Christopher Snyder, Sanghoon Lee, Huy Thong Nguyen, and Michael Edward Duffy Smith, ”Power Amplifiers Performance Survey 2000-Present,” [Online]. Available: https://gems.ece.gatech.edu/PA_survey.html

在线阅读

本文摘自慧智微微信公众号，扫描以下二维码，在线阅读



2.3 5G 射频 PA 的 Load-line 与 Load-pull

作者：Solomon

说到射频 PA（Power Amplifier，功率放大器）的设计和应用，有两个名词经常被大家提及：Load-line 与 Load-pull。在使用中，这两个名词太过常用了，以至于对这两个名词后面的理论依据反而讨论不多。接下来我们就对 Load-line 和 Load-pull 背后的知识做一个讨论。

2.3.1 Load-line：负载线，PA 的生命线

Load-line 中文名称为负载线，是 PA 设计中最为重要的设计参考。

对于 PA 来说，最重要的目的是进行功率有效的放大与输出，一切的设计理念均是为此服务。此时，以最大功率“传输”为首要目标的共轭传输无法完成最大功率“输出”的目的：PA 需要实现的是将射频功率最大程度的从直流功率中榨取出来，而不是将已经产生的信号传输出去。实现射频功率最大化输出用到的就是“最佳负载匹配”，而不是用来最大化传输已有功率的“共轭匹配”。

为了直观分析 PA 在不同负载下的功率输出，在 PA 设计中就引入了“负载线（Load-line）”的概念，用以观测 PA 在不同负载下的射频电压与电流摆幅，从而得出不同负载下的功率输出，找到最佳负载。

2.3.1.1 晶体管的 DC-IV 曲线

PA 工作是基于晶体管的特性工作的，所以在讨论 PA 的 Load-line 之后，首先对晶体管特性做一个分析。

PA 可以采用 HBT、MOSFET、pHEMT 等多种半导体工艺进行设计，不同工艺设计在 PA 在 Load-line 理论中分析方法基本相同。以下将以射频 PA 中最为常用的 HBT（Hetero-Junction Bipolar Transistor, 异质结双极型晶体管）器件为例进行分析。

HBT 是一种特殊的 BJT（Bipolar Junction Transistor，双极型晶体管），由两个背靠背连接的 PN 结构成。图2.16为 HBT 器件与其所设计的放大器的基本结构。

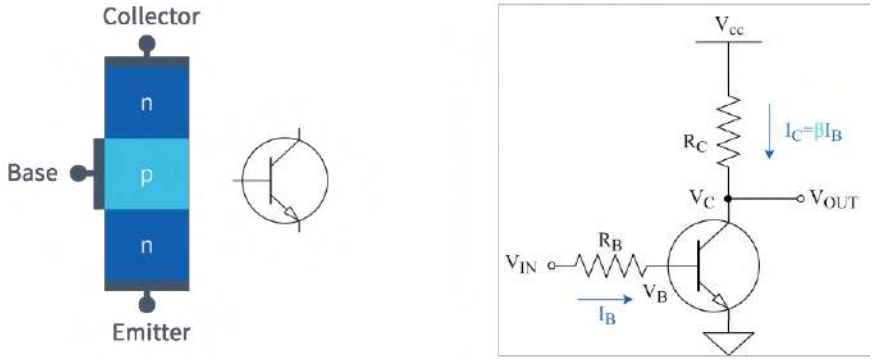


图 2.16: HBT 器件及共发射极放大器基本结构

对于 HBT 器件来说，由于 Base 的小电压（或电流）可以对 Collector 的大电流进行控制，所以一个小的 Base 输出信号，就可以经由 HBT 器件在 Collector 产生大的输出信号，这就是 HBT 作为放大器的基本原理。

理想情况下，HBT 器件在工作中有以下几个特点：

1. V_{CE} 要足够的大，才能建立起 C_E 之间的电流
2. V_{CE} 足够大之后，I_{CE} 不受 V_{CE} 控制，只受 Base 电流 I_B 控制
3. I_{CE} 与 I_B 呈 β 倍关系

简单起见，为了理解这种控制关系，如图2.17及图2.18所示，可以把 HBT 器件理解为一个水龙头,Collector 是水龙头的输入，Emitter 是水龙头的输出，而 Base 就是水龙头的控制：

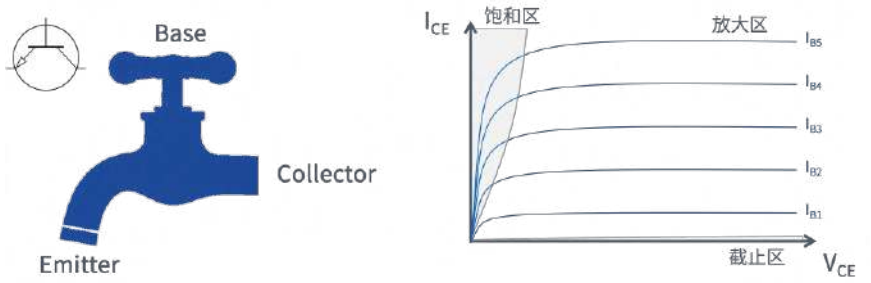


图 2.17: 图：将 HBT 器件理解为“水龙头”；HBT 器件的 DC-IV 曲线

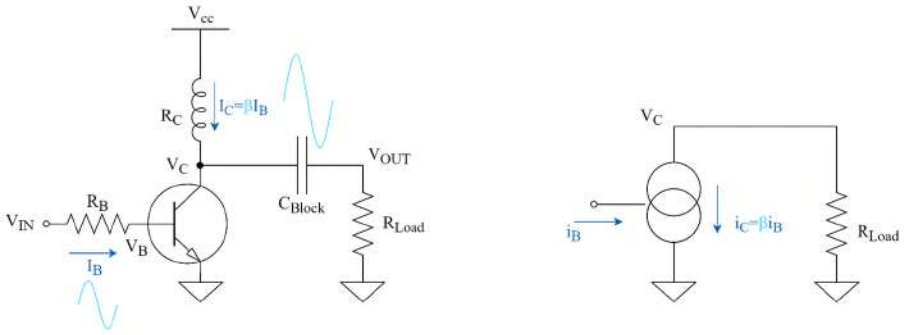


图 2.18: 带有负载的放大器基本电路及其小信号等效模型

1. 水压要足够大，水龙头才可以有水流出（V_{CE} 要足够大）
2. 一旦水压足够大，水龙头的出水量就不再由水压控制，而是由控制龙头 Base 控制（I_{CE} 不受 V_{CE} 控制，只受 Base 电流 I_B 控制）
3. 出水量与龙头控制间有一个比例关系（β 倍）

将 HBT 器件在不同的 V_{CE} 电压及不同的 I_B 控制下的 I_{CE} 在直角坐标系中描绘出来，就得到 HBT 器件的 DC-IV 曲线（直流 DC 状态下的电流电压曲线），DC-IV **曲线是半导体器件中的重要特性曲线。**

2.3.1.2 放大器的 Load-line

实际放大器在工作中，需要驱动负载阻抗。放大器对负载阻抗的驱动作用以及其小信号电路等效如下图所示。在小信号等效电路中，HBT 可以等效为电流控制电流源。

考虑负载 R_{Load} 后，在满足欧姆定律的条件下，V_C 点电压与电流的相互转换以电阻 R_{Load} 为斜率进行相互转换。将这种转换关系对应到前节所述 DC-IV 曲线上，**就得到一根斜率为 $1/R_{Load}$ 的直线，这根线就称之为此时放大器的** Load-line。如图2.19所示。Load-line 的斜率与负载呈反比，中文称为负载线。

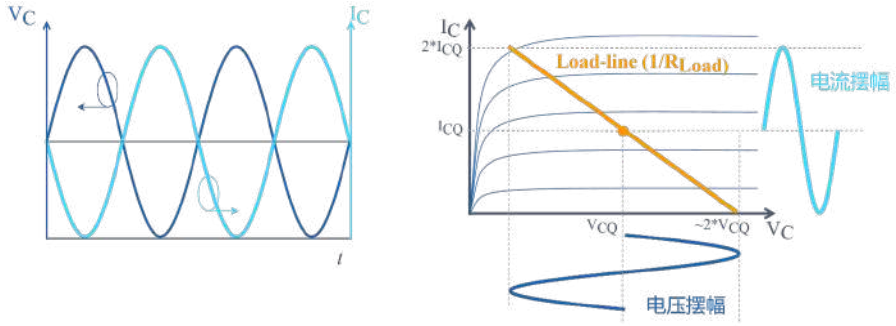


图 2.19: 输出电压与电流摆幅关系及 Load-line

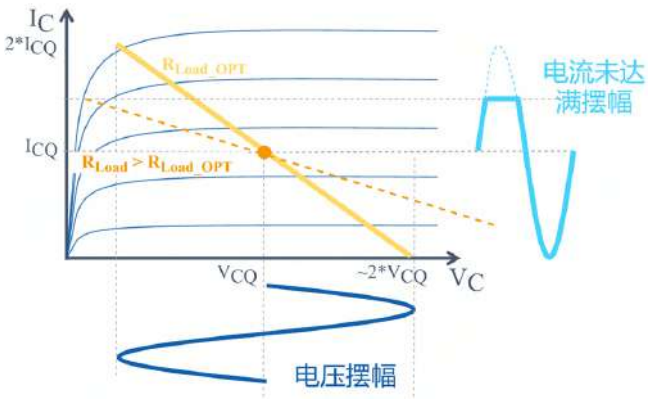


图 2.20: 当 $R_{Load} > R_{Load\_OPT}$ 时的负载线

2.3.1.3 Load-line 与输出功率间的关系

Load-line 之所以重要，是因为其直接决定了放大器最大输出功率。

为了分析 Load-line 与输出功率的定性关系，在手算直观分析中，一般以 A 类放大器进行简化分析。B 类、AB 类等放大器结论相同，不过分析过程更为复杂，若考虑到谐波阻抗、负载阻抗虚部影响，就需要借助仿真软件仔细分析。

对于匹配到最佳负载的功率放大器来说，最大输出功率时其电压与电流摆幅均达到最大，交流峰值分别等于为 V_{CQ} 与 I_{CQ}，此时放大器的输出功率 P_{OUT} 可以有多种表达方式，如：

$$P_{OUT} = \frac{V_{CQ} \cdot I_{CQ}}{2} = \frac{1}{2} \cdot \frac{V_{CQ}^2}{R_{Load}} = \frac{1}{2} \cdot I_{CQ}^2 \cdot R_{Load}$$

此时：

$$R_{Load} = \frac{V_{CQ}^2}{2 \cdot P_{OUT}}$$

并有关系：

$$R_{Load} = \frac{V_{CQ}}{I_{CQ}} = 2 \cdot \frac{P_{OUT}}{I_{CQ}^2}$$

由于在此负载线下，放大器有最大的输出功率，所以此负载线又叫放大器的最佳负载线：R_{Load_OPT}。对于某 5G PA，以 $V_{CQ} = 4.2V$ 电压为例，若目标输出功率为 34.5dBm (2.82W)，则此时的最值 Load-line 在：

$$R_{Load} = \frac{V_{CQ}^2}{2 \cdot P_{OUT}} = \frac{4.2V^2}{2 \cdot 2.82W} = 3.1\Omega$$

当 R_{Load} 远离 R_{Load_OPT} 时，需要分两种情况进行讨论，分别为 $R_{Load} > R_{Load\_OPT}$（高 Load-line）及 $R_{Load} < R_{Load\_OPT}$（低 Load-line）。

- 当 $R_{Load} > R_{Load\_OPT}$

  $R_{Load} > R_{Load\_OPT}$ 时，如图2.20所示，将此时负载线记为 R_{Load_H}。当电压达到满摆幅时，电流并没有达到满摆幅。此时输出功率受限在电压摆幅上，所以输出功率只能以电压摆幅计算，即：

  $$P_{OUT} = \frac{1}{2} \cdot \frac{V_{CQ}^2}{R_{Load}}$$

- 当 $R_{Load} < R_{Load\_OPT}$

  当 $R_{Load} < R_{Load\_OPT}$ 时，如图2.21所示，将此时负载线记为为 R_{Load_L}。当电流达到满摆幅时，电压并没有达到满摆幅。此时输出功率受限在电流摆幅上，所以输出功率只能以电流摆幅计算，即：

  $$P_{OUT} = \frac{1}{2} \cdot I_{CQ}^2 \cdot R_{Load}$$

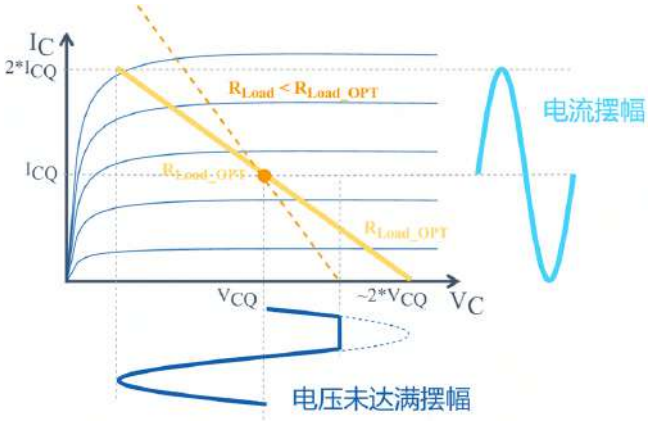


图 2.21: 当 $R_{Load} < R_{Load\_OPT}$ 时的负载线

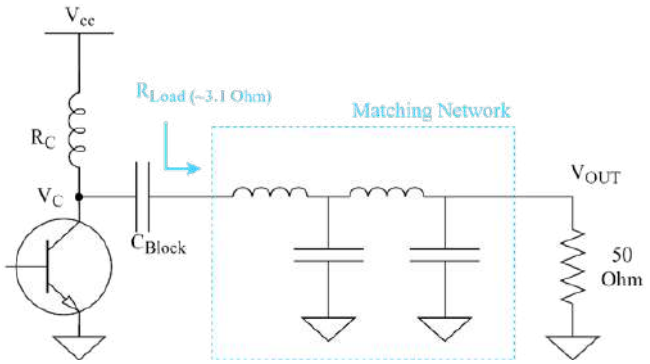


图 2.22: 5G PA 的输出匹配网络

2.3.1.4 Load-line 的阻抗匹配

通过以上分析，可以得到某 5G 射频 PA 的最佳负载阻抗约在 3.1Ω 左右，在设计中，需要设计匹配网络将 50Ω 负载匹配至该目标负载阻抗。输出匹配网络在放大器中的位置与基本结构如图2.22所示。

输出匹配网络将较高的 50Ω 阻抗匹配至较低的负载阻抗 3.1Ω，可以证明，一个匹配网络的损耗，和转换网络的 Q 值（$Q_{Network} = \sqrt{R_H/R_L - 1}$ 成正比，和器件的 Q 值（主要为电感 Q 值 Q_L，以 25 作为估计）成反比，即：

$$Loss \propto \frac{Q_{Network}}{Q_L}$$

若采用单级 L-C 匹配网络，则 50Ω 到 3.1Ω 的损耗为 15.6%，即 0.74dB。

在转换比过大时，可采用多级匹配的方式减小损耗。比如采用两级匹配的方式进行匹配，若将 50Ω 先匹配至 12Ω，再匹配至 3.1 Ω，则两级匹配网络的损耗分别为 7.1% 及 6.8%，整体损耗为 13.4%，约 0.63dB。

2.3.1.5 “高 Load-line”与“低 Load-line”PA

由于输出功率是由 V_{CQ} 与 R_{Load} 两个参数共同决定，所以在设计时可以采用高 V_{CQ} 与高 R_{Load} 的方式（即高 Load-line），也可以采用低 V_{CQ} 与低 R_{Load} 的方式（即低 Load-line）进行设计。

采用低压（配合低 Load-line）方式设计的 PA 优势显而易见：更低工作电压使得供电只需要 Buck 电路，不用 Boost，这样供电电路会简单并且电源转换效率更好。虽然对于同样的输出功率来说，电压降低后电流会升高，但二者乘积相同，总功耗相同。

另外由于手机电源供电约为 3.8V，若采用更高的电压，必须增加升压电路，造成终端成本提高。低压 PA 不需要外部升压电路对电源，降低成本，方便使用。

总结来说，低压 PA 的优势有：

Vcc (V)	Load-line (Ω)	匹配网络损耗增加 (dB)
4.2	3.1	0.00
4.0	2.8	0.02
3.8	2.5	0.04
3.6	2.3	0.06
3.4	2.0	0.08
3.2	1.8	0.11
3	1.6	0.14
2.8	1.4	0.17
2.6	1.2	0.21

图 2.23: 不同电压及 Load-line 下所对应的匹配网络损耗

- 只需要 Buck，省掉外部升压电路，节省成本
- 供电电路会更简单，且电源转换效率更好

低压 PA 虽然优势明显，但对设计的挑战增大。低压 PA 所使用的 Load-line 较低，匹配网络进行转换时匹配网络的 Q 值变大，损耗增加。图2.23列出了采用 4.2V 设计时 Load-line 为 3.1Ω，以此为标准，输出功率相同时，3.4V 低压 PA 的 Load-line 为 2.0Ω，损耗增加 0.08dB，此额外损耗需要在设计中予以克服。对此损耗的克服一般通过优化 PA 设计可以实现。

随着电压再降低，输出匹配网络损耗快速增加，所以 Load-line 不能一直降低。另外，低压使用时，器件本身所占用的膝电压（V_{Knee}，Knee Voltage）占的比重开始增加，进一步限制电压摆幅，影响输出功率。

2.3.1.6 实际应用中的 Load-line 配

以上分析采用简单 Class A PA 进行简化分析，在实际应用中手机 PA 通常用 Class F，Class AB，Class E, 和 Doherty 等。这些 PA 需要的最佳负载可能是复数，并且需要考虑谐波负载，负载线的表现与 Class A PA 会有不同；

以上分析中，最佳负载线根据最大功率进行设计，通常 PA 设计需要综合考虑 PAE 与 ACLR 等其他指标；

另外，最佳负载匹配网络需要综合考虑阻抗变换比/网络元件 Q/频率带宽三个要素。

2.3.2 Load-pull：负载线理论的最佳实践

虽然 Load-line 理论可以对 PA 特性进行简单清晰的分析，但在实际使用中，由于阻抗并非只有实部，并且加入导通角、匹配网络以及谐波影响后会变的非常复杂。Load-line 理论对于清晰理解 PA 的设计思路很重要，但在实际设计与应用中显得心余力绌。

于是，Load-pull 的概念被引入了进来。

说起 Load-pull，射频人都不会陌生：在每本教科书里都会提到；Load-pull 在 Smith 圆图上的等高线（Contour）也是 PA 设计和应用中的必备材料。图2.24为典型的 Load-pull 在 Smith 圆图上的结果呈现。

相比于“Load-pull”名字的熟悉，大家对它背后的理论谈论较少。Load-pull 的测试过程也像是“暴力破解”，好像没什么理论可依。

不过实际并非如此，Load-pull 背后有详细的理论分析，PA 届的大神 Cripps 于 1983 年发表的 IEEE MTT-S 的论文“A theory for the prediction of GaAs FET Load-pull Power Contours”[2] 就曾用纯理论的方式对 PA 的 Load-pull 进行预测。通过测试验证，Cripps 的理论预测完美的匹配了测试结果，如图2.25所示。

接下来，就让我们沿着 Cripps 的思路，仔细理解 Load-pull。

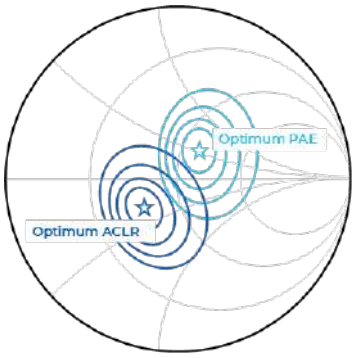


图 2.24: 典型的 Load-pull 结果在 Smith 圆图上的呈现

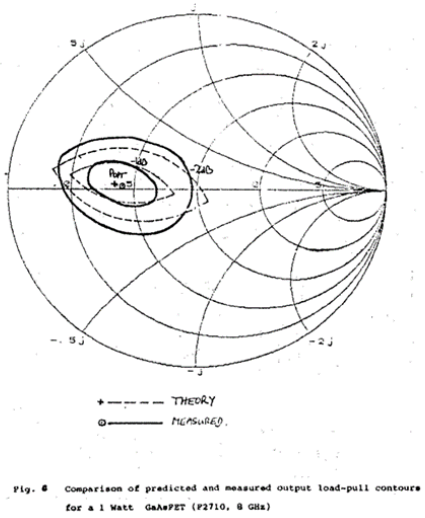


图 2.25: Cripps 在 1983 年对 PA Load-pull 进行的理论难测及验证

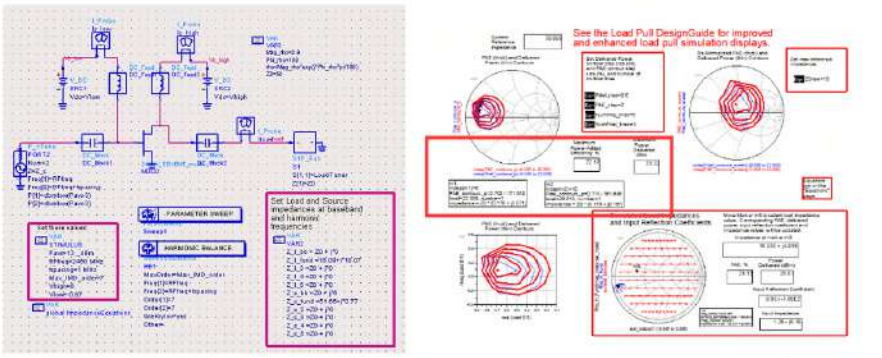


图 2.26: 采用 ADS 软件仿真得到的 Load-pull 结果

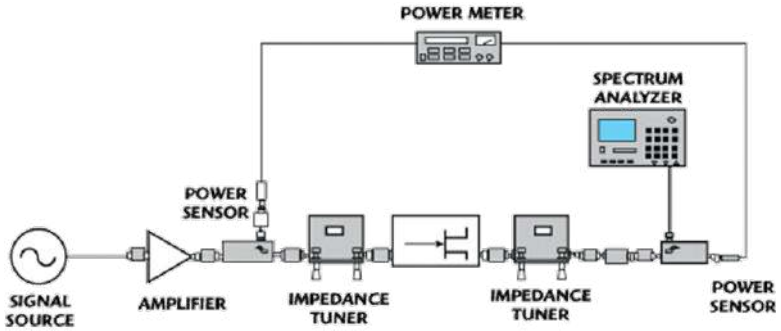


图 2.27: Load-pull 测试系统原理图 [4]

2.3.2.1 什么是 Load-pull

Load-pull 的中文名翻译为“负载牵引”，是指将被测器件（DUT，Device under Test）的负载阻抗进行遍历，同时测试记录不同负载阻抗时的器件特性，从而得到最优阻抗的方法。

在 CAD 仿真软件中，Load-pull 结果的获取较为容易，只需要将 DUT 的负载进行扫描，就可以绘制出多种多样的 Load-pull 图形，通常几秒中，就可以将 Load-pull 结果扫描出来。下图为使用 ADS 软件进行 Load-pull 仿真以及得到的结果 [3]，如图2.26所示。

在实际测试中，想要精准的遍历各个阻抗就不如仿真中容易了，需要借助 Tuner（阻抗调谐器）来实现负载阻抗的控制，Tuner 也是整个 Load-pull 系统中最为重要的组成部分。Tuner 可以理解为阻抗调谐匹配单元，可以将固定的负载阻抗有控制的匹配至 Smith 圆图上的其他位置。Load-pull 测试系统的原理图及实际测试系统如图2.27及图2.28所示 [4][5]。



图 2.28: Load-pull 测试系统（前置左右两侧白色器件分别为源 Tuner 及负载 Tuner）[5]

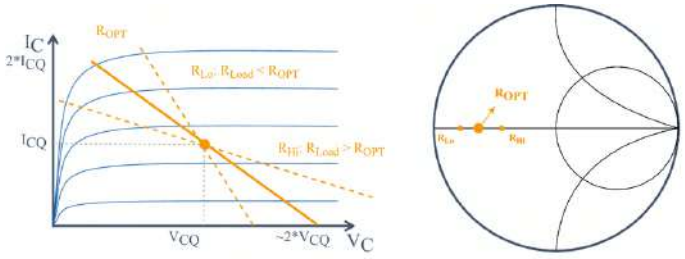


图 2.29: Load-line 的负载与 Load-pull 的阻抗标注

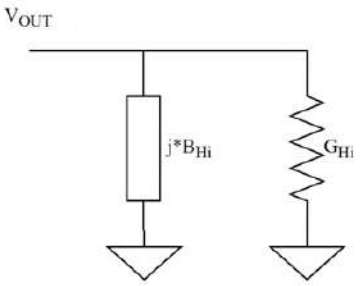


图 2.30: 用于电压驱动时的负载并联电路等效

2.3.2.2 Load-pull 理论

通过对 PA 的仿真或测试，可以得到不同负载下 PA 不同输出功率的等高线图。为何 PA 输出功率会呈等高线形状，另外等高线形状是圆形吗？等高线一定是闭合的吗？接下来将进行详细讨论。

实阻抗在 Load-pull 中的表示

通过 Load-line 理论，可以得到 PA 负载在最佳负载线（R_{Load_OPT}），以下以 R_{OPT} 表示）时有最大的输出功率。将此时的 R_{Load} 表示在 Smith 圆图上，就得到 Load-pull 的中心点。

同理，对于高 Load-line 时的负载 R_{Hi} 及低 Load-line 时的 R_{Lo}，同样可以在 Smith 圆图上标注出来。R_{Hi} 及 R_{Lo} 处分别电压及电流受限，功率均小于 R_{OPT} 处功率。如图2.29所示。

高 Load-line 区：电压受限；等电导圆上功率不变

在高 Load-line 区，Load-line 阻抗大于最优负载阻抗，Smith 圆图表示为在 R_{OPT} 的右侧。此时电压摆幅受限，输出功率以最大电压摆幅计算，为：

$$P_{OUT} = \frac{1}{2} \cdot \frac{V_{CQ}^2}{R_{Hi}}$$

由于电压固定，当输出带有虚部时，采用并联等效电路进行功率计算更为方便，将此时负载电路等效如图2.30所示。

此时，V_{OUT} 的最大值保持恒定，峰值为 V_{CQ}。若电导 G_{Hi} 也保持一致，则输出功率恒定为：

$$P_{OUT} = \frac{V_{CQ}^2 \cdot G_{Hi}}{2} = \frac{1}{2} \cdot \frac{V_{CQ}^2}{R_{Hi}}$$

即在高 Load-line 区，在等电导圆上输出功率恒定一致。此时电压电流波形及在 Smith 圆图上的阻抗位置如图2.31所示。

需要注意的是，当负载阻抗在远离 G_{Hi} 向短路点移动时，虽然电压摆幅保持一致，但电流摆幅会逐步增加。若电流摆幅增加至 I_{Max}，则电流开始受限，不再能使用前述电压与 G_{Hi} 的方式计算功率，即功率无法保持恒定。有关电流达到受限的阻抗点后续将详细讨论。

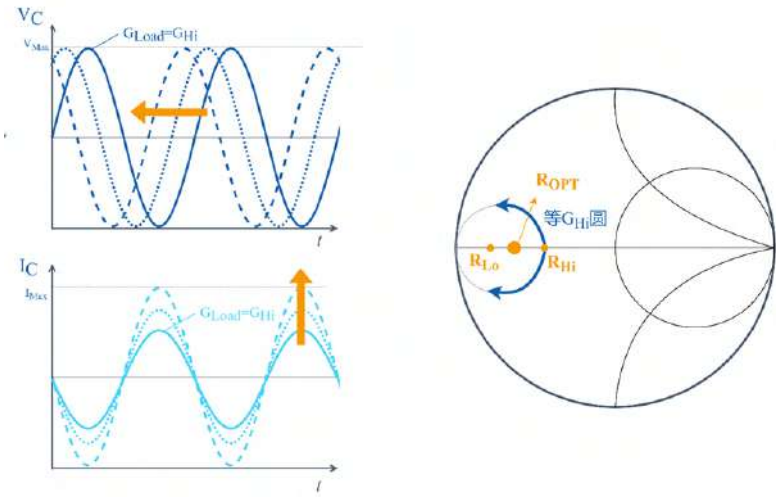


图 2.31: 高 Load-line 区域时，虚部增加对电压电流波形的影响

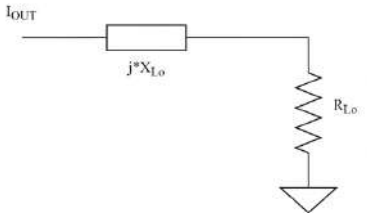


图 2.32: 用于电流驱动时的负载串联电路等效

在此讨论另外一个现象：当负载阻抗沿等 G_{Hi} 圆移动时，可以看到电流摆幅明显增加。为何电流摆幅增加不会带来功率的增加呢？欢迎扫描文章最后二维码，参与留言，讨论对此现象的理解。

低 Load-line 区：电流受限，等电阻圆上功率不变

在低 Load-line 区，Load-line 阻抗小于最优负载阻抗，Smith 圆图表示为在 R_{OPT} 的左侧。此时电流摆幅受限，输出功率以最大电流摆幅计算，为：

$$P_{OUT} = \frac{1}{2} \cdot I_{CQ}^2 \cdot R_{Lo}$$

当输出带有虚部时，由于电流受限，采用串联等效电路进行功率计算更为方便，将此时负载电路等效如图2.32所示。

此时，I_{OUT} 的最大值保持恒定，峰值为 I_{CQ}。若电阻 R_{Lo} 也保持一致，则输出功率恒定为：

$$P_{OUT} = \frac{1}{2} \cdot I_{CQ}^2 \cdot R_{Lo}$$

即在低 Load-line 区，在等电阻圆上输出功率恒定一致。此时电压电流波形及在 Smith 圆图上的阻抗位置如图2.33所示。

与高 Load-line 区域分析类似，在此时同样要注意，随着负载阻抗远离 R_{Lo} 向移动时，电压摆幅也会逐步增加，直至电压受限。也同样可以讨论：为何电压摆幅的增加，没有带来功率的增加？

有关“受限”的讨论

在以上分析中，高 Load-line 区域电压受限，但当负载沿等 G_{Hi} 圆移动时，电流摆幅增加，直至受限；低 Load-line 区域电流受限，但当负载沿等 R_{Lo} 圆移动时，电压摆幅增加，直至受限。这个受限点在哪里呢？

对于高 Load-line 区域，设 $R_{Hi} = A \cdot R_{OPT}$，则在等 G_{Hi} 圆上，输出功率以 P'_{OUT} 表示，为：

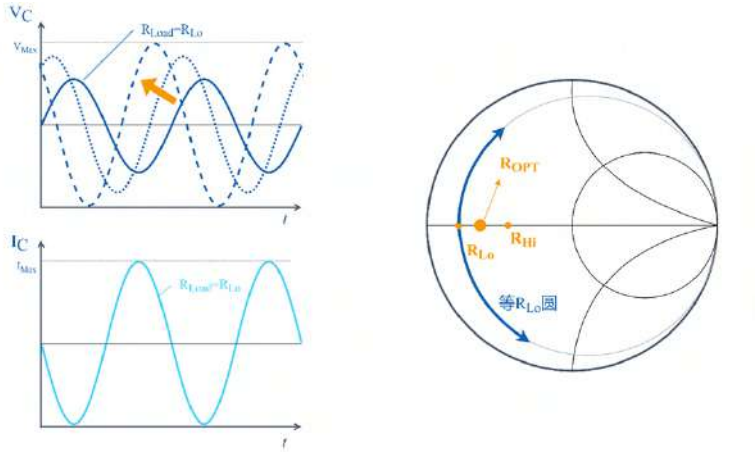


图 2.33: 低 Load-line 区域时，虚部增加对电压电流波形的影响

$$P_{OUT}^{'} = \frac{V_{CQ}^2 \cdot G_{Hi}}{2} = \frac{V_{CQ}^2 \cdot G_{OPT}}{2 \cdot A}$$

此时，输出功率为最大输出功率的 1/A。

在低 Load-line 区域，若得到与此相同的输出功率，根据

$$P_{OUT}^{''} = \frac{1}{2} \cdot I_{CQ}^2 \cdot R_{Lo}$$

可得低 Load-line 区域的阻抗点阻抗

$$R_{Lo} = \frac{1}{A} \cdot R_{OPT}$$

计算高 Load-line 区域沿等 G_{Hi} 圆变化时电流随阻抗实部的变化。此圆上的电流及阻抗实部分别以 $I^{'}$ 及 $R^{'}$ 表示，以电流及阻抗的方式计算等 G_{Hi} 圆上的输出功率为：

$$P_{OUT}^{'} = \frac{I'^2 \cdot R^{'}}{2}$$

令其与电压、导纳计算方式得到的功率相同，则：

$$\frac{V_{CQ}^2 \cdot G_{OPT}}{2 \cdot A} = \frac{I'^2 \cdot R^{'}}{2}$$

即：

$$R^{'} = \frac{V_{CQ}^2}{I'^2} \cdot \frac{1}{A \cdot R_{OPT}}$$

当电流摆幅达到最大，即 $I^{'} = I_{Max} = I_{CQ}$ 时：

$$R^{'} = \frac{V_{CQ}^2}{I'^2} \cdot \frac{1}{A \cdot R_{OPT}} = R_{OPT}^2 \cdot \frac{1}{A \cdot R_{OPT}} = \frac{1}{A} \cdot R_{OPT} = R_{Lo}$$

即：在当等 G_{Hi} 圆与等 R_{Lo} 圆相交时，电流取到最大值，电流与电压同时受限。此时用 Smith 圆图表示的 Load-pull 曲线闭合，等功率圆呈现橄榄球形状的闭合曲线。

当 Smith 圆图上的阻抗远离 R_{OPT} 时，输出功率变小。所以，Smith 圆图上的等功率圆呈现出一组闭合的等高线，如图2.34所示。

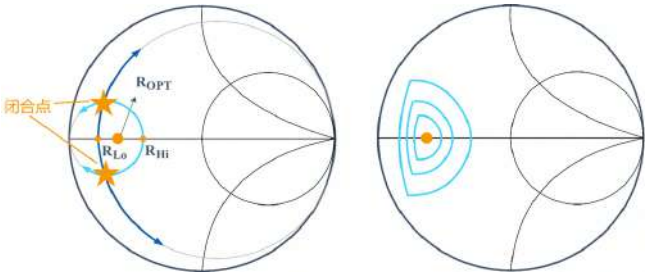


图 2.34: Load-pull 曲线的闭合，以及 Load-pull 的等高线结果

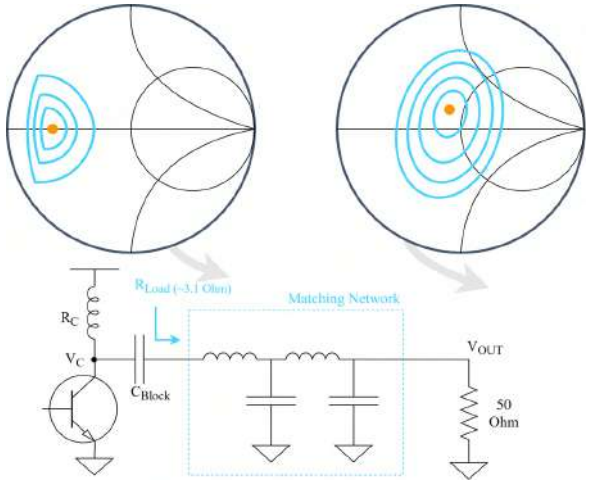


图 2.35: PA 电路中不同平面观测到的 Load-pull 形状不同

2.3.2.3 实际中的 Load-pull

在实际应用中，观测到的 Load-pull 曲线和理论分析曲线可能存在差异，有以下几点需要注意：

- 匹配网络对 Load-pull 的转移
  以上分析均是以 PA 晶体管输出平面计算，由于匹配网络及寄生效应的影响，在芯片输出端口观测到的 Load-pull 可能会有不同。图2.35为不同平面看到的不同 Load-pull 示意图。
- 谐波对 Load-pull 影响
  以上分析中均为简化分析，只考虑基波（Fundamental）阻抗的影响，在 PA 设计中，其他高次谐波，如 2f0、3f0 等阻抗均会对 PA 功率、效率以及线性度产生影响。考虑谐波影响好，Load-pull 形状会有差异。
- 其他指标的 Load-pull
  以上分析针对 PA 中最为重要的指标：功率的 Load-pull 进行分析，PA 的其他指标如线性度等，采用带有带宽的调制信号进行测试，其 Load-pull 形状大致相同。一般不会再针对其他指标进行详细分析。

2.3.3 总结

Load-line 与 Load-pull 是 PA 设计中最重要的两个基础概念，在过去几十年的射频 PA 设计中，前人专家也积累了许多经典的分析方法。

虽然 5G 等高阶通信协议的到来对射频 PA 提出了新的要求，近年来也涌现出如低压 PA、高效率 PA、高/低 Load-line PA 等不同 PA 产品，但射频 PA 的一些基础原理仍然是在 PA 设计中被广泛遵循的，期待和您一起对这些基础原理有更好的理解。

参考文献

1. RF Power Amplifiers for Wireless Communications (second edition), Steve C. Cripps
2. Cripps, S. C. . ”A Theory for the Prediction of GaAs FET Load-Pull Power Contours.” International Microwave Symposium Digest IEEE, 1983.
3. https://www.ebaina.com/
4. https://www.microwavejournal.com/
5. https://www.maurymw.com/

在线阅读

本文摘自慧智微微信公众号，扫描以下二维码，在线阅读

2.4 高效率 PA 设计：从 Class A 到 Class J

作者：彭洋洋 博士

射频 PA（Power Amplifier，功率放大器，简称功放）芯片耗电大，可靠性要求高，是无线通信系统中的重要组成部分。随着 5G 系统频率和功率进一步推高，对射频 PA 的要求也进一步增强。为此，我们整理了一期材料：《5G 射频 PA 架构》，专门讨论在 5G 射频 PA 设计中所采用的主要架构。

在以上材料中，我们讨论到：PA 设计中可以依靠“简单功率合成”来进行功率的合并，也可以利用“特殊功率合成”，达到高效率、负载不敏感等设计需求。材料推出去之后，一些朋友发来问题：提升效率只能依靠架构上的改变吗？一些 PA 没有用到 Doherty 等复杂的架构，依然取得了不错的效率，是怎么做的呢？

这些问题没有一个简单的答案可以回答。于是我们想，那就不如把 PA 在高效率设计中的考量讨论的再细节一些：不止是定性的讨论到宏观架构，而是深入到单个 PA 设计的思考取舍中。

在 PA 设计的细节考量中，最重要的一个概念就是 PA 的“类”别了，也即 PA 的 Class。比如我们经常提到的 Class A、Class AB、Class E 及 Class F 等。不同 Class 的 PA 有不同的特点，图2.36为常见的 PA 类别分类，我们这份材料就以 PA 的 Class 为线，讨论手机射频 PA 的“高效率（High Efficiency）”设计。

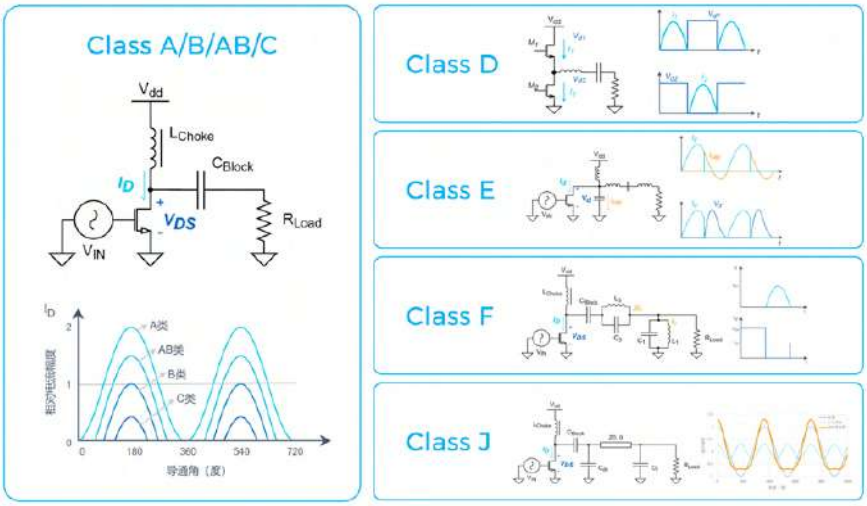


图 2.36: PA 设计的“类”别分类

2.4.1 PA 设计的重要指标

在讨论 PA 的类别之前，首先将 PA 设计中的重要指标做一个探讨。

PA 是用来将直流能量转化射频信号的器件。PA 通常工作在大信号状态，其工作状态与偏置点、负载阻抗、谐波控制、功率的过驱动等都有关系，设计和分析起来非常复杂。

尽管一款性能优异的 PA 设计复杂，但 PA 的评价指标却相对简单，最重要的两个指标就是：线性度，效率。

2.4.1.1 线性度

在现代 PA 设计中，线性度反应输出信号的失真程度，是一个重要指标。

在衡包络通信系统中，如 FM，FSK，GMSK，信号包络并不承载信息，所以并不需要信号的线性放大，“线性度”并不是一个重要指标。但在现代通信系统中，为了最大化的利用频谱资源，出现了如高阶 QAM 等非等幅信号调制方法，也诞生出了 CDMA 及 OFDM 等多工方式，其幅度不再恒定不变，而是承载通信信号信息。图2.37为带有幅度调制信号的 OFDM 信号在时域的波形构成。

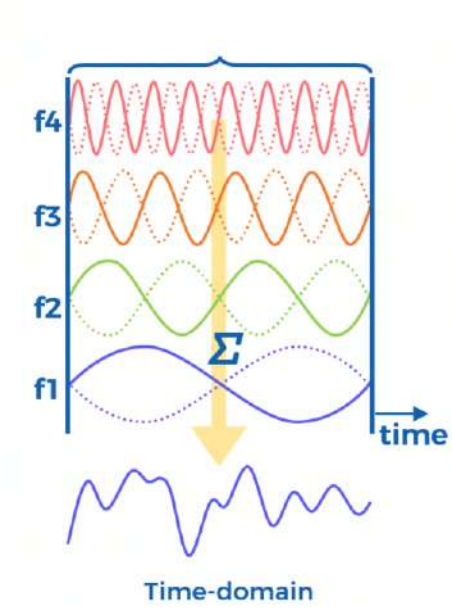


图 2.37: OFDM 信号在时域的波形构成，幅度不再恒定

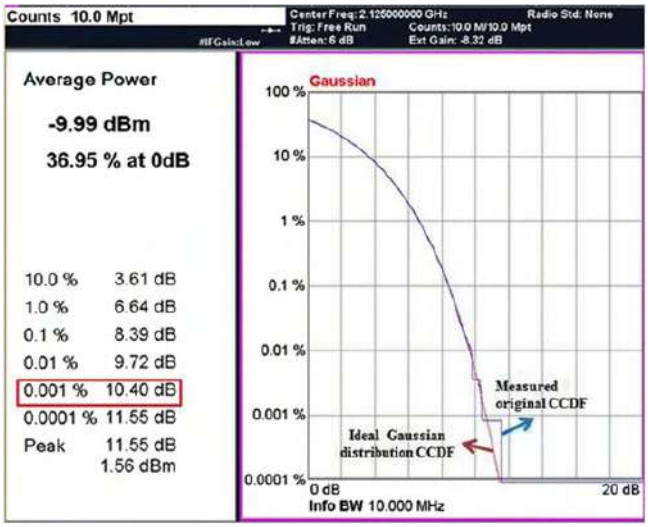


图 2.38: 典型信号的概率分布

在这些信号中，为了表征信号幅度变化的大小，信号“功率峰均比”的概念被引入了进来。峰均比的英文是 Peak to Average Power Ratio，简称 PAPR，是指信号峰值功率与均值功率的比值。一般线性信号的 PAPR 在 3 ~9dB 之间。图2.38为某典型 OFDM 信号能量谱分布，可以看到信号幅度在 0 ~11.55dB 之间呈现概率分布。

由于这些高阶调制方式的信号中的幅度与相位均带有信息，这时必须要对信号进行线性放大，才能使幅度信息不失真的进行传递。若 PA 本身存在非线性，则 PA 对不同大小信号展现出不同的响应，就会产生非线性失真，使 ACLR、EVM 等指标产生恶化，影响系统的通信能力。

2.4.1.2 效率

PA 的本质是把直流能量转化为大功率的射频信号，PA 的效率反应了转换能力的大小。所以“效率”成为 PA 设计中另一个重要指标。

PA 的效率由两种定义方式，分别为功率转换效率（Efficiency），和功率附加效率（Power Added Efficiency）。

功率转换效率的定义是输出功率与消耗直流之比，通常用 η 表示，计算公式为：

$$\eta = \frac{P_{out}}{P_{DC}}$$

PAE 的定义将输入功率也考虑进来，只计算经过 PA 后“增加的”功率部分与直流耗散之间的比值，所以被称作功率“附加”效率。PAE 的计算公式为：

$$PAE = \frac{P_{out} - P_{in}}{P_{DC}}$$

当 PA 增益足够大时，P_{out} 将远大于 P_{in}。比如当 PA 增益为 30dB 左右时，P_{in} 只有 P_{out} 的千分之一。这时 PAE 与功率转换效率的计算结果基本相同。

PA 的效率和线性度之间是一对折中。在 PA 设计中，不断的对这两个指标进行优化或者取舍。

效率优化的理念

根据效率定义的公式可以看出，如果直流能量 PDC 被完全转化为了基波射频能量 P_{out}，则 PA 将达到梦寐以求的 100% 的效率，能量一点也不被浪费。

但这个 100% 的效率是不可达成的，因为在实际电路中，一定会有一部分直流能量转换为谐波射频能量（$P_{out,nf}$）以及直流耗散能量（P_{diss}）。三种能量与直流能量的关系如下公式所示：

$$P_{DC} = P_{out,f} + \sum_{2}^{\infty} P_{out,nf} + P_{diss}$$

高效率 PA 的设计理念就是：尽量使得以上公式的后两项：直流耗散能量、谐波耗散能量，趋近于零。

直流耗散能量（P_{diss}）趋向于 0

直流耗散能量的计算方法为流过晶体管的电压与电流的乘积对时间的积分，其公式可以表示为：

$$P_{diss} = \frac{1}{T}\int_{0}^{T} v(t) \cdot i(t) dt = 0$$

若使此能量等于零，则时域表现为二者之间在时域尽量少的交叠。以上设计观念在高效率 PA 中经常被采用，尤其是在 F 类、E 类等开关类 PA 设计中。

谐波能量（$P_{out,nf}$ 趋向于 0

PA 是一个大功率工作器件，一般 PA 都会工作在带有一定非线性的工作状态，这时就会产生谐波能量。

对于 n 次谐波，其能量可以表示为：

$$P_{out,nf} = \frac{1}{2} \cdot V_n \cdot I_n \cdot cos(\phi_n)$$

以上公式中，V_n 与 I_n 分别为 n 次谐波的电压与电流摆幅幅值，ϕ_n 为其相位。

为了达到 100% 的效率，需要控制能量不被谐波能量所消耗。如果有谐波能量，即使直流消耗为 0，也无法达到 100% 的效率。一个典型的例子是当 PA 的电压与电流均是完美方波的情况下，电压与电流无交叠，直流耗散 P_{diss} 为 0。但此时电压与电流均存在奇次谐波能量，奇次谐波能量不为 0。通过傅里叶展开可计算，此时 PA 的最大效率只有：

$$\eta = \frac{8}{\pi^2} \approx 81.1\%$$

即在电压电流均是完美方波的情况下，谐波部分抢走了近 1/5 的能量。图2.39为此时电压与电流的时域波形，以及能量在频域的分布。

为了减少直流耗散能量、谐波耗散能量这两大损耗，在高效率 PA 中不断要考虑电压与电流波形之间的关系，对波形进行塑造，提升 PA 效率。这种工作在 PA 设计中被称为“波形工程（Waveform Engineering）”。

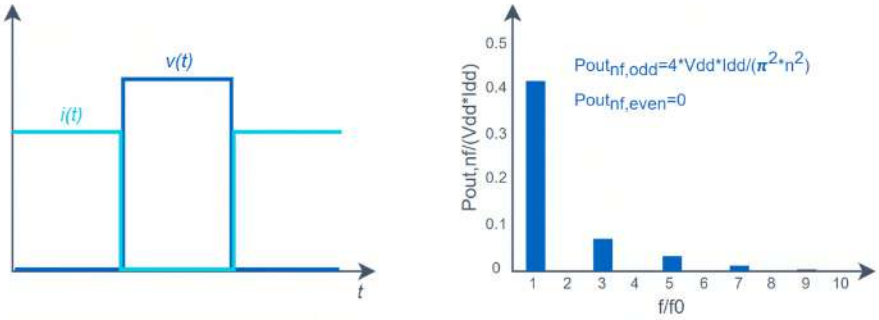


图 2.39: 电压与电流均为方波时，时域波形与频域能量分布

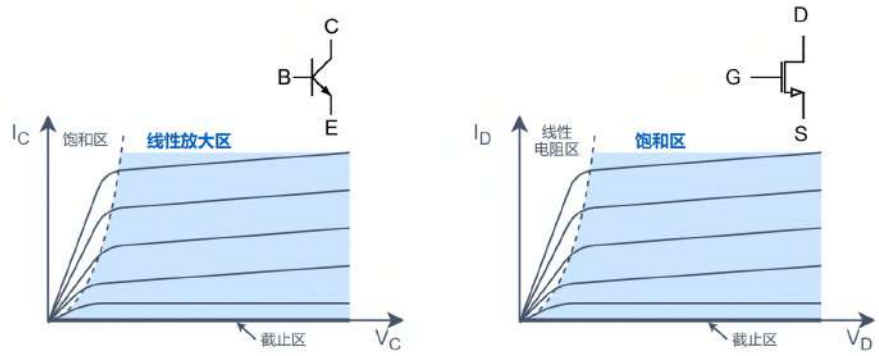


图 2.40: BJT 与 FET 器件用于实现线性 PA 的工作区域

2.4.2 PA Class 的分类

为了优化 PA 的线性度和效率，在射频 PA 先驱者过去 100 多年的工作中，根据工作类型的不同，将 PA 分为不同的类型（Class）进行分析和设计。

射频 PA 设计中所涉及的主要类别有：

1. A 类
2. B 类
3. AB 类
4. C 类
5. D 类
6. E 类
7. F 类
8. J 类

2.4.2.1 线性与开关类 PA

以上类别中，可以根据工作方式的不同，分为线性 PA 类与开关 PA 类。

线性 PA 类是指利用晶体管线性区特性，实现功率放大的放大类型。在这个工作区域内，**晶体管的输入和输出成大致线性的转移关系**，输出信号幅度的大小可以反应输入信号幅度的大小。

需要注意的是，对于 BJT 与 FET 的工作原理不同, 二者“线性区域”有不同的称呼方法。在 BJT 器件中，这个区域被称为线性放大区，或线性区；而在 FET 器件中，这个区域被称作饱和区，如图2.40所示。

线性 PA 的主要类型有 A 类、B 类、AB 类和 C 类等。

需要说明的是，**线性 PA 并不是完全没有非线性**。对于除 A 类外的 AB 类、B 类和 C 类线性放大器，因为波形的非完整周期导通，输出均存在非线性分量。即使对于全周期导通的 A 类放大器来说，由于晶体管输入和输

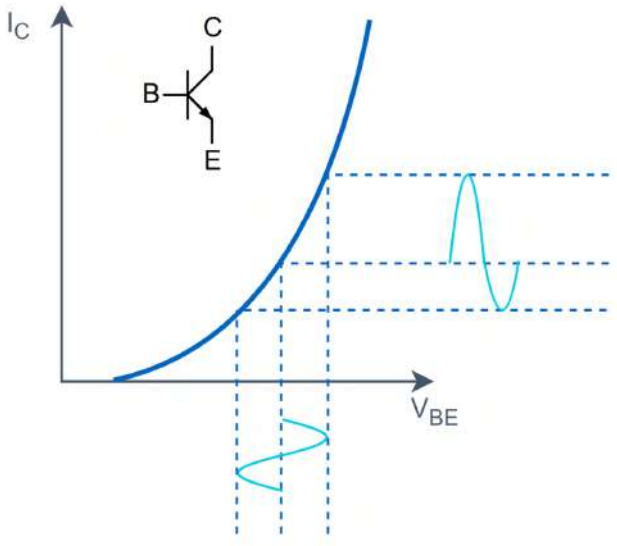


图 2.41: 晶体管的转移特性

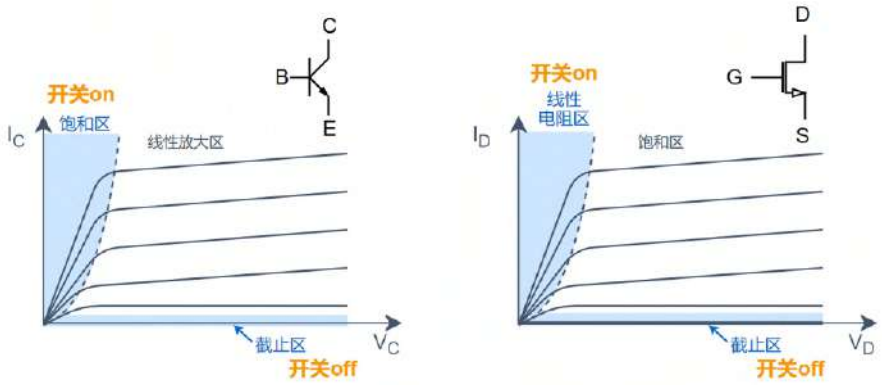


图 2.42: 用于实现开关 PA 的工作区域

出的非线性转移关系（指数或者二次方关系），输出也会出现非线性分量，如图2.41所示。但这些非线性分量的存在，不影响晶体管工作在线性工作区的实质，也不影响输出信号幅度与输入信号幅度呈正向比例的相对关系，所以这些 PA 都属于线性 PA。

因为理解直观，设计容易，线性 PA 是放大器设计中的首要选择。但线 PA 很难做到高效率，于是 PA 设计先驱者开始将开关类 PA 的设计理念引入进来。图2.42为用于实现开关 PA 的工作区域。

开关类 PA 类型主要有 D 类、E 类、F 类和 J 类等，其特点是晶体管工作在类似开关状态。在这些 PA 设计中，对 PA 的设计考虑不再局限于从 A/B/AB/C 类放大器中偏置的角度讨论，而是把输出端负载对波形的调制影响也考虑进来。

“高效率”是开关类 PA 的传统优势，但其线性度较差。并且，进入到高频之后，由于寄生效应的存在，并不能实现完美的开关特性，谐波控制也变的困难，这些都给开关类 PA 的设计带来挑战。

2.4.2.2 线性类 PA

线性类 PA 简化电路如图2.43所示。

根据晶体管导通角的不同，可以将线性 PA 分为 A 类、AB 类、B 类和 C 类。不同类型 PA 的偏置状态与导通角的关系如图2.44所示。

A 类 PA

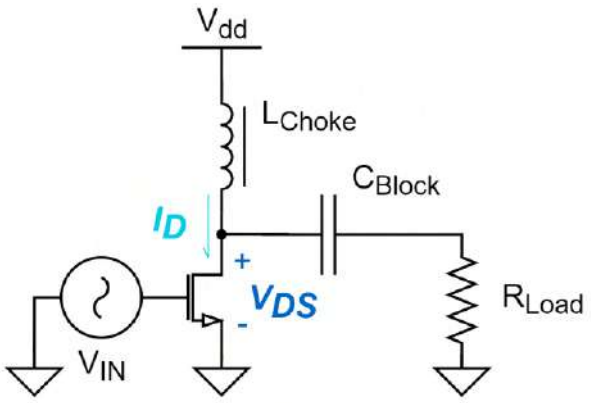


图 2.43: 线性类 PA 简化电路示意图

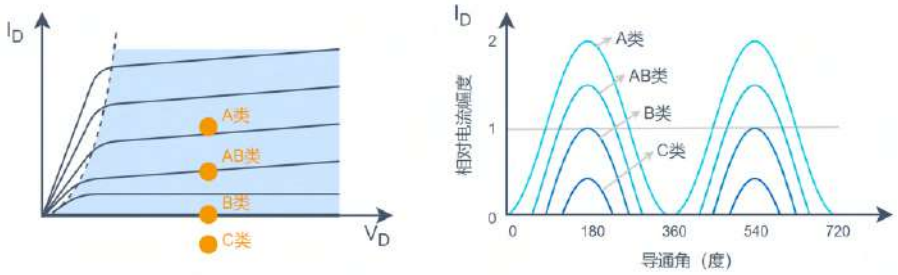


图 2.44: 各类线性放大器偏置状态与导通角的关系

A 类 PA 的晶体管在信号的全周期导通。其电压与电流的波形如图2.45所示。

假设 A 类 PA 的直流工作点为 VDC 与 IDC，射频信号摆幅为 VRF 与 IRF，则根据 PA 效率的定义，A 类 PA 的效率可表示为：

$$\eta = \frac{P_{RF}}{P_{DC}} = \frac{\frac{1}{2} \cdot V_{RF} \cdot I_{RF}}{V_{DC} \cdot I_{DC}} = \frac{1}{2} \cdot \frac{V_{RF} \cdot I_{RF}}{V_{DC} \cdot I_{DC}}$$

A 类 PA 最大效率在 VRF 与 IRF 达到最大摆幅 VDC 与 IDC 时取得，为 50%。

在输出功率变小时，A 类 PA 射频功率变小，但 DC 功耗不变，使得其效率明显下降。在输出功率下降至最大 A 类输出功率的一半时，其效率也将下降一半。即在功率回退时，其效率以 2 倍/3dB（10 倍/10dB）的速度下降，如图2.46所示。

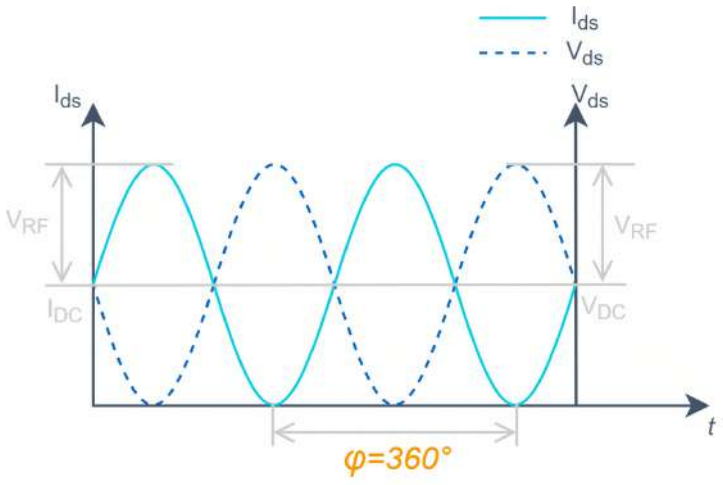


图 2.45: A 类 PA 的电压与电流波形

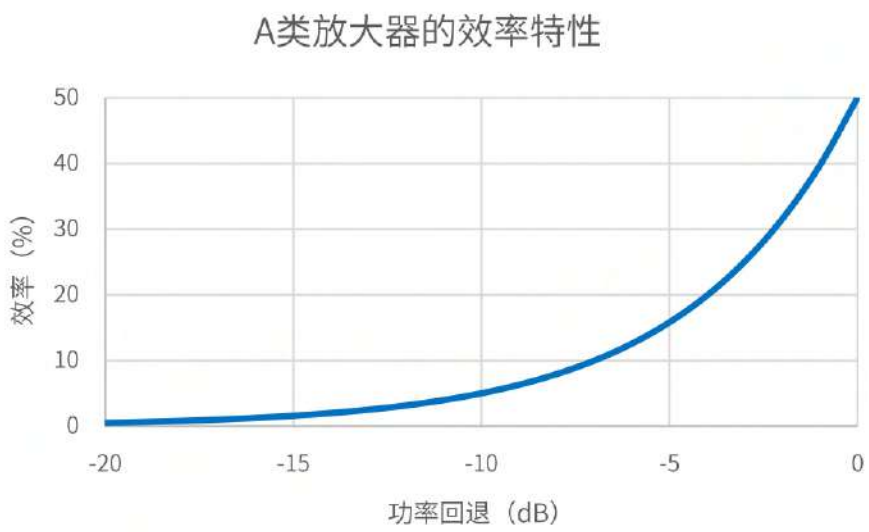


图 2.46: A 类 PA 功率回退时的效率特性

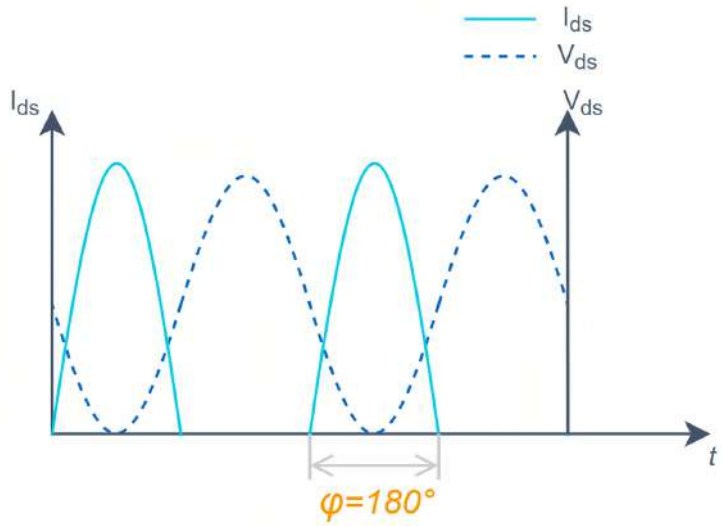


图 2.47: B 类 PA 的电压与电流波形

B 类 PA

B 类 PA 的直流偏置点放置在晶体管导通开启电压附近，只有当输入信号为正向摆幅时晶体管导通，导通角只有全周期的一半。B 类 PA 的电压与电流波形如图2.47所示。

虽然输出电流只有半周期导通，波形出现了强烈的失真，但由于器件仍然工作于线性工作区域，输入信号的幅度仍然基本正比于输入信号的幅度，所以 B 类 PA 仍然是线性类 PA。

分别对基波和 DC 进行积分，可以推导得出导通角为 2θ（2θ=φ）的 PA 最大效率表达式为：

$$\eta = \frac{P_{RF}}{P_{DC}} = \frac{2\theta - sin(2\theta)}{4[sin\theta - \theta \cdot cos\theta]}$$

对于 B 类 PA，$2\theta = \pi$，则上式推导为：

$$\eta = \frac{2\theta - sin(2\theta)}{4[sin\theta - \theta \cdot cos\theta]} = \frac{\pi - sin(\pi)}{4 \cdot [sin(\frac{\pi}{2}) - \frac{\pi}{2} \cdot cos(\frac{\pi}{2})]} = \frac{\pi}{4} \approx 78.5\%$$

所以相比于 A 类放大器 50% 的最高效率，B 类放大器的最高效率可提高至 78.5%。

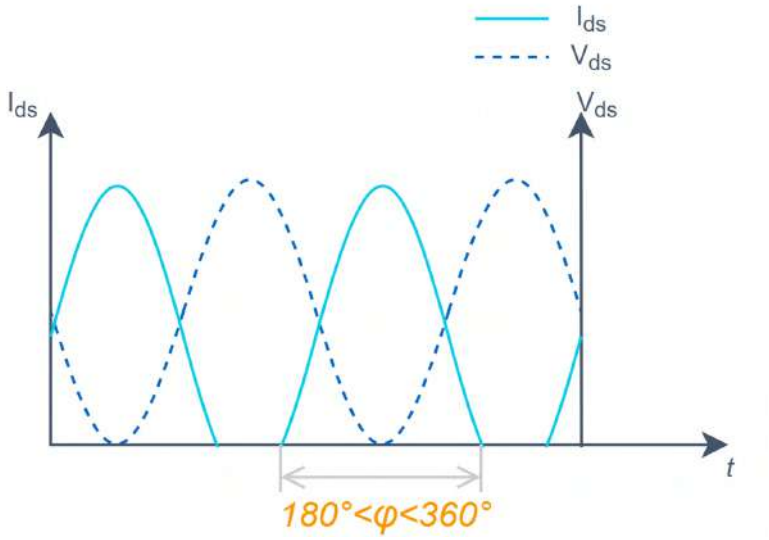


图 2.48: AB 类 PA 的电压与电流波形

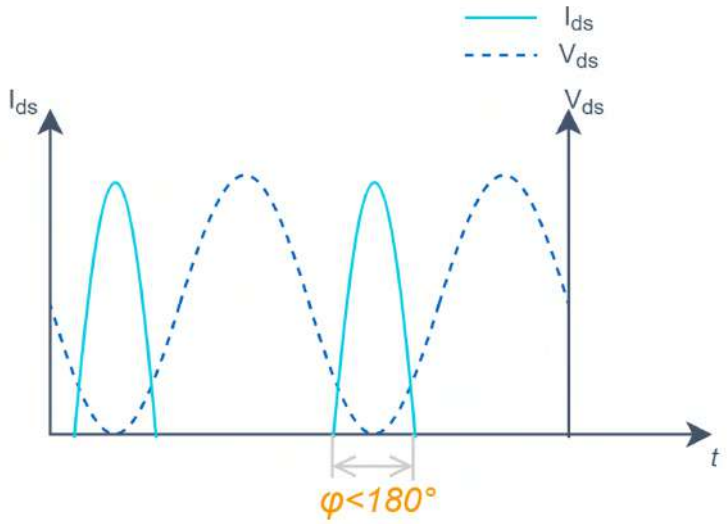


图 2.49: C 类 PA 的电压与电流波形

AB 类 PA

AB 类 PA 是指处于 A 类和 B 类之间的 PA 类型，导通角为在 π 与 2π 之间，其电压与电流波形如图2.48所示。

AB 类 PA 的效率根据导通角的不同而不同，处于 A 类的 50% 与 B 类的 78.5% 之间。

C 类 PA

通过带有导通角的效率计算公式可以看到，当导通角变小时，PA 的效率提升。在 B 类 PA 的基础上继续减小导通角，全周期的导通角减少至 180 度以下时，就形成了 C 类 PA。C 类 PA 的电压与电流波形如图2.49所示。

对于 C 类 PA，将导通角取 0 为极限，可以得到 PA 的极限效率可达 100%。

虽然理论上 C 类 PA 可以达到 100% 的效率，但达到 100% 效率时导通角为 0，此时也无功率输出。所以 C 类 PA 中的 100% 效率是可望而不可及，无法为实际使用的 PA 设计提供有效帮助。

线性类 PA 比较

由于导通角的不同，A 类、AB 类、B 类及 C 类 PA 在效率、基频功率，以及谐波功率上有不同的表现。采用傅里叶变换可以对几种不同类型的 PA 进行分析，得到几种类型的 PA 特性如图2.50所示 [8]：

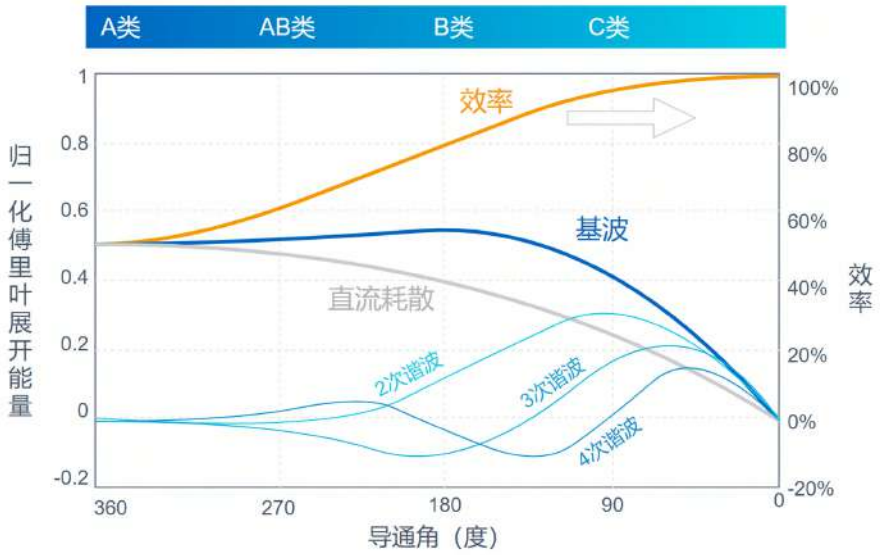


图 2.50: 不同导通角下 PA 的能量分布及效率特性

依靠上图，可以很好的理解不同类型的线性 PA 的特性关系，可以看到：

- 随着导通角变小，PA 效率由 50
- 从 A 类到 B 类的变化中，导通角的变小并没有引起基波能量的减小
- 从 B 类到 C 类变化中，基波输出功率迅速减小

正是由于以上特性，应用中线性 PA 大多设计在 AB 类，这时有较好的效率，较高的基波输出功率，同时也有可接受的谐波特性。

2.4.2.3 开关类 PA

在理想开关类 PA 中，当输入电压为正，晶体管打开，电流通过晶体管，此时晶体管两端电压为 0；当电压为负，晶体管关闭，此时输出电压开始建立，但流过晶体管的电流为 0。

因为作为开关使用的晶体管不消耗功耗，所有的能量都可以转化为射频能量，所以理论上开关 PA 可以达到 100% 的效率。

然而在实际使用中，由于开关的非理想特性和器件的寄生效应，晶体管并不能达到理想开关状态，也就造成了电压与电流的部分交叠，造成效率的损失。另外，如果谐波能量处理不当，也会引起能量损失。

为了使开关类 PA 的效率不断逼近 100%，PA 设计先驱们采用了多种不同的设计分析方法，于是就产生出多种不同的开关 PA 类型。在射频 PA 中，最主要的有 D 类、E 类、F 类以及 J 类 PA。

D 类 PA

D 类 PA 是 1959 年由 Baxandall 教授首先提出 [2]，其构成为一个成对的开关管 M1 及 M2，以及谐振在基波频率的调谐负载。如图2.51所示。

理想情况下，D 类 PA 晶体管的电压与电流无交叠，使得 D 类 PA 直流功率耗散为 0。

对于谐波功率耗散，D 类 PA 的输出电压波形为方波信号；由于谐振负载的存在，输出电流波形为半正弦波。对方波信号进行傅里叶展开，所有的偶次谐波分量为 0；对半正弦波进行傅里叶展开，所有奇次谐波分量为 0。谐波分量中电压与电流交替为 0，使得理想 D 类 PA 的无谐波功率耗散。

以上这种电压与电流分别为方波与半正弦波时，在谐波功率耗散为 0 的特性，在高效率 PA 设计中经常被使用。

由于直流功率耗散与谐波功率耗散均为 0，所以理想 D 类放大器可以达到 100% 的效率。

D 类 PA 也有一系列变形，例如可以将电流构建成方波，将电压构建成半正弦波。这种就是电流模式的 D 类

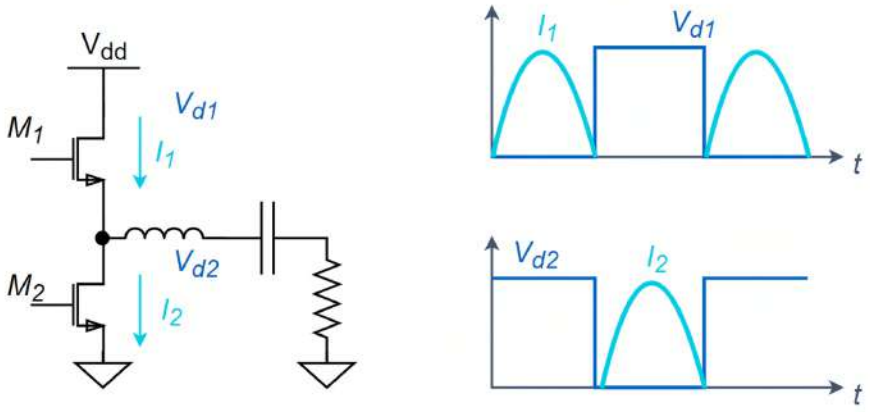


图 2.51: D 类 PA 工作原理及波形

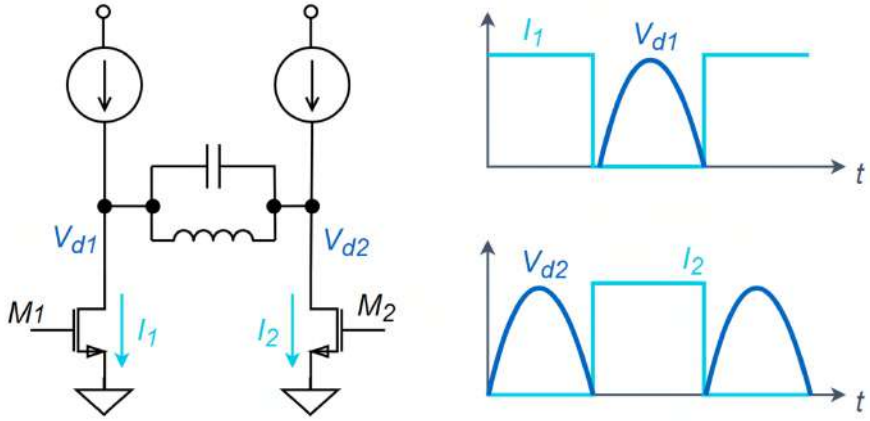


图 2.52: 电流模式的 D 类放大器

放大器。其简化电路及波形图如图2.52所示。

E 类 PA

D 类 PA 看起来可以完美实现 100% 的转换效率，但其高度依赖两个开关管的完美切换，在高频应用时显的无能为力：感性的负载会使开关“关不断”，容性负载会使开关“打不开”。并且 D 类 PA 需要的成对推挽结构，也使得其在部分高频应用时力不从心。

为了使开关类 PA 可以应用在高频频率，Sokal 教授在 1975 年发明了 E 类 PA[3]。E 类 PA 由工作于开关状态的单端晶体管、负载匹配网络构成。其架构如图2.53所示。

E 类 PA 的架构看起来与普通的 A 类、AB 类 PA 并无大的不同，但设计思路却相差很大。E 类 PA 的设计理念是：

1. 首先将晶体管的偏置和驱动功率进行合理设计
2. 然后对输出波形提出一系列约束条件，这些约束条件使晶体管工作于开关状态

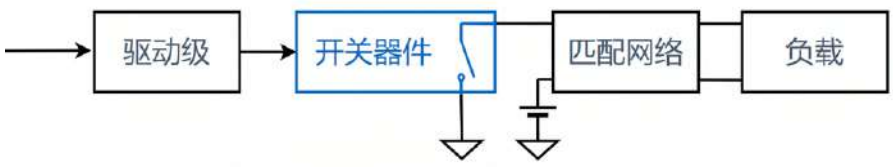


图 2.53: E 类放大器框图

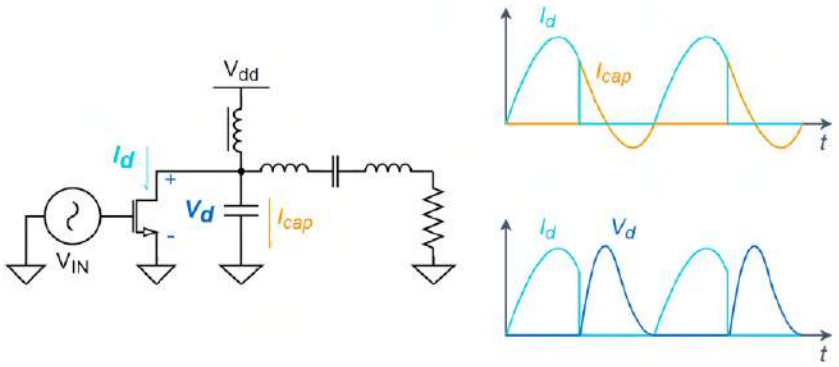


图 2.54: E 类放大器的简化电路，以及波形示意图

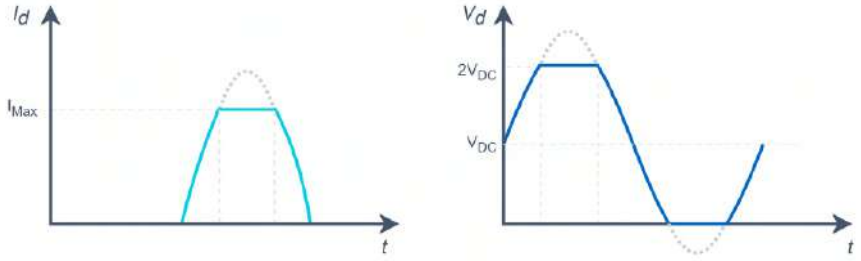


图 2.55: 理想过激励下 B 类放大器的电流与电压波形

3. 基于这些约束条件，就可以计算不同拓扑中器件取值，从而完成设计

E 类 PA 约束条件使作为开关等效的晶体管在合适的时间进行 off 到 on 的切换，从而减小开关切换过程中带来的充放电损耗。Sokal 教授提出 E 类 PA 设计的两个重要约束条件：

- 当开关从 off 到 on 转换的瞬间，漏极电压为零（Zero voltage switching，简称为 ZVS）
- 当开关从 off 到 on 转换的瞬间，漏端电压波形的斜率为零（Zero voltage derivative switching，简称 ZVDS）

E 类放大器的简化等效电路与波形示意如图2.54所示。

F 类 PA

F 类 PA 的发明源自对过激励的 B 类 PA 的研究。

在线性 PA 分析中，B 类放大器的最大效率为 78.5%，但如果对其进行过激励驱动，其电压、电流波形出现削峰，形成类似方波的波形，这时电压电流交叠变少，从而效率得到提高。如图2.55所示。

过驱动的 B 类放大器虽然有更高的效率，但由于电压电流均都包含奇次和偶次谐波，有部分能量仍然消耗在了谐波上，所以效率无法达到 100%。通过计算，过驱动的 B 类放大器可以达到 88.6% 的效率。这就是为何有些 B 类放大器能看到高于理论值 78.5% 的峰值效率的原因。

基于对过驱动的 B 类 PA 的研究，D. M. Snider 教授于 1967 年提出 F 类 PA 的概念 [4]，随后 F. H. Raab 等教授将 F 类 PA 的原理和设计方法进行了更进一步阐述 [5][6][7]。

F 类 PA 通过对谐波阻抗的控制，使奇次谐波阻抗为开路，偶次谐波阻抗短路，于是得到方形的电压波形，和半正弦的电流波形。由于二者分别只含有奇次和偶次谐波，谐波消耗为 0，在理想情况下可得到 100% 的效率。图2.56为理想 F 类 PA 的电压电流波形。

在实际应用中，由于谐波频率高，无法对各次谐波阻抗均进行完美控制，这在一定程度上恶化了 F 类 PA 的实际应用效果。图2.57为对 3 次谐波进行控制的 F 类 PA 简化电路 [5]。

同理，也可以将电流构造成方波，电压构造成半正弦波，同样可以达到理论上 100% 的效率，这样就形成了逆 F 类 PA 的设计。逆 F 类 PA 设计的条件与 F 类放大器相反：偶次谐波阻抗需要设计成开路，奇次谐波阻抗需要设计成短路。如图2.58所示。

虽然 F 类和逆 F 类在理论上都可以达到 100% 的峰值效率，但在实际应用中，二者表现还是会稍有不同。考

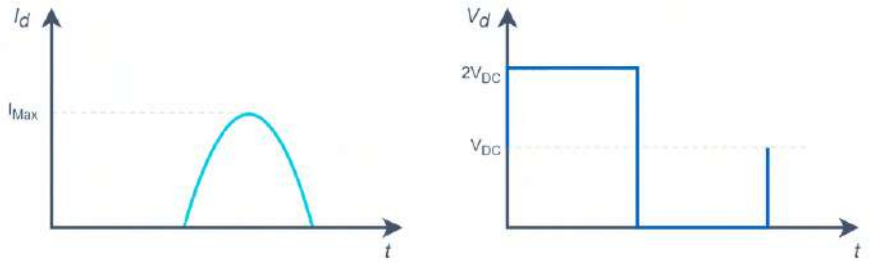


图 2.56: 理想 F 类 PA 的电流及电压波形

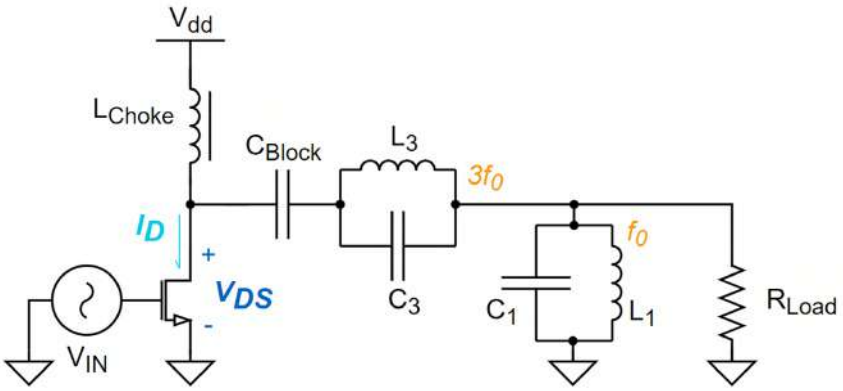


图 2.57: 典型的 F 类 PA 电路

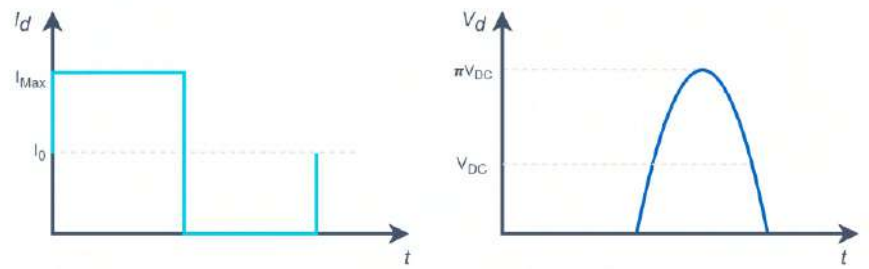


图 2.58: 理想逆 F 类 PA 漏极电流及电压波形

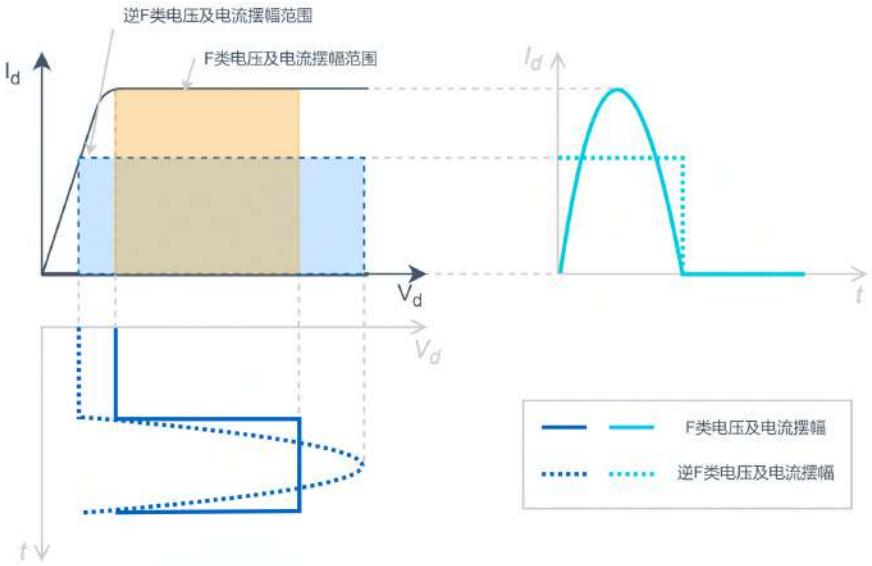


图 2.59: F 类与逆 F 类 PA 电压及电流摆幅关系比较

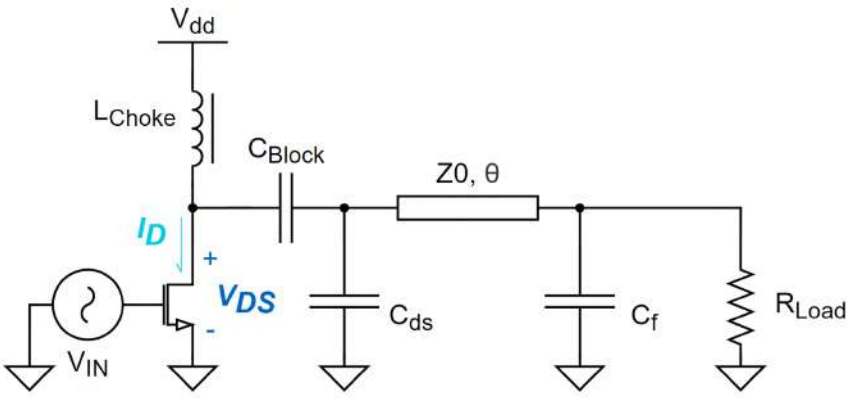


图 2.60: J 类 PA 简化电路

虑晶体管 Knee 电压之后，F 类和逆 F 类 PA 的工作区域如下图所示，从图2.59对比可以推导出系列 F 类与逆 F 类 PA 之间性能比较的关系。以上推导在本文中不再赘述。

J 类 PA

J 类 PA 是 Cripps 教授于 2006 年提出的一种高效 PA 设计方法，其设计思路仍然是利用谐波阻抗的控制，实现电压与电流交叠部分的减少，从而减少直流损耗，提升 PA 效率。图2.60为 Cripps 教授提出的 J 类 PA 的简化电路图 [8]。

与 F 类通过控制不同次谐波的幅度来构建完美的方波与半正弦波不同，Cripps 教授的想法是通过控制二次谐波的相位，使二次谐波与基波之间形成叠加，减小电压与电流之间的交叠。如图2.61所示。

2.4.3 不同类型 PA 的比较

2.4.3.1 相对功率输出力

在 PA 设计中，评价不同结构的特点时，除了效率、功率指标，还有一个重要指标，称为“归一化功率输出能力”，英文为 Normalized Output Power Capability，一般简写为 P_{NOPC}。其定义为：功率放大器的最大输出功率和最大瞬态工作电流与电压乘积之比。公式可以表示为：

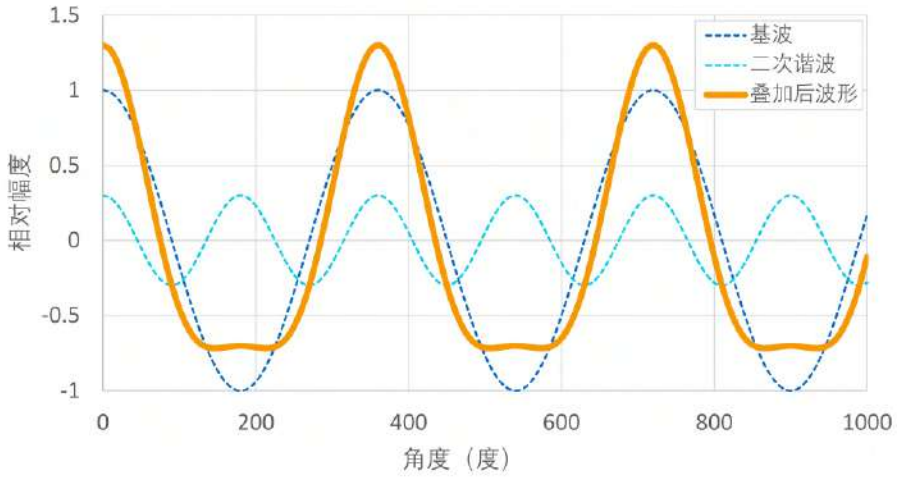


图 2.61: J 类 PA 电压波形示意图

$$P_{NOPC} = \frac{P_{out,Max}}{I_{D,Max} \cdot V_{DS,Max}}$$

归一化功率能力指标可以反应在给定的电压与电流摆幅的情况下，功率放大器输出最大功率的大小；也可以反应在给定输出功率的情况下，所需要电压与电流摆幅的大小。

对于 A 类放大器，其归一化功率输出能力可计算为：

$$P_{NOPC,ClassA} = \frac{P_{out,Max}}{I_{D,Max} \cdot V_{DS,Max}} = \frac{\frac{1}{2} \cdot V_{DQ} \cdot I_{DQ}}{2I_{DQ} \cdot 2V_{DQ}} = \frac{1}{8} = 0.125$$

同理，可计算出其他类型 PA 的相对功率输出能力。

2.4.3.2 不同类型 PA 性能对比

将以上讨论到的 PA 特性进行总结，得到性能对比如图2.62所示。

2.4.4 总结

在很多人看来，PA 设计像是“黑魔法”，难以捉摸。

说 PA 设计是“黑魔法”，是因为在很多时候，PA 的状态看起来是不受控的。甚至有时 PA 的性能已经出来了，Designer 都说不清它是怎么工作的。

PA 虽然看起来这么的不可控，但它并不是“魔法”，它也是严格遵循物理规律的。在过去 100 多年发展中，PA 设计的先驱和专家们不断想办法对它进行理解。“不同的 Class”分类，是先驱专家们总结出理解 PA 特性的经典方法，在高效率 PA 设计中得到有效使用。

本文对常见射频 PA 的类别进行简单总结，希望对大家在高效率 PA 的设计和应用中有所帮助。

本文整理过程中得到众多行业专家的帮助，在此表示诚挚感谢！

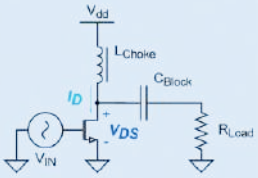
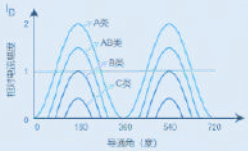
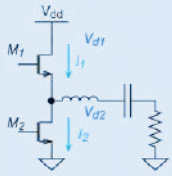
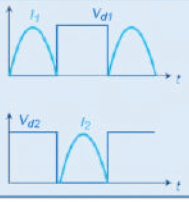
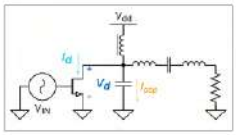
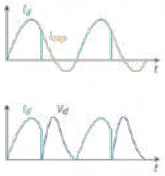
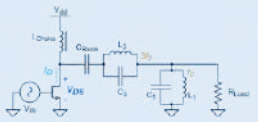
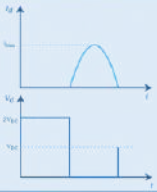
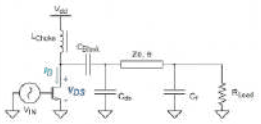
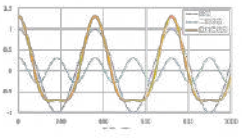
类型	架构示意	波形示意	理论效率	归一化功率输出能力
A			50%	0.125
AB			50~78.5%	0.125
B			78.5%	0.125
C			78.5~100%	0.125~0
D			100%	0.318
E			100%	0.098
F			100%	0.159
J			78.5%	0.086

图 2.62: 不同类型 PA 特性对比

参考文献

1. Anoh K, Adebisi B, Rabie K, et al. On Companding and Optimization of OFDM Signals for Mitigating Impulsive Noise in Power-Line Communication Systems[J]. IEEE Access, 2017.
2. Baxandall P J. Transistor sine-wave LC oscillators. Some general considerations and new developments[J]. Proceedings of the IEE - Part B: Electronic and Communication Engineering, 1959,106:748-758.
3. N. O S, A. D S. Class E-A new class of high-efficiency tuned single-ended switching power amplifiers[J]. IEEE Journal of Solid-State Circuits, 1975,10(3):168-176.
4. Snider D M. A theoretical analysis and experimental confirmation of the optimally loaded and overdriven RF power amplifier[J]. IEEE Transactions on Electron Devices, 1967,14(12):851-857.
5. Raab F H. Class-F power amplifiers with maximally flat waveforms[J]. IEEE Transactions on Microwave Theory and Techniques, 1997,45(11):2007-2012.
6. Raab F H. Maximum efficiency and output of class-F power amplifiers[J]. IEEE Transactions on Microwave Theory and Techniques, 2001,49(6):1162-1166.
7. Raab F H. Class-E, Class-C, and Class-F power amplifiers based upon a finite number of harmonics[J]. IEEE Transactions on Microwave Theory and Techniques, 2001,49(8):1462-1468.
8. Cripps S C. Rf power amplifier for wireless communications[M]. RF power amplifiers for wireless communications /, 2014.

在线阅读

本文摘自慧智微微信公众号，扫描以下二维码，在线阅读

2.5 5G 射频 PA 架构

作者：彭洋洋 博士

自 2019 年 5G 元年开始，过去 3 年 5G 建设如火如荼的进行。5G 快速发展中，受益最大的就是射频前端芯片。根据 Yole 的预测，至 2026 年，全球射频前端总市场将达到 216.7 亿美金，与 2019 年的 124.1 亿美金相比，7 年增长率达 74.6%。

射频前端芯片是无线通信的核心器件，是指天线之后、收发机之前的功能模块，因为位于通信系统的最前端，所以被称为“射频前端”，如图2.63所示，射频前端一般包含功率放大器、滤波器/双工器、开关以及低噪声放大器。

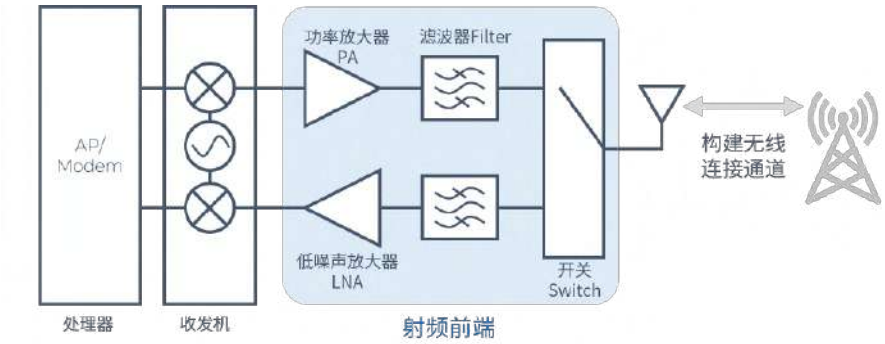


图 2.63: 射频前端的构成及功能

射频 PA（Power Amplifier，功率放大器）是射频前端中的核心器件，其性能直接决定信号的强弱、稳定性、功耗等重要因素，决定用户体验。其核心参数包括增益、带宽、效率、线性度、最大输出功率等，众多平衡的性能指标非常考验设计能力，如图2.64所示。

随着 5G 的发展，对射频 PA 的技术性能需求再次拉升，需要 PA 有更高的工作频率、更高的功率、更大的带宽，同时模组化的到来也需要 PA 设计满足高集成度模组化的需求。

于是，不同架构设计的 PA 被大家关注起来，除了日常听说的单端 PA，近年来手机应用中还出现了如 Push-pull PA、Balance PA、Doherty PA 等等多种 PA 架构。这些 PA 架构分别指的是什么？它的原理是什么？应用时有什么特点？本文将对以上问题进行讨论。

2.5.1 PA 的设计理念

PA 是 Power Amplifier（功率放大器）的缩写，是一类主要用于功率放大输出的放大器类型。与其他放大器，如低噪声放大器、驱动放大器不同，功率放大器的主要目的是**从直流榨取出尽量多的射频信号**，而不是仅仅关



图 2.64: 射频 PA 在芯片设计中的位置

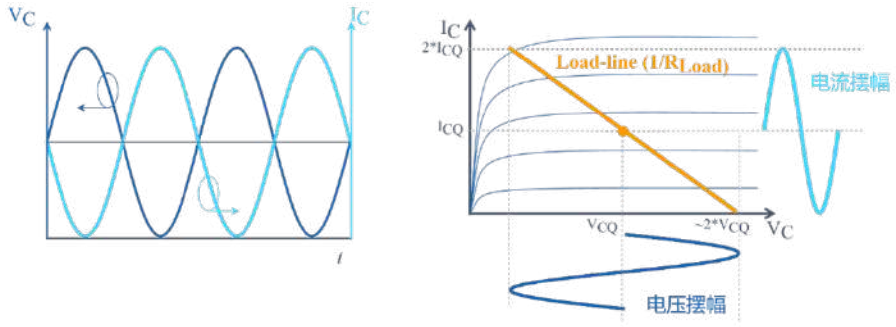


图 2.65: 负载线理论与电压/电流摆幅

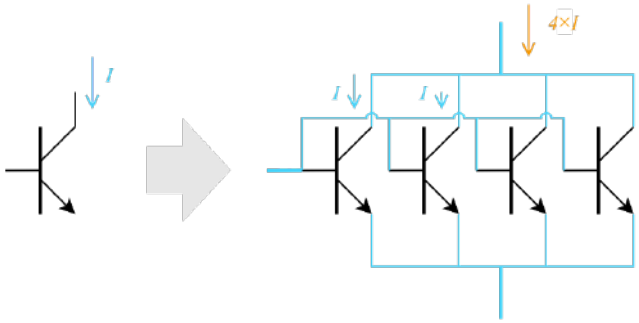


图 2.66: 功率放大器的电流合成

注增益。

为了达到足够高的功率输出，PA 的输出放弃了用来最大信号“传输”的共轭匹配，而采用了可以实现最大功率“输出”的负载线匹配。另外，为了实现可控的直流消耗，PA 设计必须考虑效率优化设计。

负载线匹配使得 PA 的电压与电流摆幅均达到最大，也即达到了最大功率输出的目标。有关 PA 负载线设计的理念，可参考文章《5G 射频 PA 的 Load-line 与 Load-pull》，如图2.65所示。

以上仅为单个 PA 单元的设计思路与方法，在实际应用中，有可能单个器件的电流、电压摆幅不足以满足 PA 整体指标需求；也有可能需要结合架构设计，实现改善效率、改善驻波的需求。这时，就需要在 PA 架构侧做改进。

2.5.2 PA 的架构设计

PA 的核心目标就是“功率”，不同 PA 的架构也是围绕“功率合成”这一理念进行的。在功率合成中，根据功率合成类型的不同，可以将功率合成分为“简单功率合成”与“特殊功率合成”。

“简单功率合成”是指将多个小功率器件进行合成，直到合成至足够大的功率。简单功率合成的方式有：电流合成、电压合成、功率合成。

“特殊功率合成”是指利用较为特殊的合成方法，在合成的时候完成一些特殊的特性设计。常见的合成方法有 Push-pull（推挽）、Balance（平衡）、Doherty（多尔蒂）等。

以下将对几种 PA 架构的设计进行讨论。

2.5.2.1 简单功率合成

电流合成

电流合成是最简单的功率合成方式，实现方式是将多个较小的器件进行并联连接，将电流进行并联，如图2.66所示。

在实现上，电流合并即在 Layout 中将多个晶体管的基极、集电极、发射极分别连接，由于实现简单，在功率放大器的设计中被广泛采用。

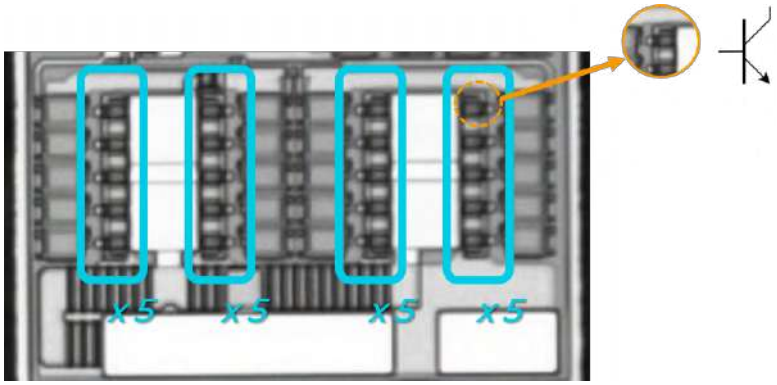


图 2.67: 典型功率放大器功率级版图

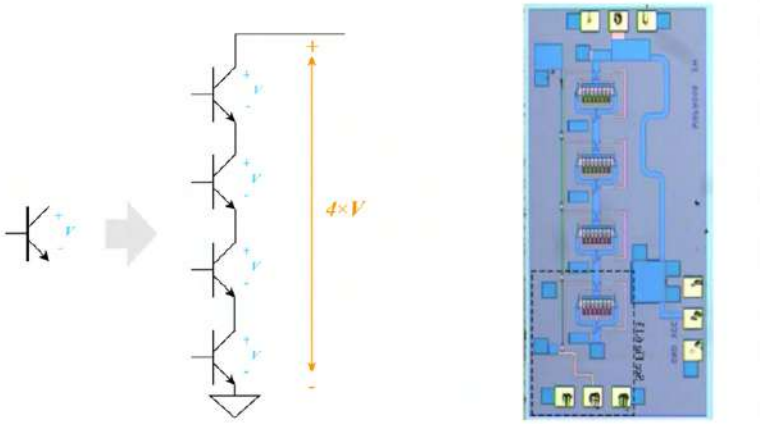


图 2.68: 采用电压合并的典型电路，以及典型实现案例

图2.67为典型的 GaAs 功率放大器功率级版图，图中末级的功率输出级由 4 列功率阵列（Power Array）构成，每个功率阵列包含 5 个功率单元（Power cell），共计 20 个功率单元，一起完成末级功率放大输出。

虽然电流合并实现简单，但在设计的时候也有许多要点需要注意：

- 单个功率阵列不易过长，以确保电流是同相叠加
- 不同阵列之间合并时要注意走线的对称性，保证电流合成时的相位相同
- 用于电流合并的走线较宽，要注意合理设计，减少走线带来的寄生电容效应

电流合并由于设计简单，实现方便，在功率放大器中被广泛使用。

电压合成

除了电流合并方式外，还有一种简单合并方式是将电压进行合并。电压合并的优点是可以提高最优输出阻抗点，在做阻抗匹配时可以实现更低的阻抗匹配损耗。

图2.68为电压合并方式的典型实现电路，以及实现示例 [1]。

电压合并的方式适用于供电电压远大于器件耐受电压的场景，如有高压供电环境；或者采用低压器件进行 PA 设计时。

当前手机 PA 设计所使用的 GaAs HBT 击穿电压 VBCEO 一般在 10 ~25V 之间，单个器件适用于 5V 以内的直流 Vcc 偏置电压。在 2016 年，业界公司尝试将电池电压由 3.8V 先升至 11V 左右，再供电给 Cascode 电压合并的 PA 设计，以达到提高输出最优阻抗，进而提升 PA 整体效率的设计目的。此方案在 2016 年前后的一些旗舰手机中得到了成功采用。不过由于需要额外的升压电路，影响了方案的通用性，此方案并未在业界推广开来。

功率合成

其实电流、电压的合并都是在做功率合成，不过，还有一些功率合成方式不能严格的说是在做电流合成还

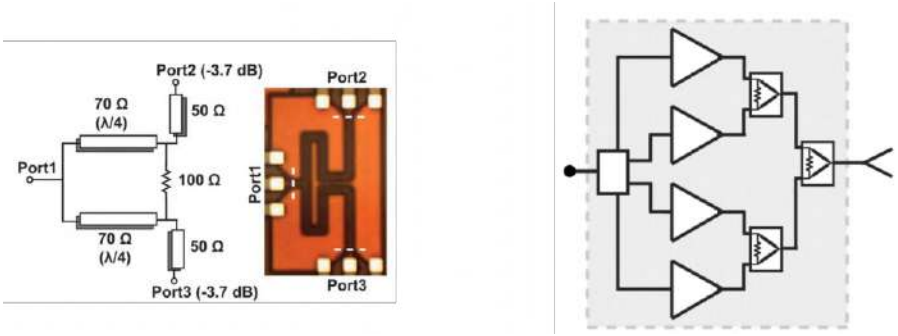


图 2.69: 典型的威尔金森功分器设计，以及其设计的功率合成 PA

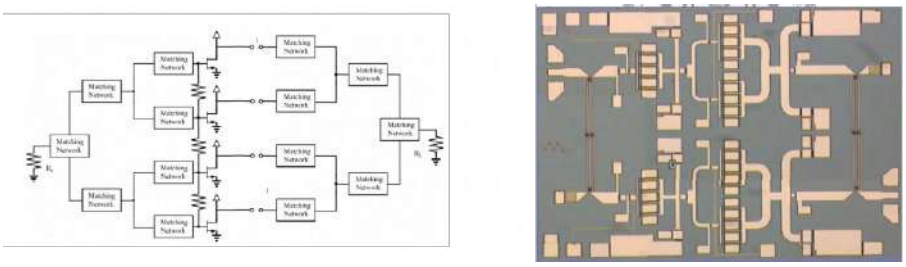


图 2.70: 采用二合一直接合并的方式实现的功率放大器

是在做电压合成，就将其归类为“功率合成”。

典型的功率合成方式是用功合器进行功率合并，威尔金森（Wilkinson）功合器（功分器）就是一种简单的功合器。威尔金森功率合器与其设计的功率合成 PA 如图2.69所示 [2]：

由于威尔金森功合器有三端口可以同时匹配的特性，所以采用其设计的功率合成 PA 设计简单，只需要单独设计好各路 PA，再进行单独进行功率合成即可。不过，由于威尔金森功合器需要两段 λ/4 传输线，并且两支路之间需要并接 100 Ohm 电阻，在 Layout 实现时并不经济，在对面积有限制的设计中无法有效采用。

为了改善威尔金森功合器的面积制约，在一些设计中采用了直接二合一合并匹配（Binary Combine）的方式进行设计，虽然采用这种方式不能达到三端口的完全匹配，但由于所占面积小，实现方便，在 MMIC 设计中得到了广泛采用。图2.70为采用直接二合一合并方式实现的 PA 设计 [3]。

除了在功率合成器上的研究外，一些研究还采用变压器（Transformer）或者直接空间功率合并的方式进行功率合成 [2]。如图2.71所示。

2.5.3 特殊功率合成

在“简单功率合成”中，功率合成的思路是将功率进行简单合并，完成“1+1=2”的功率输出。除了功率的简单合成外，还可以在功率合成中加入特殊设计，完成功率合成的同时，也实现更复杂的特性。

一些常见的特殊功率合成方法有 Push-pull（推挽）、Balance（平衡）、Doherty（多尔蒂）等。

2.5.3.1 Push-pull PA

Push-pull PA 一般在中文中被译为推挽 PA。Push-pull PA 的设计是将两个分别正、反导通的放大器合并，完成整个周期波形合成输出。这样，每个单独的 PA 就可以设计为高效率的 Class B 工作模式，PA 整体有高的效率。如图2.72所示。

上图所示的 Push-Pull PA 由 NPN 和 PNP 型两个晶体管构成，分别负责正半周期及负半周期的信号导通。在实际设计中，由于 PNP 型双极型晶体管一般不易做成高速，而且在集成电路实现中，一般的 Epi（外延层）只含有一种类型的晶体管，所以在射频中常采用双 NPN 型晶体管设计电路。

这个时候输入和输出就需要用到不平衡到平衡的转换电路，即巴伦（Balun，Balance to Unbalance），来将两路信号进行反向。采用巴伦和双 NPN 晶体管设计的 Push-pull PA 如图2.73所示。不同于采用 NPN+PNP 设计的

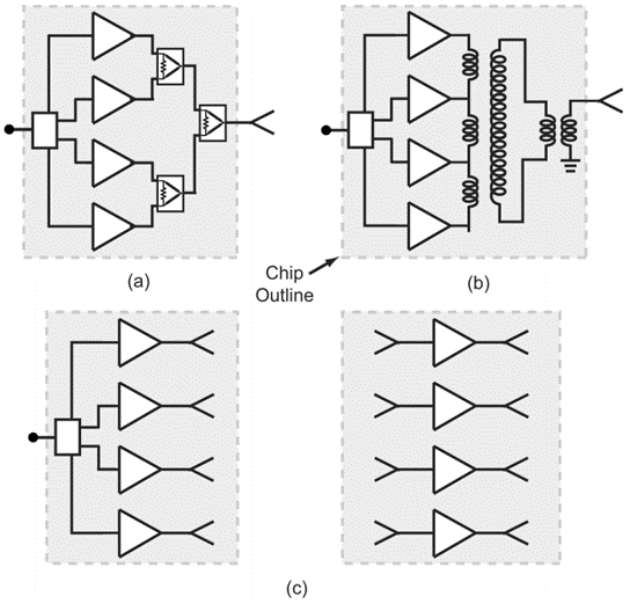


图 2.71: 采用威尔金森、变压器，以及空间合并的功率合成示意图

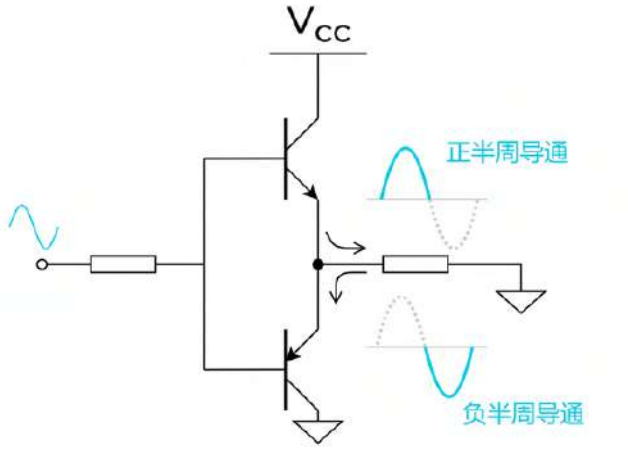


图 2.72: Push-pull PA 原理图示

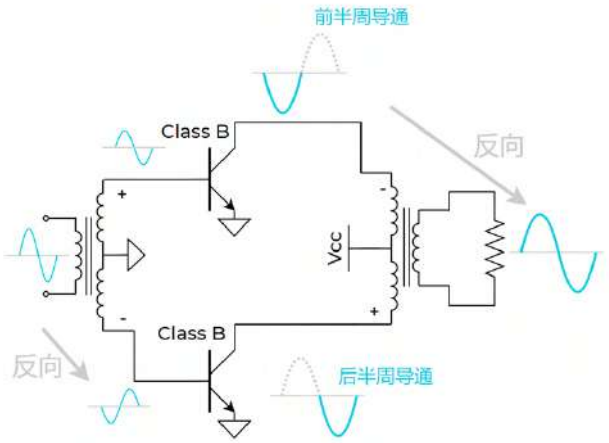


图 2.73: 采用双 NPN 型晶体管及巴伦设计的 Push-Pull PA

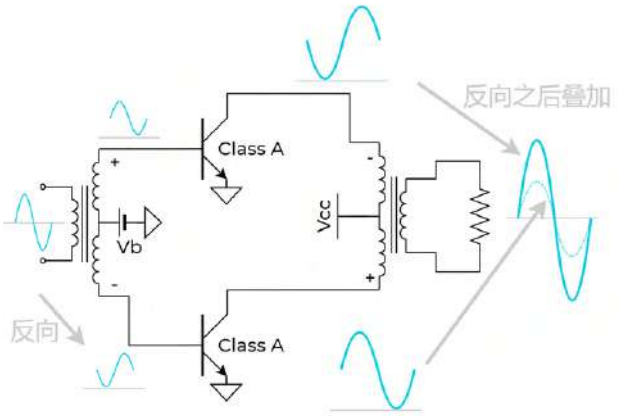


图 2.74: 采用 Class A PA 设计的功率合成 Push-Pull PA

推挽 PA 中有物理接地、两路功放均与地为参考流动，采用巴伦和双 NPN 晶体管设计的 Push-pull PA 电流在两路之间差分流动，射频以二者中间点为虚拟参考地。

需要说明的是，Push-Pull PA 不止可以用两个 Class B 的 PA 进行 PA 效率的提升，还可以用两个 Class A 的 PA 合并，进行功率的合成提高。采用 Class A 进行设计时，单个 PA 的效率并没有提升，但输出功率合成增加。采用两个 Class A PA 进行功率合成的示意图如图2.74所示。

巴伦是 Push-pull PA 的重要器件，巴伦是完成平衡信号（差分信号）与非平衡信号（单端信号）相互转换的电路。在平衡信号侧，信号以差分形式传输，相位相差 180 °；在非平衡侧，信号以地为参考，单端传输。双线变压器绕线法是常见的一种巴伦实现方法，采用双绕线耦合的形式，可实现信号不平衡到平衡的相互转换 [4]。并且，还可以改变线圈的比例，实现不同阻抗的变换。如图2.75所示。

在芯片设计中，通常利用多层金属边缘耦合（Edge Couple）或者宽边耦合（Broad Side）的方式实现金属线圈间的耦合。图2.76为集成电路中的变压器巴伦的实现 [5]。

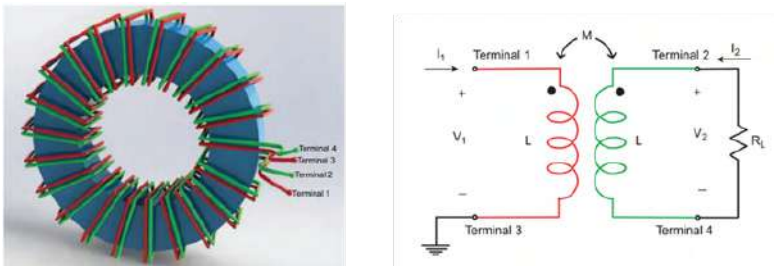


图 2.75: 双线变压器绕线法实现的巴伦

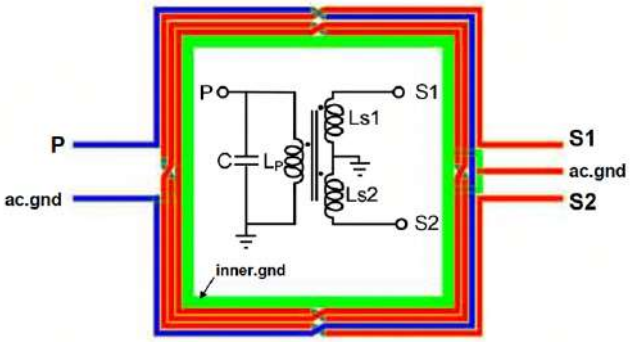


图 2.76: 集成电路中变压器巴伦的实现及其等效电路

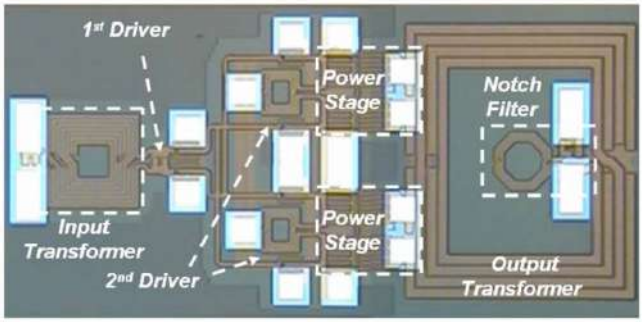


图 2.77: 典型的 Push-pull PA 芯片设计

为了产生对称的差分信号，巴伦在设计中一般注意线圈绕线的对称；另外 Push-pull PA 两个放大通路也需要对称设计，这也就使得 Push-pull PA 较易识别：

- 有对称的两个 PA 放大通路
- 两个 PA 放大通路的前后有对称的绕线巴伦

图2.77为典型的 Push-pull PA 芯片设计版图 [6]。

2.5.3.2 Balance PA

Balance PA 一般翻译为平衡放大器，是另外一种特殊的功率合成方式。Balance PA 与 Push-pull PA 相同，也是采用两路 PA 进行功率合并。不过与 Push-pull PA 的 180° 功率分配与合成不同，Balance PA 采用的是 90° 的功率分配与合成。图2.78为 Balance PA 的设计框图。

Balance PA 最大的特点是，只要两路 PA 是完全对称的，则两路 PA 的反射信号将在输入与输出端口完全抵

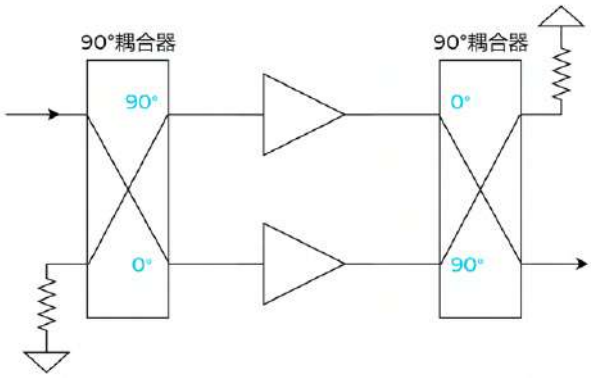


图 2.78: Balance PA 设计框图

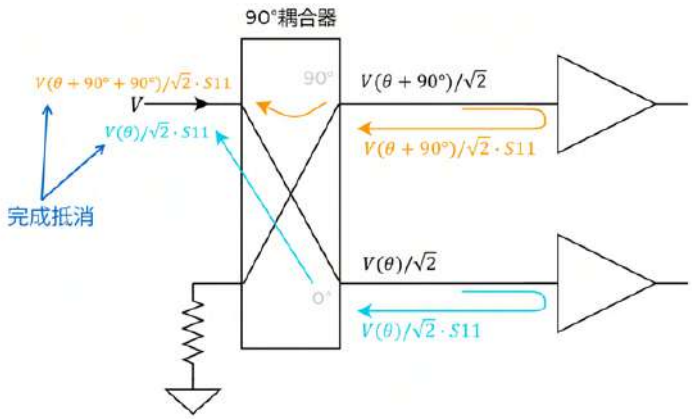


图 2.79: 平衡放大器对输入驻波的改善作用

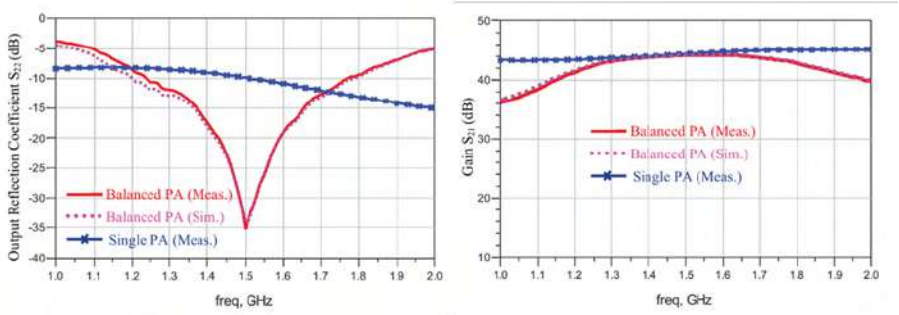


图 2.80: Balance PA 与单端 PA 的特性比较

消，实现输入和输出在较宽范围内有较好的 VSWR，适合于宽带微波毫米波放大器设计，以及对 S11/S22 有特殊需求的场景。Balance PA 中反射信号的抵消原理可用图2.79表示。

因为有以上特性，Balance PA 有很好的 S11/S22 特性，文章 [8] 显示，Balance PA 架构可以将以 1.5GHz 为中心频率 PA 的 S22 实现由-10dB 至-35dB 的优化，在 1.4 ~1.6GHz 范围内均实现-15dB 以下的 S22。

但由于 Balance PA 对 S11/S22 的抵消以及功率的合成依赖于耦合器 90° 的精准相移，而耦合器的相移精度与频率相关，所以 Balance PA 只能在一定频率范围内控制有效。根据文章 [8] 显示，在 1.2 ~1.7GHz 之外，Balance PA 的 S11/S22 与 S21 与单端相比反而出现了恶化，在 1.0GHz 及 2.0GHz 处，S22 恶化 5 ~10dB，S21 恶化 5 ~7dB。如图2.80所示。

由于在抵消点附近 Balance PA 表现出良好的 S11/S22，同时有较好的负载不敏感特性，Balance PA 在一些业界公司也被称为 Load Insensitive PA（负载不敏感 PA），简称 LIPA。

Balance PA 设计的关键器件是 90° 功率分配及合成器，其实现方式有多种方式，可以采用移相器的方式进行设计，也可以采用定向耦合器，或者 90° 正交混合网络的方式进行设计。图2.81为采用移相的方式实现 90° 功率分配与合成的 Balance PA 设计 [9]。

在微波/毫米波频段，若采用集总的电感、电容移相，则器件值较小，较难实现。一般在微波/毫米波频段会采用 90° 耦合器的方式来实现 90° 的功率分配与合成。图2.82为采用定向耦合器方式实现 90° 功率分配与合成的 Balance PA 设计 [10]。

Balance PA 设计和特性的特殊性也可以用来识别 Balance PA：

- 有对称的两个 PA 设计通路
- 有不对称的功率分配/合成网络
- 测试 S11/S22，在抵消点附近有较好的 S11/S22

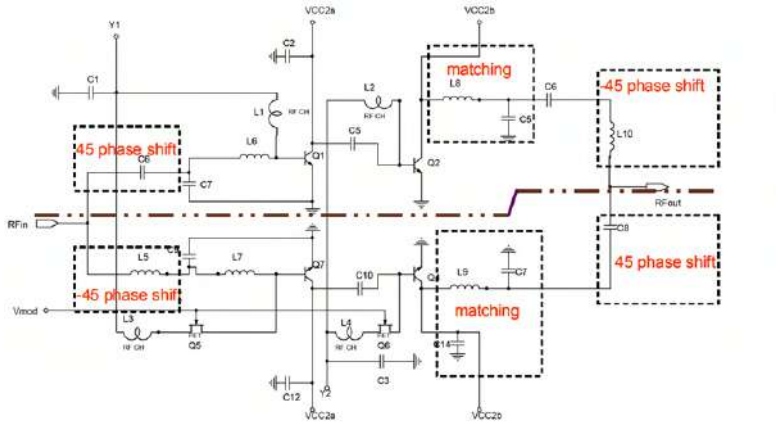


图 2.81: 采用移相的方式实现 90° 功率分配与合成的 Balance PA 设计

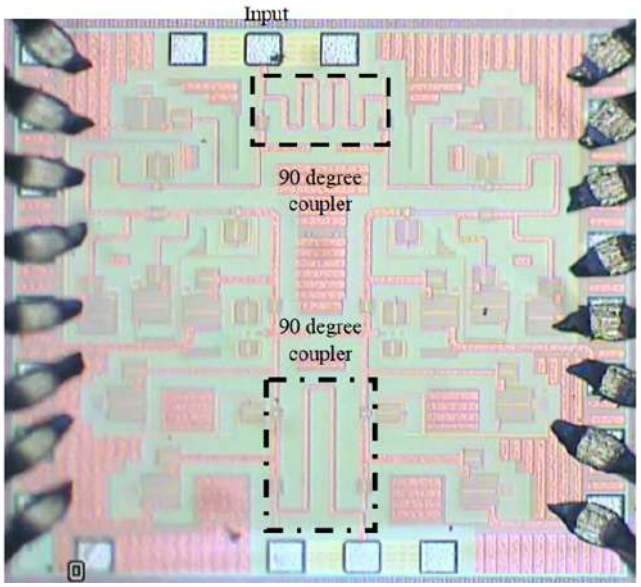


图 2.82: 采用定向耦合器方式实现 90° 功率分配与合成的 Balance PA 设计

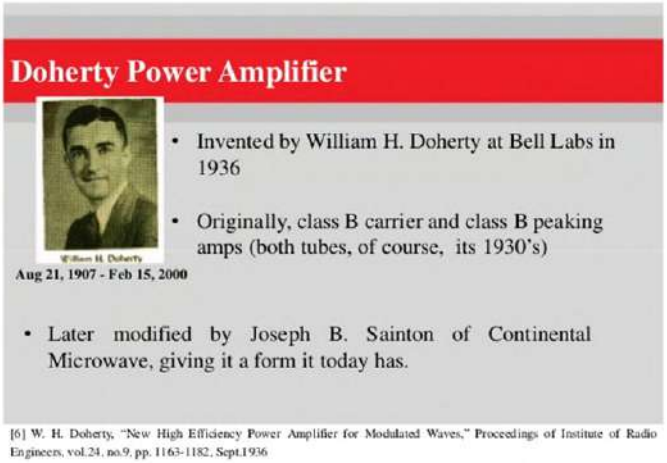


图 2.83: William H. Doherty 个人简介

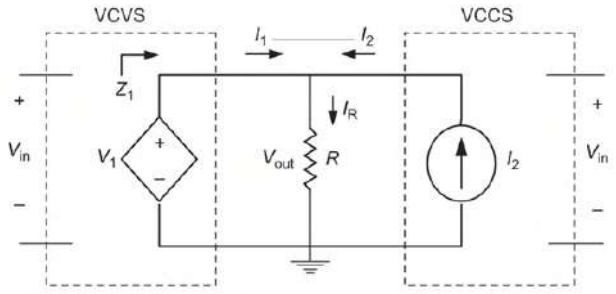


图 2.84: 电压源及电流源激励下的负载调制效应

2.5.3.3 Doherty PA

Doherty PA 近年来可谓是在手机射频 PA 使用中备受关注的架构，话题性非常强。Doherty PA 引人注目的一个原因是它的高效特性，可以很好的缓解 5G 因为功率提升带来的功耗提升问题。

Doherty PA 对于效率的优化是通过“动态负载调制效应”，是靠“两路工作在不同状态的 PA 相互配合，使得 PA 的负载发生变化，从而优化 PA 的回退特性”。虽然教科书上都是如此解释，但一个 PA 的负载竟然可以靠另外一个 PA 的工作状态来改变，听上去总是有些黑科技的意味，更加深了 Doherty PA 的神秘感。

Doherty PA 近年来被手机 PA 设计所关注，但却不是最近才被研制出来的新架构。Doherty PA 是由享有盛名的实验室：贝尔实验室的工程师 William H. Doherty 在 1936 年所发明的，距今已经有近 90 年历史。William H. Doherty 最初研究的目标是开发 KW 量级的高效跨大洋传输功率放大器，在 Doherty PA 被发明后，由于匹配的应用场景还未出现，在整个 20 世纪的大部分时候，仅在 AM 发射机中有些应用。

虽然 Doherty PA 发明之后在很长时间里并没有广泛应用，但这一切都是藏器待时。1990 年，随着全球移动通信的迅猛发展，对于高效、高功率的 PA 有了强烈需求。同时，硅基、III-V 族半导体工艺的发展，数字信号处理技术带来的线性化技术，为 Doherty PA 在移动基站的应用提供了坚实的应用基础。Doherty PA 技术在基站侧迅速发展，目前几乎统治了整个宏基站 PA 市场。

Doherty PA 的核心原理是“负载调制”效应，基原理如下图所示。若一个负载 R 由两个源进行激励，分别为电压源及电流源，则电压源 V1 看到的阻抗可以由欧姆定律表示为：

$$Z_1 = \frac{V_1}{I_1} = \frac{V_1}{I_R - I_2}$$

若电流源的输出电流 I2 发生变化，则电压源 V1 看到的阻抗也将发生变化。即可以用电流源 I2 的大小，控制 V1 看到的阻抗，这就是“负载调制”的基本原理。

在实际 PA 设计中，一般是将 PA 等效为电流源。于是在实际设计中需要加入阻抗反转网络，将电流源转换

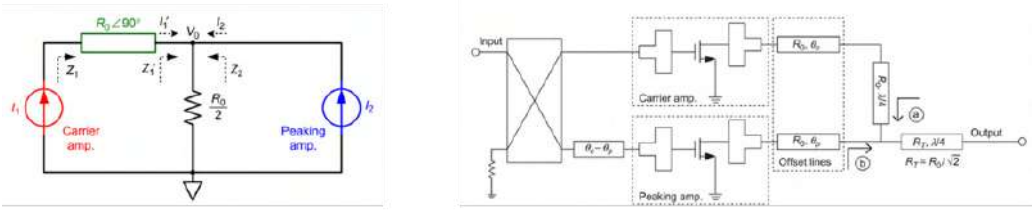


图 2.85: ：Doherty PA 的原理和具体实现

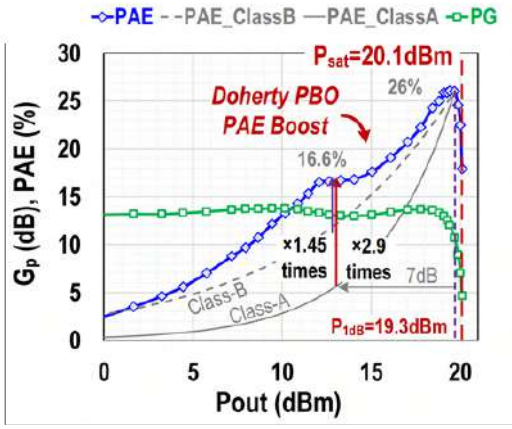


图 2.86: Doherty 在功率回退时的效率改善特性

为电压源。另外，为了满足阻抗匹配的需求，和相位对齐的需求，还会加入补偿线等，完成整个 Doherty 功放的设计。如图2.85所示。

Doherty PA 由两个 PA 通路构成，分别是 Carrier 通路和 Peak 通路。在小功率工作时，只有 Carrier 功放开启，其负载线维持在较高位置，保持高的效率；在大功率时，Peak 功放开启，Carrier 功放负载线被调制至较低位置，产生较高功率输出。以此完成功率回退时效率的提升。（有关负载线与输出功率的关系，详见文章《5G PA 的 Load-line 与 Load-pull》。

实际在分析 Doherty PA 的工作时，并不像以上说明这么简单，在慧智微 PA 问答群中，Doherty PA 设计专家表示：Doherty PA **的推导需要在功率、电压、阻抗三个维度去考虑。推导 Doherty 理论，首先要有一个概念，就是要站在整体上把 Doherty 分解，从功率角度就是两个节点，一是回退功率，二是饱和功率，站在这两个点上去分析两个管子的阻抗变化，功率变化，电压变化。**

图2.86展示为一个典型的 Doherty PA 效率曲线图，此设计为一款饱和功率为 20dBm 的 60GHz Doherty PA。可以看到，在回退 7dB 时，Doherty 功放的效率相比于 Class-B 功放有 1.45 倍的提升，相比于 Class-A 有 2.9 倍的提升 [13]。

在手机应用中，Doherty PA 并不常见。最主要的原因是手机使用环境复杂，需要支持的频段和模式众多，而 Doherty PA 负载敏感、较窄带，并且需要强大算法加持的特性使得传统意义 Doherty PA 在手机应用中并不兼容。手机与基站侧应用环境的比较如图2.87所示。

近年来随着 5G 手机中 PA 耗电的持续增加，Doherty PA 被重新考虑是否可应用于手机应用。因为手机应用复杂，需要考虑在高低温、不同天线驻波比下均满足系统指标，Doherty PA 在满足这些指标设计时，需要牺牲回退效率优化的特性，用以折中兼容以上手机应用中的必须特性。

图2.88为文章 [14] 中展示的 Doherty PA 设计实例。

在实际分析中，可以根据 Doherty PA 在的特性来识别是否是 Doherty PA：

- 至少有两路功放构成
- 两路功放呈非对称状态，或者是设计的不对称，或者是偏置状态的不对称
- 功率合成部分为非对称结构

	基站	手机
负载	固定	复杂变化
频段带宽	窄带 (针对运营商)	宽带， 全频段均需支持
频段和模式数目	较单一 (针对运营商)	全球多模式、 多频段均需支持
算法支持	有强大算法支持， 可做复杂校正和线性化	无法提供强大算法支持

图 2.87: 基站与手机应用环境差异

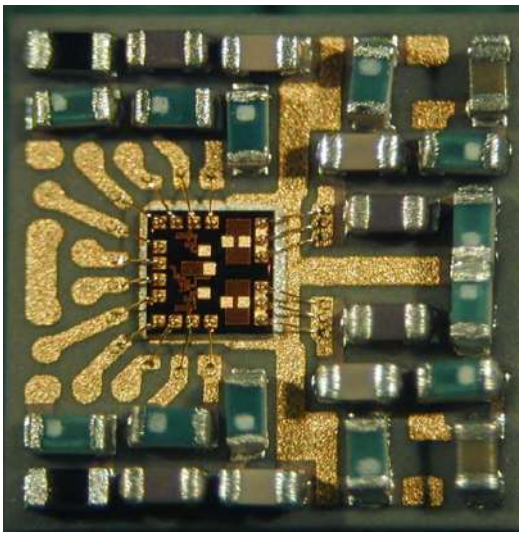


图 2.88: Doherty PA 设计实例

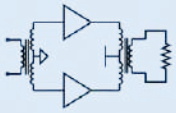
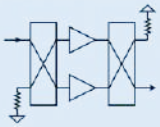
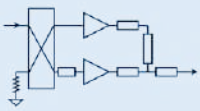
	Push-pull	Balance	Doherty
架构			
功能特点	1. 两个A类功放可做功率合成 2. 两个B类功放可用于效率提升	1. 可应用于对S11/S22有高要求的场景 2. 可做功率合成	1. 可改善回退效率 2. 窄带，负载敏感
功放通路	对称	对称	非对称
功率合成	180°	90°	90°及补偿线

图 2.89: 特殊功率合成 PA 架构总结整理对比

- 功率合成部分为非对称结构

以上不同的功率合成架构有不同的应用，将特殊功率合成将特殊功率合成 PA 架构总结整理对比如图2.89所示。

2.5.3.4 多种架构混合设计

不同 PA 架构之间也不是非此即彼，也是可以相互结合的。比如在 Doherty PA 或者 Push-pull PA 的每个 PA 单元中，都可以选择电流合并或者电压合并的结构进行设计。

甚至不同的特殊功率合成之间也可以相互组合，例如，Doherty PA 的 Carrier 和 Peak PA 就可以选择 Push-pull PA。甚至还可以将两个 Doherty PA 组合设计成 Balance PA，减少 Doherty PA 的负载敏感特性。

图2.90为文献 [15] 中所设计的 PA，其中利用到了威尔金森功分器、90° 正交耦合器、差分放大器以及 Doherty PA 架构。

2.5.4 总结

本文对工程中常见的 PA 架构做了简单整理，介绍了不同架构 PA 的设计思路，架构特点，以及主要的实现方式。

另外需要说明的是，PA 架构并无优劣之分，只有合适与否。只有了解当前的需求和限制，了解不同 PA 的特性，才能选择恰到好处的 PA 架构。

本文写作过程中，得到众多业界专家的指导，在此表示感谢！

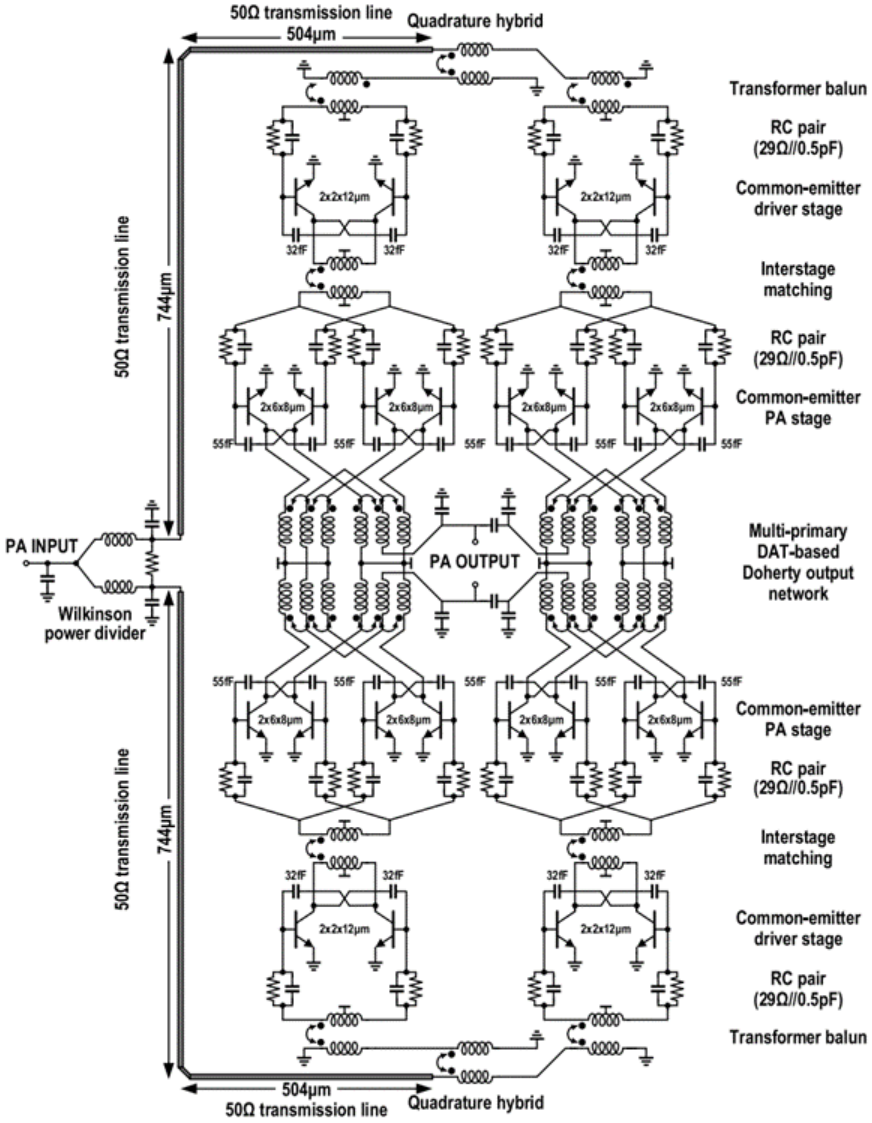


图 **2.90:** 多种架构组合的 PA 设计

参考文献

1. Lei M F , Tsai Z M , Lin K Y , et al. Design and Analysis of Stacked Power Amplifier in Series-Input and Series-Output Configuration[J]. IEEE Transactions on Microwave Theory & Techniques, 2007, 55(12):2802-2812.
2. Atesal Y A , Cetinoneri B , Chang M , et al. Millimeter-Wave Wafer-Scale Silicon BiCMOS Power Amplifiers Using Free-Space Power Combining[J]. IEEE Transactions on Microwave Theory and Techniques, 2011, 59(4):954-965.
3. Ping-Sung Chi, Research of HEMT and CMOS Microwave Power Amplifier, 2009, NTU.
4. Mclean J . Analysis of the Distributed Bifilar Isolation Transformer and Current Balun[C]// 2020 IEEE International Symposium on Electromagnetic Compatibility & Signal/Power Integrity (EMCSI). IEEE, 2020.
5. Lin Y C , Chen K S , Horng T S , et al. A 2.5-2.7 GHz pseudo-cascode CMOS receiver front-end with integrated unequal-turn-ratio transformer balun. 2010.
6. An K H , Kim Y , Lee O , et al. A Monolithic Voltage-Boosting Parallel-Primary Transformer Structures for Fully Integrated CMOS Power Amplifier Design[C]// Radio Frequency Integrated Circuits. IEEE, 2007.
7. Cripps S C . The evolution of the push-pull RFPA[C]// Microwave Symposium. IEEE, 2015.
8. Olukoya O , Kodogiannis V , Budimir D . Highly Efficient Balanced Power Amplifiers for 4G Applications[C]// 2018 41st International Conference on Telecommunications and Signal Processing (TSP). 2018.
9. Zhang G , Chang S , Chen S , et al. Dual mode efficiency enhanced linear power amplifiers using a new balanced structure[C]// Radio Frequency Integrated Circuits Symposium. IEEE, 2009.
10. Lee J , Chen C C , Tsai J H , et al. A 68–83 GHz power amplifier in 90 nm CMOS[C]// International Microwave Symposium Digest. IEEE, 2009.
11. Staudinger J . Doherty history, principles and key developments[C]// 2016 IEEE/MTT-S International Microwave Symposium - MTT 2016. IEEE, 2016.
12. Kim B . Doherty power amplifiers: From fundamentals to advanced design methods. 2018.
13. Nguyen H T , Wang H . A Coupler-Based Differential Doherty Power Amplifier with Built-In Baluns for High Mm-Wave Linear-Yet-Efficient Gbit/s Amplifications[C]// 2019 IEEE Radio Frequency Integrated Circuits Symposium (RFIC). IEEE, 2019.
14. Kato T , Yamaguchi K , Kuriyama Y . A 4-mm-square 1.9 GHz Doherty power amplifier module for mobile terminals[C]// Microwave Conference. IEEE, 2005.
15. Wang F , Wang H . A High-Power Broadband Multi-Primary DAT-Based Doherty Power Amplifier for mm-Wave 5G Applications[J]. IEEE Journal of Solid-State Circuits, 2021, 56(6):1668-1681.

在线阅读

本文摘自慧智微微信公众号，扫描以下二维码，在线阅读

2.6 高效率 PA：系统级设计

作者：彭洋洋 博士

“效率”在射频功率放大器（PA）设计中占据举足轻重的地位。

在之前的两篇文章中（《高效率 PA 设计：从 Class A 到 Class J》与《5G 射频 PA 的架构》），我们已经深入探讨了高效率 PA 设计的两大核心：PA 的“Class”设计以及功率合成架构。然而，在实际的射频前端系统中，PA 并非孤立存在，而是与整个系统紧密相连。系统层面的设计方案对 PA 功耗有着深远影响。

当单体 PA 的效率提升至一定极限后，PA 厂商、平台厂商以及终端系统厂商开始将研究焦点转向包含 PA 在内的系统级设计。试图通过更高层次的系统解决方案，进一步突破 PA 的性能瓶颈。在这一领域，既有大家耳熟能详的包络跟踪（ET）和数字预失真（DPD）技术，也有相对陌生的负载调制（Load Modulation）和波峰因子减少（CFR）等技术。

这些技术究竟是如何神奇地提升 PA 效率的？未来又有哪些新技术可能应用于 PA 效率的提升？本文将试着从系统级的角度，理解高效率 PA 系统设计的奥秘。

2.6.1 效率与线性度：PA 设计的双重挑战

在 PA 的单体设计中，效率和设计所选的“Class”、拓扑架构乃至匹配网络的损耗等诸多因素紧密相连。然而，从系统层面来看，**各种 PA 的效率曲线都可大致描绘为一根随功率提升而上升的曲线**。简而言之，随着功率的增加，射频摆幅扩大，越来越多的能量被转化为射频信号，因此在高功率状态下，PA 将直流电转换为射频信号的效率达到最高。

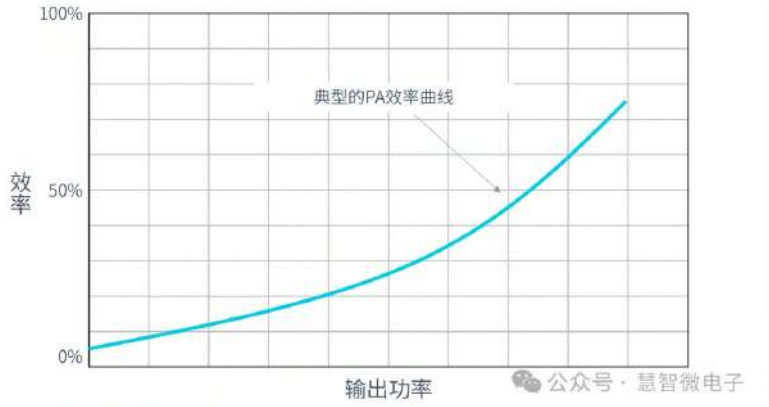


图 2.92: 典型的 PA 效率曲线

尽管不同 PA 的峰值效率点和功率回退时的斜率存在差异，但效率随功率变化的大致趋势是固定的。基于这一效率曲线，我们可以得出一个 PA 高效率应用的基本原则：**尽量让 PA 工作在接近高功率的状态**。

然而，这种高效率状态并非没有代价。当 PA 越接近高功率饱和区，其线性度就会逐渐恶化。典型的增益随输出功率变化曲线显示，在饱和区，PA 增益可能会出现几 dB 甚至十几 dB 的压缩，对于线性信号而言，这几乎使得 PA 处于不可用状态。

这一困境该如何应对？这就需要引入“功率回退”（Power-back off）的理念。所谓功率回退，即为了保持输出信号的线性度，在使用时将功率从饱和点适度降低。这意味着为了保证线性度，我们必须牺牲一部分功率输出，也就牺牲了一部分的效率。

在 PA 的应用中，效率和线性度始终是一对需要权衡的因素。这种权衡体现在多个方面：

1. 在调整输出功率时，效率随功率增加而提升，但线性度却随之下降；
2. 在进行负载匹配时，PA 的最佳效率点和最佳线性度点往往并不重合；
3. 在选择供电电压时，更高的电压虽然有利于改善线性度，但也会降低效率。

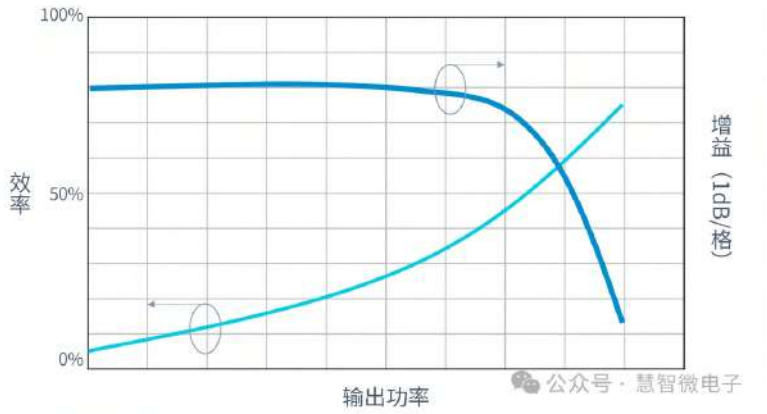


图 2.93: 典型的 PA 效率与增益曲线

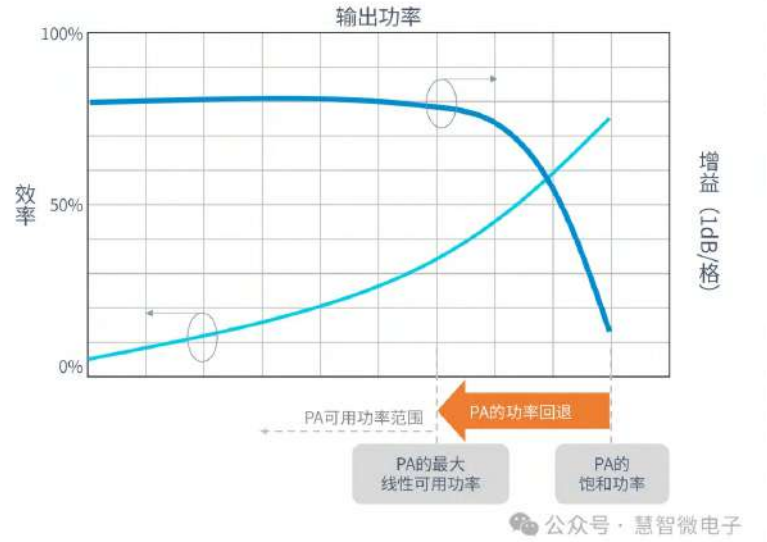


图 2.94: PA 的功率回退使用

因此，在 PA 设计中，如何平衡效率和线性度成为了一个永恒的挑战。

2.6.2 改善输入信号：使 PA 在更高效率区域运行

在 PA 设计中，会致力于扩展 PA 的工作区域，以获得更出色的效率表现。例如，设计时会尽量推迟 P1dB 点的出现，使 PA 增益的压缩尽量延迟，从而让 PA 的工作点更接近饱和点，以减少效率损失。

同样，在不需要极高功率的设计中，也可以通过降低饱和功率的设计，也能使 PA 工作点更加接近饱和功率处的高效区域。

在系统设计层面，也可以采用类似的理念来提升效率。这里，CFR 和 DPD 两项关键技术发挥着重要作用。

2.6.2.1 CFR：降低信号的峰均比

CFR（Crest Factor Reduction），即波峰因数降低技术，实质上是一种信号“削峰”技术。在现代通信系统中，传输的信号通常是幅度和相位均带有信息的编码信号，这导致信号存在较大的峰均比（PAPR）。峰均比越大，意味着信号中的峰值与均值之间的偏差越大。

为了保证信号不失真，必须确保信号的最高功率点不超过 PA 的饱和区。因此，均值功率的工作点需要比饱和功率低至少一个峰均比。所以，信号的峰均比越大，需要回退的功率就越多，从而降低 PA 的效率。

CFR 技术通过算法处理原始信号，尽量减少高峰均值信号的出现，或通过限幅、削峰等方法对输入信号进行

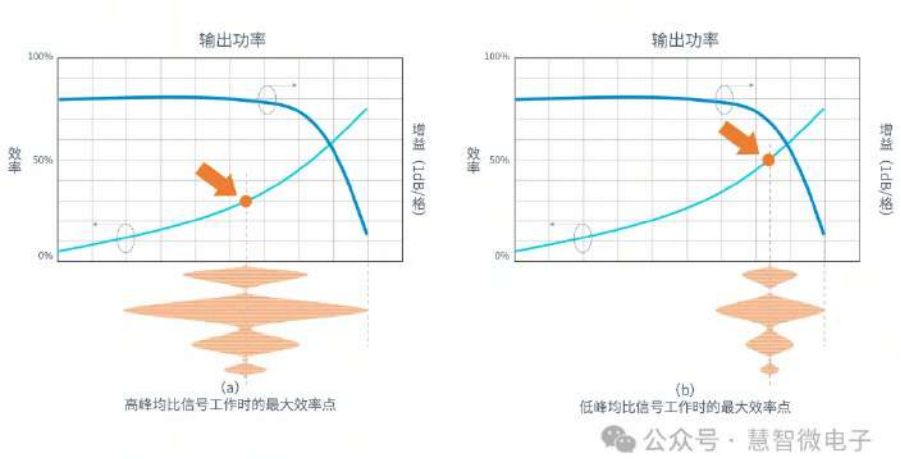


图 2.95: 高低两种峰均比信号对 PA 使用影响对比

限制。这样，进入 PA 的信号就不会有过高的峰值，使 PA 能够在更高的功率范围内工作，从而提升效率。图2.96为经过 CFR 修改后的信号与原始信号对比图。

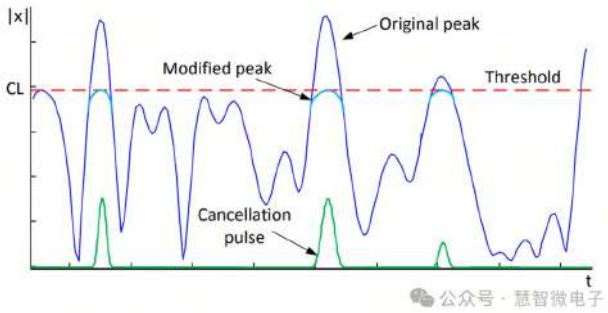


图 2.96: CFR 技术对信号峰值的降低

由于 OFDM 信号的特性，4G/5G 信号在时域中具有显著的峰均比（原因见文章《5G PA“记忆效应”的现象、形成与消除》），这使得 CFR 技术显得尤为重要。系统 CFR 能力的强弱直接影响到整个系统的性能表现。因此，在基站和手机终端等平台中，CFR 技术已经得到了广泛的推广和应用。

2.6.2.2 DPD：校准 PA 失真

尽管 CFR 能够整形信号并降低峰均比，但信号的峰均比始终存在。当 PA 的功率向上提升时，必然会遇到线性度恶化的问题。此时，DPD 技术便发挥作用。

DPD（Digital Pre-distortion），即数字预失真技术，其核心理念是让输入信号提前产生一个与 PA 相反的失真，以抵消 PA 引入的失真，工作原理如图2.97所示。

在 DPD 的帮助下，PA 的线性度得到提升。系统也就能容忍更多的 PA 线性度恶化，从而使其能够在更高的功率范围内工作，进而提高效率。有 DPD 和没有 DPD 时，PA 可工作的功率区域如图2.98所示。

DPD 操作完全在数字域进行，无需对 PA 进行特殊调整即可看到效果。然而，DPD 并非万能，它也有自身的局限性。DPD 的效果强烈依赖于算法，而移动终端由于功耗和算力的限制，无法支持过于复杂的 DPD 算法。此外，PA 的特性会随温度、负载、偏压和批次等因素的变化而变化，为了获得良好的校正效果，DPD 需要不断调整以适应这些变化。同时，DPD 需要与 PA 进行联合调试和特定开发，以充分发挥二者的潜力。

尽管有所限制，DPD 仍在蜂窝和 Wi-Fi 等应用中得到了广泛应用。最近 Wi-Fi 应用中较为热门的“非线性 PA”，就是利用 DPD 技术和非线性 PA 技术结合实现的高效率 PA 技术。

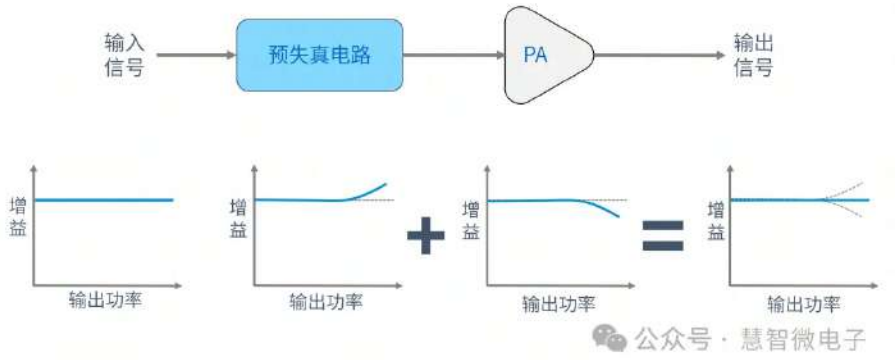


图 2.97: DPD 的原理

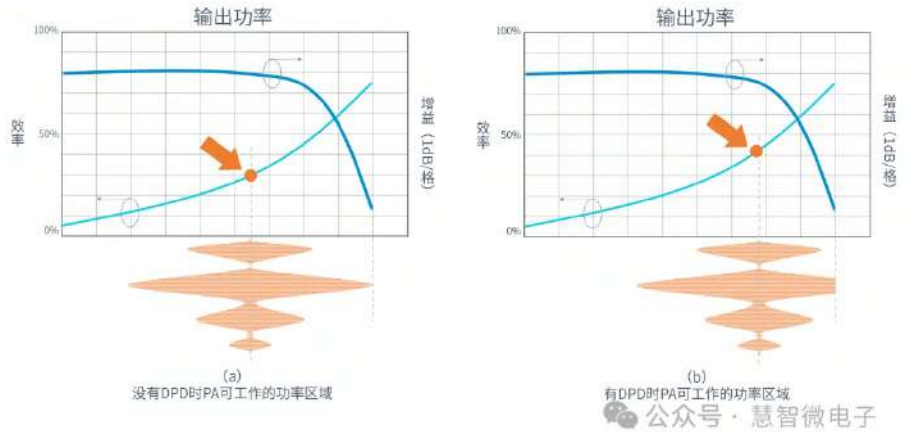


图 2.98: DPD 下，PA 工作的功率区域可以提升

无论是 CFR 还是 DPD，它们都旨在通过算法降低系统对 PA 线性度的要求，从而提升 PA 的工作功率和效率。这些技术对平台的算法能力提出了更高的要求。

除了 CFR 和 DPD 之外，系统中还采用了另一种方法来提升 PA 在回退功率点的效率，即“动态调制”。这种方法通过动态调整系统参数，来优化 PA 的工作效率。

2.6.3 参数的动态调制

动态调制是指根据信号的输出功率，实时调整系统中的某些关键参数（如供电电压、供电电流、负载阻抗等），以达到优化功耗的目的。这种调制方式要求参数的变化必须非常迅速，以适应信号输出的瞬时变化需求，达到接近实时的状态。

以 100MHz 带宽的 5G NR 信号为例，调制信号的变化需要达到纳秒级别，才能与包络信号的变化保持同步。这样的速度要求远远超出了芯片内部数字接口电路的能力范围。举例来说，射频前端标准接口协议 MIPI RFFE 的最大通信时钟频率为 52MHz，发送一个完整的 MIPI 指令需要 25 个时钟周期，即完成一个控制转变需要 0.48us。这个速度比动态调制所需的切换速度慢了上百倍。

因此，为了实现参数的动态调制，通常会利用模拟控制接口，或者直接利用 PA 的内部特性来完成。

根据调节参数的不同，动态调制主要可分为以下三种类型：

1. Vcc 动态调制；
2. Bias 动态调制；
3. 负载动态调制。

这些调制方式能够更有效地管理功耗，尤其是功率回退时的功耗，提升系统的整体效率。

2.6.3.1 Vcc 动态调制：供电电压的灵活调整

Vcc 动态调制的核心思想在于，使 PA 的供电电源能够随信号幅度的变化而灵活调整，而非保持固定值。这种动态调整有助于显著提高 PA 效率，减少不必要的能量损耗。

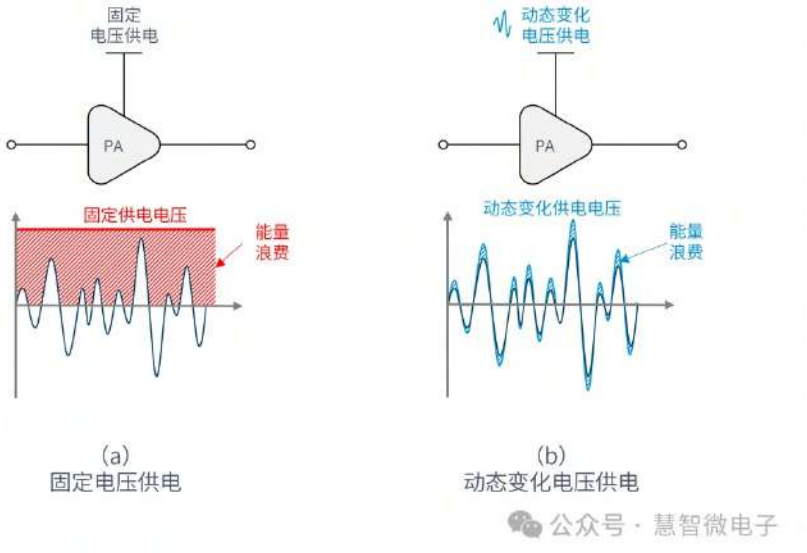


图 2.99: 固定电压供电及动态变化电压供电

上图是 Vcc 动态调制的基本原理。对于高峰均比信号，其幅度随时间快速变化。若采用固定电压供电，将会导致显著的能量浪费。而通过使电源电压跟随信号幅度变化，可以实时优化能耗。

实现 Vcc 动态调制的两种主要技术是 EER（Envelope Elimination and Restoration，包络消除与恢复）和 ET（Envelope Tracking，包络跟踪）。

EER 技术

EER 技术由 Kahn 教授于 1952 年提出，其基本原理是将输入信号分解为两部分处理 [2]。其基本架构如图2.100所示。

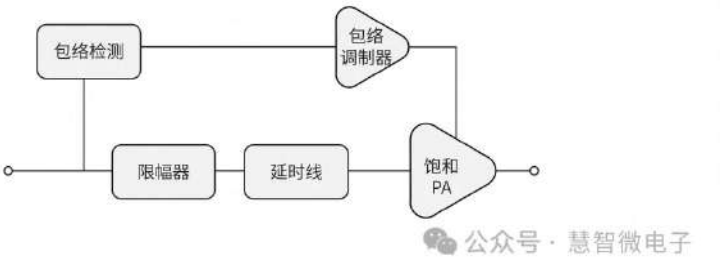


图 2.100: EER PA 系统架构

在 EER 架构中，一路信号忽略包络信息，直接传输至工作在开关状态的 PA，从而实现高效放大。这一过程中，包络信息被主动消除，故称“包络消除”通路。

另一路则负责“包络恢复”。首先，通过包络检测电路提取包络信号，然后输入包络调制器。输出的包络信号用作饱和 PA 的动态供电电源，既实现了电源的动态调制，又完成了包络的恢复。

EER 电路的优点在于巧妙地将包络信号和载波信号分离处理，使得两个通路各司其职，互不干扰，从而提高了系统效率。然而，它也存在一些缺点：

1. 由于信号的幅度和相位信息被拆分，因此需要精确对齐两个通路以确保信号的完整恢复。
2. EER 过程中存在幅度消除和恢复两个环节，若处理不当可能导致信号失真。

3. 本质上，EER 相当于包络信号与高频恒包络信号的混频，可能引发频谱扩展问题。

鉴于 EER 技术的复杂性和挑战性，另一种更为简洁的 PA 系统结构——ET 技术应运而生。

ET 技术

ET，即包络跟踪技术，其核心理念是“跟踪”信号所需的电源幅度，并为 PA 提供相应电源。这样既节省了功耗，又避免了 EER 技术中复杂的信号拆分与对齐过程。结合 ET 技术，可以显著提升系统整体效率。ET 系统架构如图2.101所示。

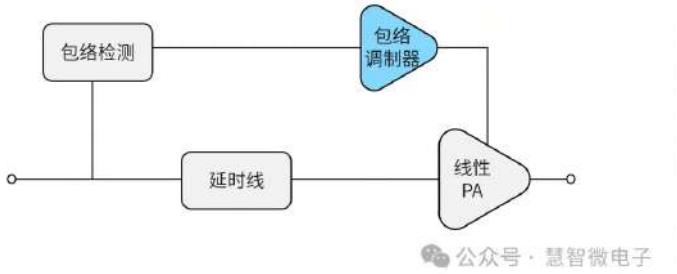


图 2.101: ET PA 系统架构

然而，ET 技术的性能高度依赖于 ET Modulator 的表现。为了输出大带宽（100MHz 及以上）、大电流（2A 及以上）的包络信号，并保持高转换效率，ET 调制器的设计面临巨大挑战。随着带宽和电流能力的增强，实现高效率变得愈发困难。

在 5G 应用中，即使信号带宽维持在 100MHz，ET 调制器的转换效率通常也仅约 80%，意味着存在 20% 左右的能量损失。因此，ET PA 系统的收益必须足以弥补这部分能量损失才能实现整体收益。随着未来信号带宽的不断增加，ET 调制器的设计难度将进一步上升。

2.6.3.2 Bias 的动态调制

除了对供电电压 Vcc 进行动态调制外，PA 的偏置电流，即 Bias，同样可以进行动态调制以优化性能。

Saleh 教授等人在 1983 年的一篇文章中，提出了一种创新的系统架构来提升 PA 的效率 [3]。该架构的核心在于，通过提取输入信号的包络来控制 PA 的 Bias，使 Bias 能够随输入信号包络的变化而动态调整。这种调整方式目标是实现 PA 效率的最优化。

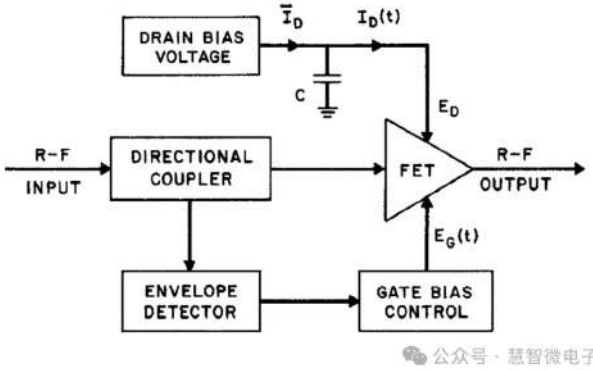


图 2.102: Saleh 教授等于 1983 年提出的 Bias 调制 PA 架构

在现代 PA 设计中，“动态偏置”特性被巧妙地融入 Bias 电路的设计中，以提升效率并优化线性度。

以 HBT PA 为例，利用其偏置电压 Vb 与 Ic 电流之间的指数关系，设计师们可确保偏置中心点 Vb 的尽量稳定。这时当大信号输入时，Ic 电流会自动增加，其工作原理如图2.103所示。这种设计使得在大功率时提升 Ic 电流，而在小功率时减少 Ic 电流，从而实现了高效的动态调制 [4]。

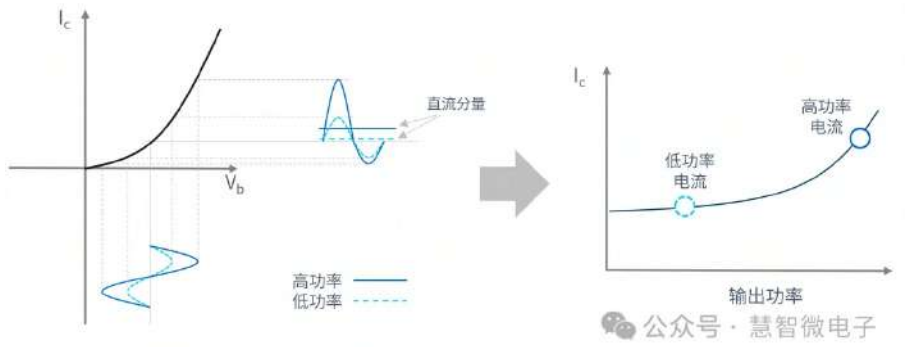


图 2.103: 电流的动态调制效应

2.6.3.3 负载的动态调制

除了电压和电流的动态调制外，还存在一种独特的调制技术，即负载动态调制（Load Modulation）。正如《5G PA 的 Load-line 与 Load-pull》一文中所探讨的，PA 的负载线（Load-line）在 PA 设计中占据至关重要的地位。负载线的高低直接影响着 PA 输出功率的大小。

为了达到理想的输出功率，PA 的负载线通常是根据最大输出功率来设计的。然而，这种设计在小功率输出时显得并不经济：此时 PA 的工作功率远低于其饱和功率。根据先前关于 PA 特性的讨论，我们知道当 PA 的工作功率远离饱和功率时，其效率会显著降低。

尽管可以通过降低 Vcc 电压来减少功率浪费，并降低输出的饱和功率以提升 PA 在小功率输出时的效率，但这种方法需要依赖 ET 技术。然而，ET 调制器的成本相对较高，并且其在大带宽下的效率是设计中的一大难题。

在这种情况下，对于小功率输出的场景，我们可以利用负载线的特性来降低输出饱和功率，从而提高效率。放大器的输出功率与供电电压成正比，与负载阻抗成反比。**在供电电压固定的情况下，可以通过增加放大器输出负载的方式来降低 PA 的饱和输出功率，从而提升小功率输出点的效率。**这正是负载动态调制的核心理念。

在实施方面，负载动态调制主要采用以下几种方法：

通过可调匹配网络实现；

1. 利用 Doherty 架构实现；
2. 借助 Out-phasing 架构实现；
3. 采用 Load Modulation Balance PA 架构。

可调匹配网络实现

可调匹配网络方法的核心在于，在 PA 的匹配网络中融入可由包络信号控制的可调无源器件。经过精心设计，这种匹配网络能够实现负载阻抗随功率的特定变化。图2.104是 Kenlen Chen 教授于 2012 年所发表论文中展示的架构图 [5]。

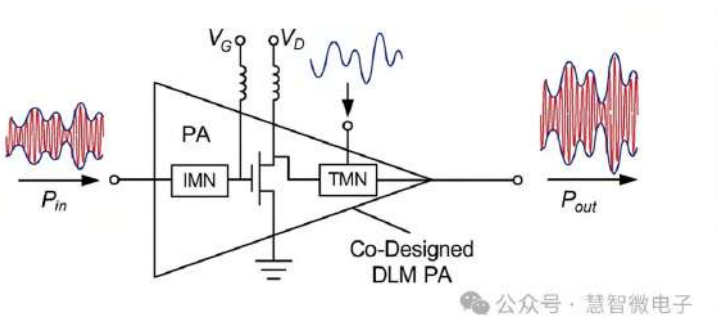


图 2.104: 可调匹配网络示意

此方法的直观性和易于理解性是其显著优点。不同功率点对应不同的阻抗，设计师只需通过匹配网络和相

应的控制信号设计到相应的目标阻抗，便可实现阻抗的特定转换。

然而，这种方法也存在局限性。其可控范围有限，且受限于匹配网络的固有特性，目标阻抗可能无法完全符合预期变化。此外，寻找满足低插损、高集成度、大范围调节和快速响应的可调器件也是不小的挑战。

因此，多种等效负载阻抗调谐方法应运而生。在这些方法中，负载阻抗并非真正被调节，而是通过电路等效方式使其发生变化，从而达到相应改变 PA 特性的目的。其中，Doherty、Out-phasing 和 Load Modulation Balance PA 三种实现方案尤为突出。

Doherty 架构实现

尽管在之前的《5G 射频 PA 架构》文章中已详细阐述 Doherty PA 的设计理念，但从负载动态调制的系统角度来看，仍有必要重新审视 Doherty PA。深入理解 Doherty PA 中的动态负载调制效应，将有助于更好地理解其他两种负载动态调制架构。

Doherty PA 由两个 PA 组成，但它们并非简单合成。通过对两个 PA 进行阻抗转换，其工作原理可简化为图2.105所示 [6]。

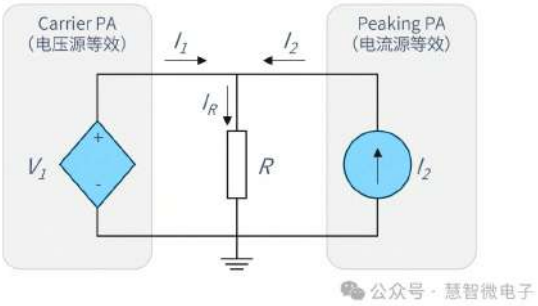


图 2.105: Doherty PA 的简化电路

在此图中，左右两个电源分别代表 Doherty PA 中的 Carrier 和 Peaking 两个 PA。左侧的 Carrier PA 通过阻抗变换被设计为等效电压源，而右侧的 Peaking PA 则设计为等效电流源。这两个 PA 共同作用于同一个负载阻抗 R 上。

由于负载 R 直接连接在电压源 V_1 两端，因此 R 两端的电压即为 V_1。根据欧姆定律，流过 R 的电流为 V_1/R。

值得注意的是，虽然电压源为电阻 R 提供电压，但它并不能控制其流向电阻 R 的电流。因为 R 上的电流由两部分组成：一部分是电流源 I_2 强制输出的电流，另一部分是向电压源请求的电流。电流源 I_2 会强制输出一个电流 I_2，因此电阻 R 向 V_1 请求的电流仅为 $V_1/R - I_2$。

对于电压源 V_1 而言，其输出电流的大小会受到另一个电流源 I_2 的影响。根据 V_1 输出的电压和电流，我们可以计算出电压源 V_1 看到的等效阻抗为 $R_{Load} = V_1/(V_1/R - I_2)$。

从上述公式中可以看出，V_1 看到的阻抗会随着 I_2 输出的大小而变化，**仿佛 V_1 的负载阻抗被外部控制了一样。这就是动态负载调制效应**。

在 Doherty PA 设计中，正是利用了这种效应。通过 Peaking PA 的开启与关闭，可以控制 Carrier PA 看到的阻抗变化。通过匹配转换网络的设计，当 Peaking PA 开启时，Carrier PA 看到的阻抗降低，从而提升 PA 的输出功率以实现高功率输出；同时，当关闭 Peaking PA 时，Carrier PA 看到的阻抗提升，此时 PA 的输出功率降低，换回了较高的效率。

通过这种方式，Doherty PA 通过负载线的动态调整实现了高功率和高效率的兼顾。图2.106及图2.107为 Doherty PA 的完整实现架构以及典型的效率特性曲线图。

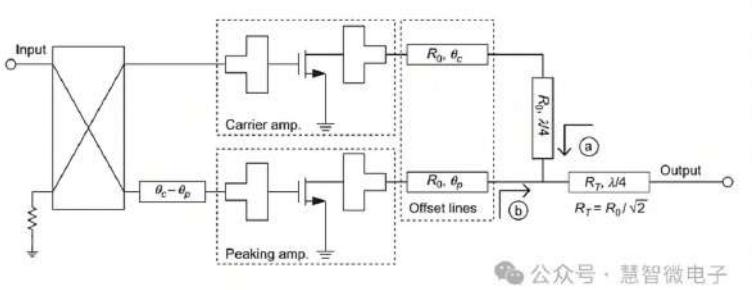


图 2.106: 完整的 Doherty PA 电路

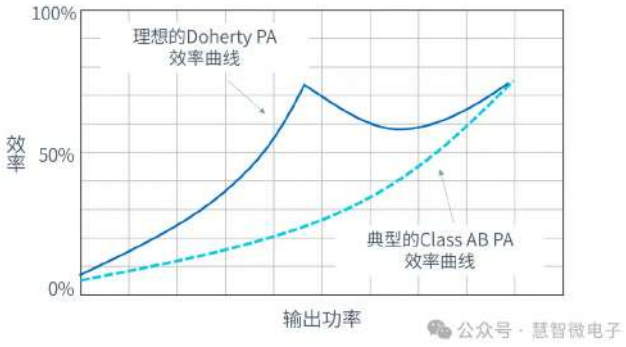


图 2.107: 理想的 Doherty PA 效率曲线

尽管 Doherty PA 在高功率和效率方面表现出色，但它也存在一些明显的缺点。首先，两路 PA 在切换时会影响线性度，因此通常需要与 DPD 技术结合使用；其次，Doherty PA 利用了大量的阻抗变换网络来实现电压源和电流源的等效变换，这些网络的相对窄带特性限制了 Doherty PA 的带宽；最后，Doherty PA 对负载变化较为敏感，需要相对固定的 PA 输出阻抗。

在过去几十年的发展中，Doherty PA 在基站中得到了广泛应用。在这些应用场景中，有强大的 DPD 支持、只需针对特定运营商的较窄频段设计、负载相对固定（如固定天线），这些条件恰好满足了 Doherty PA 的系统需求。

近年来，一些厂商尝试将 Doherty PA 应用于手机中。为了适应手机等移动终端有限的 DPD 算力、全球频段支持和不断变化的负载等挑战，手机中应用的 Doherty PA 通常会弱化其 Doherty 特性。这样一来，PA 就不会对周围环境和带宽过于敏感。然而，这种设计也会相应地削弱 Doherty PA 对效率的改善效果。

Out-phasing 架构的实现

在理解了 Doherty PA 架构之后，Out-phasing 架构就变得相对容易掌握了。

顾名思义，Out-phasing 技术是通过利用不同“相位”（Phase）的 PA 来实现特定设计的。其代表性的架构图如图2.108所示 [7]。

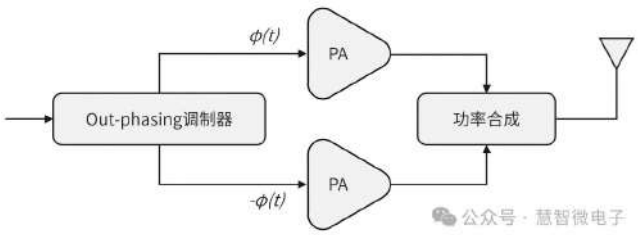


图 2.108: Out-phasing PA 架构

与理解 Doherty PA 类似，我们同样可以将 Out-phasing PA 简化，简化后的等效电路图如图2.109所示。

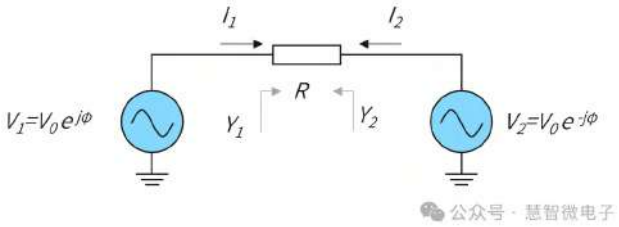


图 2.109: Out-phasing PA 简化电路图

从图中可以看出，负载阻抗 RL 两端的电压随着两个电压源的相位变化而波动，因此流过 RL 的电流也会随之调整。

需要注意的是，每个电压源所见的等效阻抗并非 RL，而是其自身电压与电流的比值。因此，每个电源的等效阻抗也随相位的变化而改变。这意味着，通过调控输入到两路 PA 的相位，我们可以实现对两路 PA 负载阻抗的精准控制。这也就是利用 Out-phasing 架构，对负载进行调制的原理。

Load Modulation Balance PA 架构

Load Modulation Balance PA（LMBA）是射频功率放大器领域的杰出专家 Cripps 在 2016 年提出的一种创新型架构 [8]。该架构的示意图如图2.110所示。

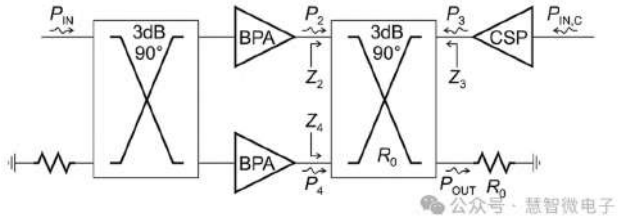


图 2.110: LMBA 架构

LMBA 是在传统 Balance PA（BPA）的基础上进行改造而来的。与传统的 BPA（通常只有两路 PA 进行 90 度合成）不同，LMBA 增加了一路 PA，Cripps 教授将其命名为 CSP（控制信号功率）。CSP 信号被注入到 90 度合成网络的隔离端口中，这一信号的引入会影响两路 BPA 所见的负载阻抗。通过调整 CSP 信号的大小，可以等效地控制 BPA 的负载阻抗，进而影响 PA 的输出功率和效率。

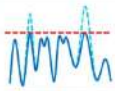
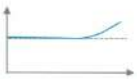
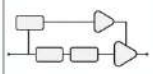
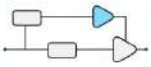
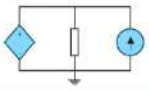
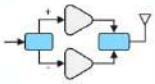
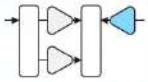
LMBA 的独到之处在于它将影响负载阻抗的 PA 独立出来，并允许通过单独的信号进行控制。这与 Doherty PA 和 Out-phasing PA 只能依赖两路 PA 自行调节的方式不同，从而极大地便利了负载调制 PA 的应用。

基于这一基本架构，一些文章已经开展了对 Doherty 与 LMBA 结合、不同类别的 PA 作为 CSP PA 的影响等方面的研究。对此感兴趣的同学可以查阅相关论文。

2.6.4 总结

在射频系统中，PA 是主要的能耗器件。为了有效降低 PA 的功耗，针对 PA 的研究也不能只停留在 PA 单体层面，还应更广泛地从系统角度来探索解决方案。

从系统角度考虑，优化 PA 效率可能的方案如图2.111。

实现方法	技术方案		架构原理	
改善输入信号	CFR			• 对输入信号进行“削峰”
	DPD			• 输入信号进行预先失真
参数动态调制	Vcc动态调制	EER		• 将包络和相位分别处理
		ET		• 使电源跟踪信号而变
	Bias动态调制	动态Bias		• Bias随功率变化而变化
	Load Modulation	匹配动态调制		• 匹配网络由包络信号控制
		Doherty PA		• 动态调制负载阻抗
		Out-phasing PA		• 动态调制负载阻抗
		LMBA		• 动态调制负载阻抗

公众号 · 慧智微电子

图 2.111: 系统级 PA 效率优化方式

随着技术平台的不断进步和 PA 技术的持续发展，未来肯定还会有更多创新的系统级手段涌现，以进一步降低 PA 的功耗。

您认为还有什么系统级的手段可以将 PA 的功耗降低下来呢？欢迎您评论区留言分享您的见解和建议。

参考文献

1. Farabegoli, A. et al. “A novel crest factor reduction technique using memoryless polynomials.” 2014 44th European Microwave Conference (2014): 825-828.
2. E Kahn, Leonard R.. “Single-Sideband Transmission by Envelope Elimination and Restoration.” Proceedings of the IRE 40 (1952): 803-806.
3. A. A. M. Saleh and D. C. Cox, ”Improving the Power-Added Efficiency of FET Amplifiers Operating with Varying-Envelope Signals,” in IEEE Transactions on Microwave Theory and Techniques, vol. 31, no. 1, pp. 51-56, Jan. 1983.
4. Lee, Byoung-Seok and Lawrence P. Dunleavy. “Understanding Base Biasing Influence on Large Signal Behavior in HBTs.” (2007).
5. K. Chen and D. Peroulis, ”Design of Adaptive Highly Efficient GaN Power Amplifier for Octave-Bandwidth Application and Dynamic Load Modulation,” in IEEE Transactions on Microwave Theory and Techniques, vol. 60, no. 6, pp. 1829-1839, June 2012.
6. B. Kim, Doherty Power Amplifiers from Fundamentals to Advanced. Academic Press, 2018.
7. T. Barton, ”Not Just a Phase: Out-phasing Power Amplifiers,” in IEEE Microwave Magazine, vol. 17, no. 2, pp. 18-31, Feb. 2016.
8. D. J. Shepphard, J. Powell and S. C. Cripps, ”An Efficient Broadband Reconfigurable Power Amplifier Using Active Load Modulation,” in IEEE Microwave and Wireless Components Letters, vol. 26, no. 6, pp. 443-445, June 2016.

在线阅读

本文摘自慧智微微信公众号，扫描以下二维码，在线阅读

2.7 5G PA“记忆效应”的现象、形成与消除

作者：彭洋洋 博士

在 5G PA 设计和应用中，有一个名词经常被大家提及：记忆效应（Memory Effect）。PA 怎么还会有“记忆”？记忆效应有什么影响，又要怎么规避？带着以上问题，本文对 PA 的“记忆效应”做一个讨论。

2.7.1 PA 的“记忆效应”

2.7.1.1 什么是 PA 的“记忆效应”？

记忆效应是指器件的特性随时间变化而发生变化的现象。器件特性随时间变化，当前时间的特性受上一时间状态的影响，就好像器件带有“记忆”一样，所以这种特性被称作“记忆效应”。

对于 PA 电路来说，记忆效应反应出来的现象是：在同样的输入功率下，PA 的增益、延时等特性在不同时刻表现不同。

PA 一旦表现出强的记忆效应，线性度将会受到明显影响，出现宽带线性度恶化、左右 ACLR 不平现象，还会影响预失真（DPD 或 APD，数字预失真或模拟预失真）电路的工作 [1]。在 PA 设计和使用中，记忆效应需要尽量规避。

2.7.1.2 PA 记忆效应的识别

方法一：观测 AM/AM、AM/PM

在 PA 输入端加入输入信号，观察 PA 的输出信号，将此时放大器的增益、相位与输入信号逐点一一对应起来，就得到不同功率下的 AM/AM、AM/PM 特性曲线。观测 AM/AM、AM/PM 是识别记忆效应的最主要方法。

未发生记忆效应时，AM/AM、AM/PM 曲线为一条平滑曲线，代表同样输入功率时不同时刻 PA 的特性保持相同。发生记忆效应时，曲线表现出一组发散分布的点，点的分散的离散与否代表了记忆效应的强弱，如图2.113所示。如果 AM/AM、AM/PM 曲线表现分散，表示对于一个同样的输入功率，PA 有不同的增益、相位变化。PA 的状态不止以此时刻的输入有关，还与其他时刻的状态有关，即 PA 表现出了“记忆效应”。

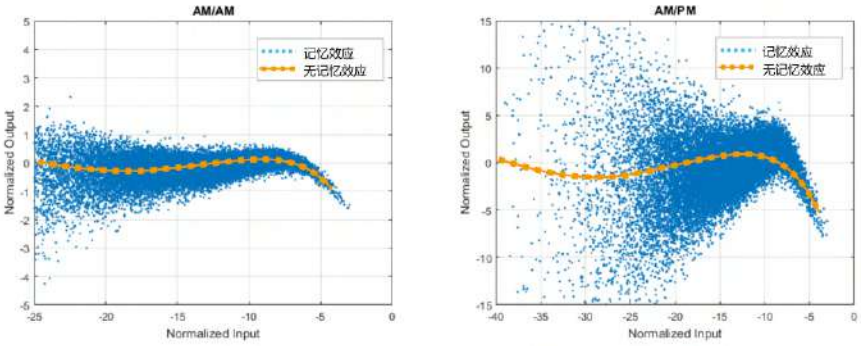


图 2.113: 带有记忆效应的 PA 特性 [1]

方法二：宽带 ACLR 的恶化或不对称

另外一个识别记忆效应的方法是观察 ACLR 随信号带宽的变化。如果观察到信号带宽变宽时 ACLR 出现快速恶化，或者开始出现明显的左右不对称时，很有可能是发生了记忆效应 [2]，如图2.114所示。

ACLR 不对称的产生也对测试造成较大困扰，当左右 ACLR 差别较大时，数据的准确记录都成问题。当发生严重左右不对称现象时，需要停下来检查 PA 是否发生了记忆效应。

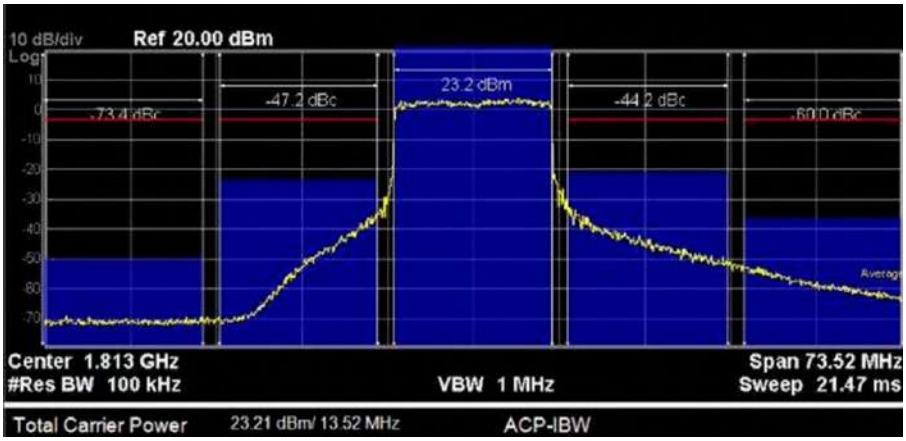


图 2.114: 典型的 PA 左右 ACLR 不对称现象

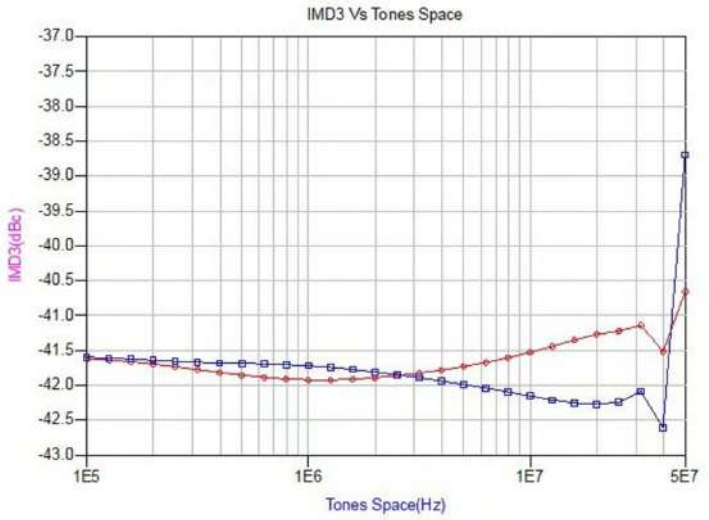


图 2.115: 采用双音测试法识别 PA 记忆效应

方法三：宽带双音信号 IMD3 恶化或不对称

与观测宽带 ACLR 是否恶化方法类似，还可以采用观察宽带双音信号（Two tone）的 IMD3 是否出现记忆效应。

采用双音对 PA 进行激励，并将双音信号的间距逐渐由小变大，测试双音信号的 IMD3。如果随着双音信号的间隔变大，IMD3 出现恶化或左右不对称，则代表 PA 开始出现记忆效应。如图2.115所示测试，当信号间隔大于 10MHz 时，PA 记忆效应开始显现。

2.7.2 PA 记忆效应的产生与规避

在对 PA 记忆效应的分析中，一般认为是输入信号中随时间变化的包络引起了 PA 特性的变化，造成 PA 不同时间的响应不同。为何输入信号会有包络呢？包络又是如何进一步引起 PA 特性变化？

2.7.2.1 4G/5G 信号的包络

4G/5G 系统中所使用的调制方式是 OFDM（Orthogonal Frequency Division Modulation，正交频分复用）调制，OFDM 是一种特殊的 FDM（Frequency Division Multiplex，频分复用）调制方式。

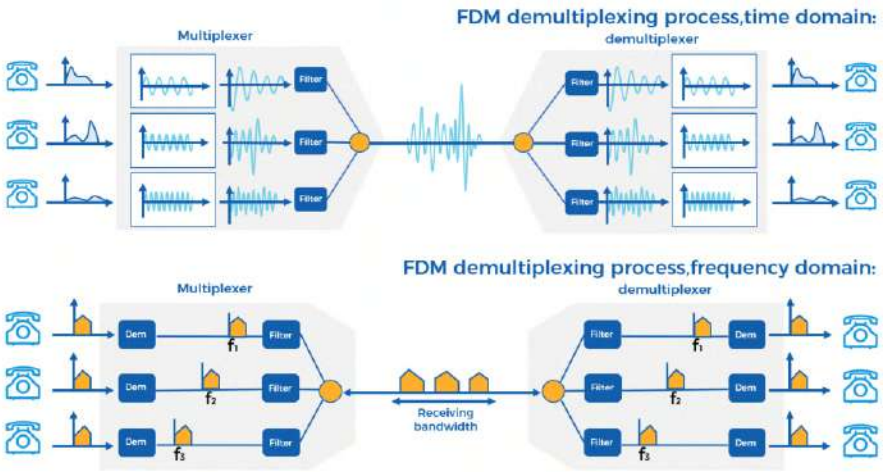


图 2.116: 图：频分系统的时域及频域图示

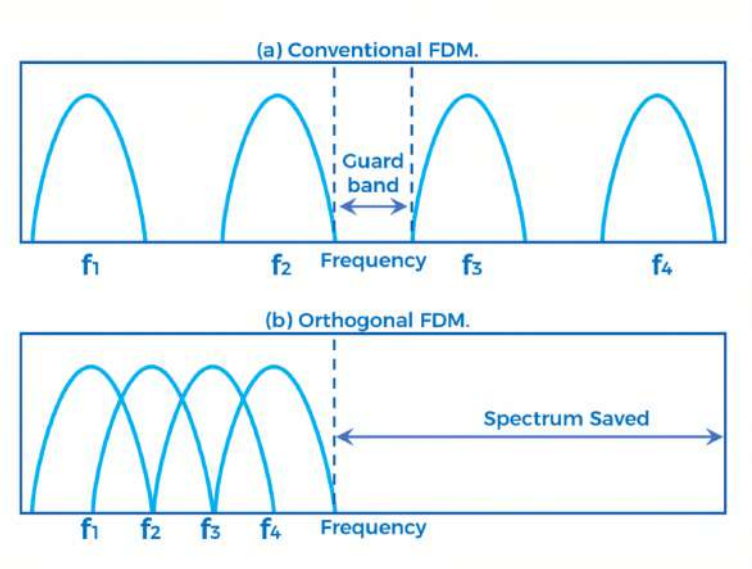


图 2.117: FDM 与 OFDM 的频谱特性 [6]

FDM 是无线通信中的重要复用技术，是指将传输信号调制到不同频率的多载波上，再进行空间传输。与单一频率的单载波信号相比，FDM 可以有效增加带宽，提升传输速率，并且可以增强信号抗干扰能力，在无线通信系统中得到广泛应用。图2.116为 FDM 系统的典型实现框图。

OFDM 本质也是一种 FDM 技术，也是利用频分的方式进行信号的空间复用传输。只不过在信号的正交上，并不只是采用如传统 FDM 相同的频分技术，还利用到了一些信号在正交上的特性，OFDM 利用到的信号正交特性可由如下积分公式表示，当积分区间为 $0 \sim 2\pi$,n,m 为正整数时：

$$\int \cos nx \cdot \sin mx \, dx = 0$$
$$\int \cos nx \cdot \cos nx \, dx = \pi$$

如果将 $\cos nx$ 看成一路信号，$\sin mx$ 或 $\cos nx$ 看成另外一路信号，可以看到相同的信号相乘后积分为 π，不同信号相乘后积分为 0。所以对于混在一起的不同信号，我们只需要用指定信号和其相乘积分，就可以将有用信号提取出来，实现信号的正交。这就是正交频分复用的原理。

利用这一特性，OFDM 中并不需要像传统 FDM 一样保留较宽的保护带，而是可以将频率成倍数关系的子载波叠加在一起，就可以完成频分复用，大大提高了频谱利用效率。FDM 和 OFDM 的频域关系如图2.117所示：

在信号传输时，频域上多个子载波频分复用进行信号传输，时域上这些信号幅度相叠加。如果某一时间点

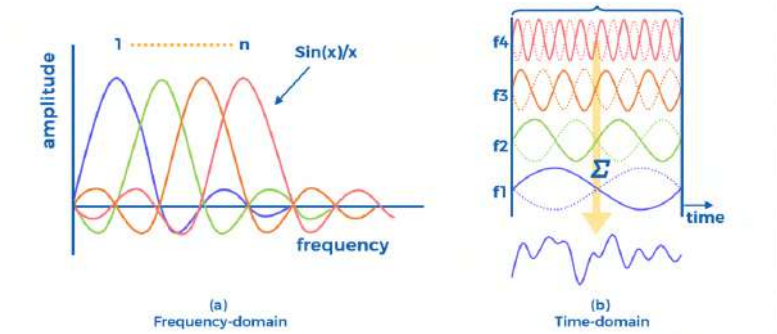


图 2.118: OFDM 信号的时域叠加：形成高峰均比（非恒定幅值）信号

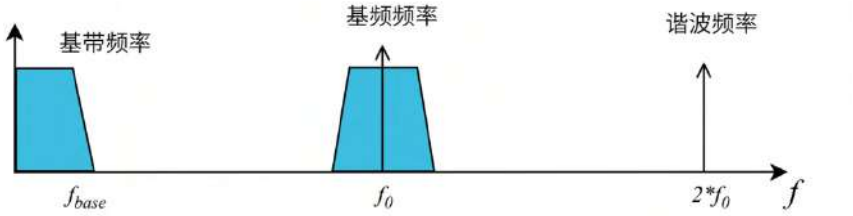


图 2.119: 基带频率、基频频率与谐波（二次谐波为代表）频率之间的关系

几个子载波的振幅均为高位，则叠加出现峰值功率，OFDM 信号的峰均比：PAPR（Peak to Average Power Ratio）就此出现。如图2.118所示。

以上 OFDM 信号的产生均在基频处发生，为了将信号在空间传输，需要将信号调制到射频频率。一般在无线通信系统中，将携带信息的 OFDM 信号称为基带信号，所对应的频率也称作基带频率。由于射频信号的调制和放大经常伴有谐波发生，为了把射频的信号做区分，一般将所用来做射频调制的频率称为基频频率，将信号的谐波称为二次、三次谐波频率。

经过射频调制的 OFDM 信号波形如图2.120所示。

由上图（右）可以看到，OFDM 基带信号被射频频率调制成包络与原始 OFDM 信号相同、载波频率与射频基频信号相同的射频信号。

以上信号将进入功率放大器放大输出，即进入功率放大器的信号并不是恒包络信号，而是包络跟随 OFDM 信号变化、带有峰均比的信号。这就是 4G/5G 中射频信号包络产生的过程。

PA 记忆效应的来源

在 4G/5G 通信系统中所使用的是非恒定包络的 OFDM 信号，如果 PA 特性随信号包络在变化，则会发生记忆效应。

在 PA 记忆效应的分析中，一般将包络信号对 PA 的影响分为三类 [3]，分别是：

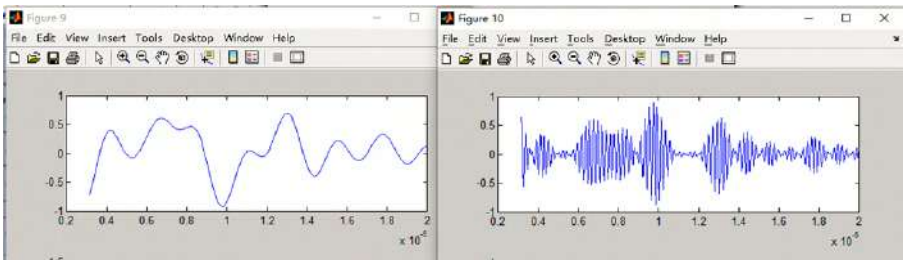


图 2.120: OFDM 基带包络信号的形成（左）及射频基频信号正交调制后的波形（右）

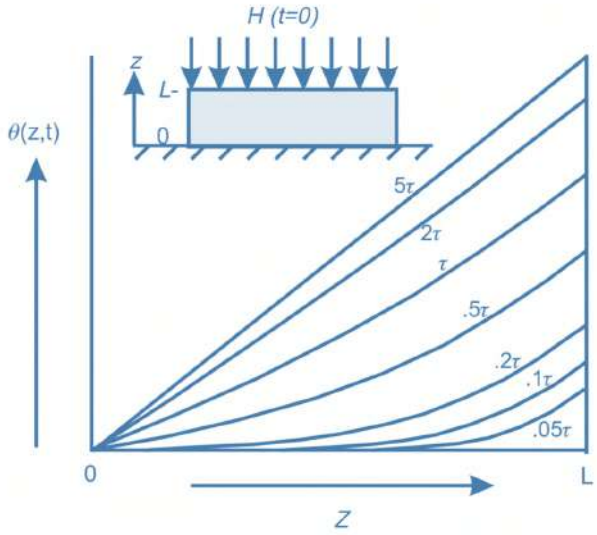


图 2.121: PA 晶体管芯片温度分析的一维等效

- 电记忆效应；
- 电热记忆效应；
- 半导体器件陷波效应（Semiconductor Trapping Effects）

电记忆效应和半导体器件的陷波效应

PA 是高耗电器件，同时 PA 特性对温度敏感。对于 PA 器件而言，热扩散与时间相关。图2.121为对 PA 芯片进行一维等效，在芯片表面加热源后，芯片各部分位置温度随时间变化图。可以看到，随着时间的不同，芯片各位置的温度发生变化 [3]。

如果 PA 的输入信号带有包络，不同包络信号的幅度将会引起 PA 温度的变化，进而引起 PA 特性的不同，使 PA 产生记忆效应。通常情况下，电热记忆效应只与低于 1MHz 的包络频率有关，因此，通常把电热记忆效应认为是长时记忆效应。

针对于电热记忆效应，需要对 PA 进行良好的散热设计，确保 PA 产生的热量可以良好扩散。

半导体的陷波效应（Semiconductor Trapping Effects）是指在半导体器件中表现出来的记忆效应，此种记忆效应与半导体制造工艺相关。对于半导体陷波记忆效应，设计上可尝试的方法并不多，需要代工厂在器件上做优化和保障。设计中可以采用较为成熟的半导体工艺，确保器件侧无记忆效应问题。

电记忆效应

电记忆效应是 PA 电路记忆效应的主要来源。

在 PA 设计中，PA 能量的来源是直流供电。直流供电需要通过偏置电路加在器件的输出（集电极或漏极）及输入（基极或栅极），这些偏置电路需要完成供电功能，并且不能影响到射频信号传输。

理想情况下，偏置电路需要对射频信号表现出高阻（射频信号才不会进入到偏置电路中），同时对其他频率保持低阻（防止其他频率出现交流电压摆幅），如图2.122所示。对于射频信号的包络频率来说，由于其与射频载波信号的二阶交调（IMD2）刚好会落在信号的旁带，所以射频信号包络频率的阻抗控制尤其重要。如果处理不好，偏置电路将会发生包络调制，引起 ACLR 恶化。

以上图所示简单偏置电路来理解包络信号对 PA 偏置的调制作用。图示为简单共射极放大电路，放大器集电极通过串联电感连接至直流供电电源，Bypass 电容提供射频接地，电感 LC 提供高的射频阻抗。

如果假设 PA 工作射频频率为 2.6GHz，电感 LC 的感值为 3nH，供电电压为 3.4V，电流为 500mA。则电感 LC 可提供的射频阻抗为 $Z_{Bias@2.6GHz}$ = 49Ω，由于此点的射频负载阻抗 $Z_{Load@2.6GHz} \approx 3.5\Omega$，在 2.6GHz 射频工作频率，$Z_{Bias}$ 远大于 Z_{Load}，偏置网络可以提供较高的射频阻抗，使射频信号不进入偏置网络内。

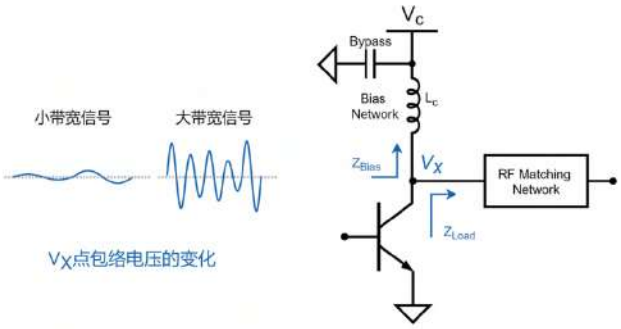


图 2.122: 简单 PA 偏置网络示意图

如果此 2.6GHz 信号带有 10MHz 信号带宽，位于基带频率的 10MHz 的包络信号在经过 LC 产生的压降为 0.1V 左右，晶体管接收到的将不是一个稳定的 3.4V 供电电压，而是一个受包络调制，摆幅 0.1V 的变化电压。如果晶体管特性在 0.1V 供电电压波动时不产生变化，输出信号将看不到记忆效应。

如果这时将输入信号带宽提高到 100MHz，则 LC 产生的包络频率压降将提升至 1V 左右，这将对一个 3.4V 供电的电源带来明显影响，如果晶体管特性在 1V 的变化下发生明显变化，记忆效应将显现。所以，随着信号带宽的增加，偏置网络引起的记忆效应现象将更加明显。这也是为什么 5G 相比 4G 记忆效应问题更加突显的原因。

以上是以输出偏置网络为例对电记忆效应其中一个成因的简单描述。在 PA 设计中，PA 输入偏置网络、阻抗匹配网络的记忆效应均需要在设计中加以考虑。

对包络信号引起 IMD3 变化的定量分析如下。

输入放大器带有一定带宽的 OFMD 信号调制信号可以简化为一组双音信号，频率分别为 ω_1，ω_2，在只考虑 HBT 器件非线性跨导时 g_{m2}，IMD3 信号产生过程如下：

- 输入端双音信号通过 HBT 器件的二阶非线性跨导 gm2 在输出端产生包络信号 $A_2(\omega_1-\omega_2)$、$A_2(\omega_2-\omega_1)$；
- 输出端包络信号通过寄生电容和源极电感等反馈通路反馈到输入端，经过反馈通路后到输入端包络信号为 $B_2(\omega_1-\omega_2)$、$B_2(\omega_2-\omega_1)$；
- 输入端的包络信号和输入端的基频信号 ω_1 再次通过 HBT 的二阶非线性跨导 g_{m2} 产生二阶交调，在输出端产生 IMD3 信号 $H_3(\omega_1-\omega_2)$、$H_3(\omega_2-\omega_1)$、，造成线性度的恶化；

过程如图2.123所示。

以上为宽带信号 IMD3 产生及影响线性度恶化的过程，通过 Volterra 级数推导，由二次跨导非线性 g_{m2} 和包络信号产生的 IMD3 为：

$$H_3(2\omega_1-\omega_2)=\frac{3}{4}\cdot A_1\cdot\omega_1^2\cdot A_1\cdot(-\omega_2)\cdot(-\frac{2}{3}\cdot\frac{g_{m2}^2}{g_m+g(\omega_1-\omega_2)})\cdot(1+\frac{Z_s(s)}{Z_1(s)})$$

通过以上公式也可以计算得到左侧 IMD3 $H_3(\omega_1-\omega_2)$、$H_3(\omega_2-\omega_1)$ 的表达式。需要说明的是，由于 $A_2(\omega_1-\omega_2)$ 及 $A_2(\omega_2-\omega_1)$ 位于零频两侧，二者阻抗不同，会导致第三步左右 IMD3 混频结果不对称，这也是左右 ACLR 不对称的来源。

当信号带宽变大时，即 ω_1 与 ω_2 频率相差变大。如果不加以控制，$A_2(\omega_1-\omega_2)$ 与 $A_2(\omega_2-\omega_1)$ 之间的阻抗差别也会变大，造成宽带信号 ACLR 变差，这是宽带信号记忆效应表现更明显的原因。

可以看到，IMD3 信号的大小与器件的线性与非线性跨导相关，同时与源端、负载端等阻抗相关。以上公式可以用来设计和控制包络信号产生 IMD3 分量的大小。

电热记忆效应的规避

电记忆效应的来源是由于在 PA 设计中，谐波及包络频率阻抗过大引起，在设计中，要注意对这些阻抗加以处理，将包络频率与谐波频率的阻抗尽量归零。

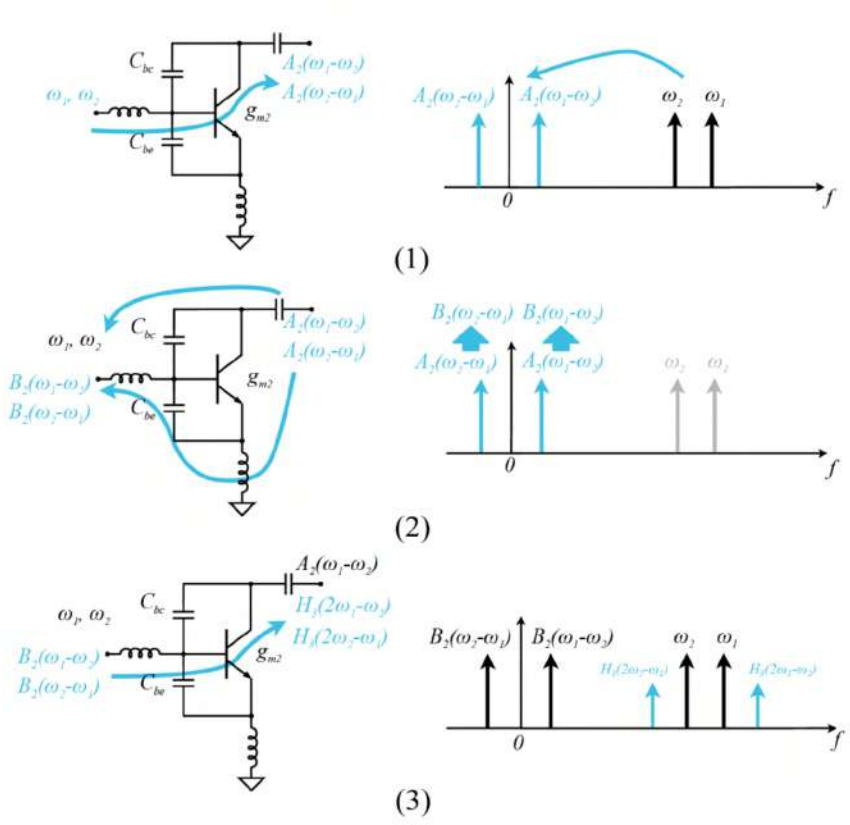


图 2.123: 宽带信号激励下，由于器件非线性引起 IMD3 的过程

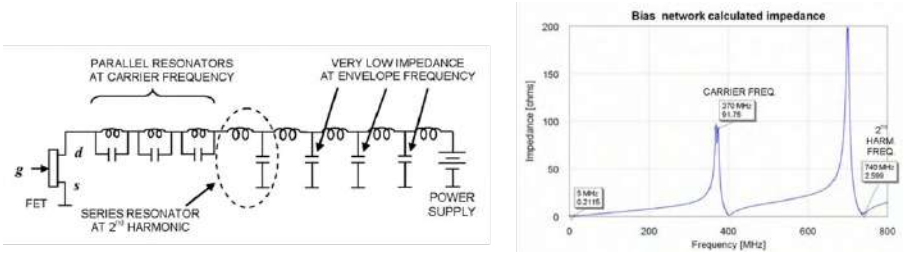


图 2.124: 文章 [4] 所提出偏置网络及其阻抗特性

文章 [4] 中提出了一种供电电路设计，此电路可以对基频信号提供较高的阻抗（91.75Ω），同时使包络信号及谐波信号保持较低的阻抗（0.2Ω 及 2.6Ω）。采用此偏置网络，PA 线性度和 ACLR 恶化现象明显改善。如图2.124及图2.125所示。

在 PA 芯片的手机应用中，由于偏置网络在芯片内部已经设计好，重新设计偏置网络不是一个可行的选项。但在应用中观测到记忆效应时，还是需要对板级 Bypass 电容进行良好设计。

一般在 PA 芯片内部，Bypass 电容为 100pF 量级（在 2.6GHz 阻抗为 0.6Ω 的低阻），起到对射频信号旁路功能。但 100pF 电容在 100MHz 处的阻抗为 16Ω 左右，无法为基带频率及包络频率提供低阻，所以，一定要在板级放置 100nF 量级 Bypass 电容，为包络信号提供近似理想接地，降低电源引起的记忆效应。如图2.126所示。

器件的非线性引起的记忆效应需要借助对器件的非线性分析进行消除，常用分析方法是 Volterra 级数法等。通过对放大器偏置点、阻抗及非线性的控制，可以使器件非线性引起的记忆效应加以改善 [7]。

2.7.3 总结

PA 记忆效应是较为抽象的概念，反应 PA 特性随时间的变化而变化，造成输出信号的失真。

PA 记忆效应是由包络信号的幅度变化引起，OFDM 调制带来的信号峰均比经过射频调制后，形成包络变化

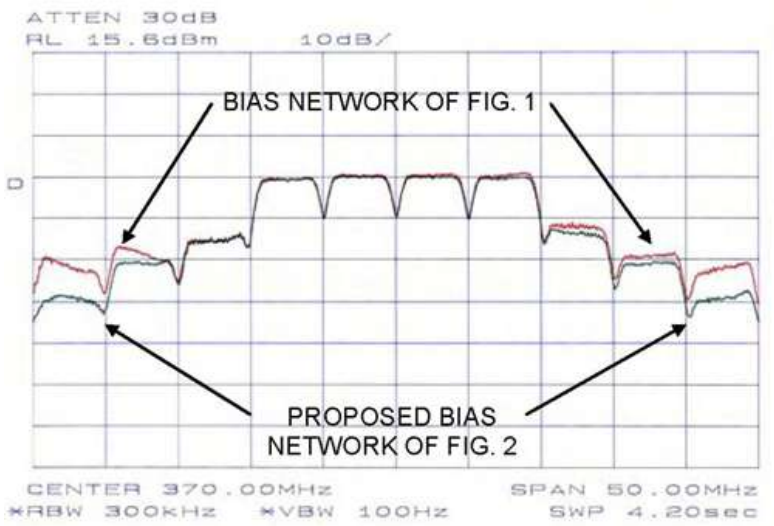


图 2.125: 文章 [4] 所提出偏置网络对线性度有明显改善

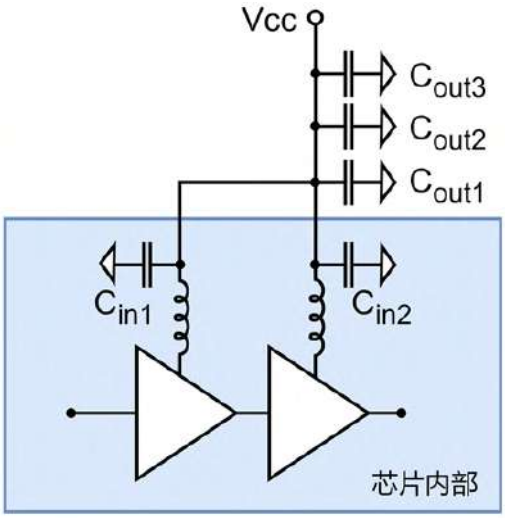


图 2.126: 图：典型 PA 芯片的 Bypass 电容放置

的射频信号。如果 PA 特性随包络变化而变化，则将发生记忆效应。

电记忆效应是 PA 记忆效应最主要的来源，通过优化偏置电路、节点阻抗可以有效地改善电记忆效应。在应用中，需要注意偏置电路中的 Bypass 电容设计，一定要在包络频率提供较低阻抗，确保包络信号在偏置电路中不被调制起来。

在 5G PA 设计中，慧智微利用可重构技术架构，可以对功率放大器各频点阻抗进行调谐控制，有效解决 5G PA 的记忆效应问题，使 5G PA 大带宽线性度明显改善。欢迎联系慧智微销售及技术支持，获取更多产品信息。

以上就是 PA 记忆效应的讨论，您在 PA 应用中是否遇到过记忆效应？看到的是什么现象，又是如何规避的？欢迎扫描文章末尾二维码，留言讨论。

参考文献

1. NandiO P . Behavioral Modeling of Power Amplifier with Memory Effect andLinearization Using Digital Pre Distortion. 2016.
2. Williams D J , Leckey J , Tasker P J . A study of the effect of envelope impedance onintermodulation asymmetry using a two-tone time domain measurement system[C] // International Microwave Symposium Digest. IEEE, 2002.
3. Cripps S C . RF PowerAmplifiers for Wireless Communications / S.C. Cripps. 2006.
4. Franco M , Guida A , Katz A , et al. Minimization of Bias-Induced Memory Effects in UHF RadioFrequency High Power Amplifiers with Broadband Signals[J]. IEEE, 2007.
5. Carvalho N B . Acomprehensive explanation of distortion sideband asymmetry[J]. IEEE Trans.Microw. Theory Tech, 2002, 50.
6. KodheliO . OFDM-based schemes for next generation wireless systems.
7. Carvalho N B , Pedro J C . Two-tone IMD asymmetry inmicrowave power amplifiers[C]// IEEE. IEEE, 2000:445-448.

在线阅读

本文摘自慧智微微信公众号，扫描以下二维码，在线阅读

2.8 如何才能不烧射频 PA?

作者：彭洋洋 博士

PA（Power Amplifier，功率放大器）是通信系统的重要组成模块，负责将射频信号的放大与功率输出。PA 的性能及可靠性对整个发射系统有重要的影响。

在手机通信系统中，PA 的输出信号约为 29~32 dBm (约 1,000 mW) 的大功率，对比收发机芯片 0dBm (1mW) 左右的功率输出，PA 输出功率大 1,000 倍。如图2.128所示。

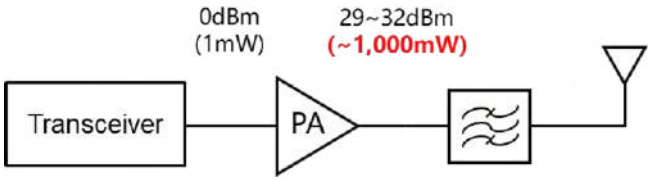


图 2.128: 典型手机通信系统链路

2.8.1 为什么会烧 PA

2.8.1.1 手机 PA 芯片中的半导体器件

PA 设计常用的半导体工艺为 GaAs HBT、GaAs pHEMT、SOI 、Bulk CMOS（体硅 CMOS）。其中，GaAs HBT 因为其功率密度大，成本低（相较于 GaAs pHEMT），成为射频功率放大器功率输出级的首选工艺。

HBT 器件有三个极限参数，分别为：

1. 集电极最大允许电流 Icmax
2. 集电极最大允许热耗散功率
3. 集电极-发射极反向击穿电压 BVCEO

在实际工作中，需要保证器件的最大工作电流、耗散功率、最大电压，在额定范围内。

2.8.1.2 PA 的电压/电流摆幅

对于射频 PA 的 HBT 器件，输出晶体管的集电极有最大电压及电流摆幅。此点电压与电流之间的摆幅随时间变化，为了更好的表征电压与电流之间的关系，通常用 Load-line（负载线）的方式，将电压与电流画在同一幅图中，如图2.129所示。

晶体管的负载线反应了在不同的负载下，晶体管电压与电流间的相互关系，一般画在 DC-IV 曲线上。在负载线中：

1. 负载线的斜率反应了负载阻抗的大小。
2. 实际电路中，由于负载虚部的存在，会形成电压/电流相位差，可能使得负载线表现为中空的环形。
3. 直流工作点的选择（电压及电流），对负载线的摆幅有影响。

在 PA 工作中，由于高电源电压、高 VSWR 的存在，会让 PA 输出电压、电流摆幅增大。图2.130为一个典型的功率放大器，分别是在 3.2V 及 5V 偏置电压下，50Ω 及 VSWR=10:1 下的动态负载线 [2]。可以看到，在大电压及 VSWR 下，PA 将需要承受更大的电压及电流摆幅。当电压及电流摆幅超过器件的耐受值时，就会造成器件烧毁。

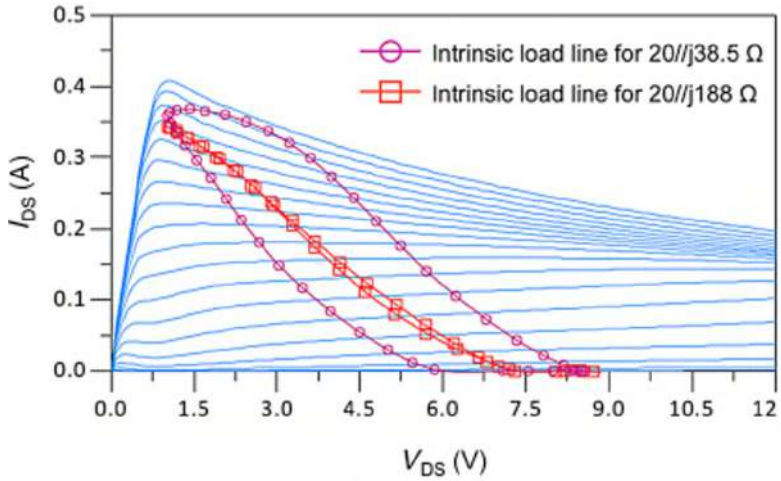
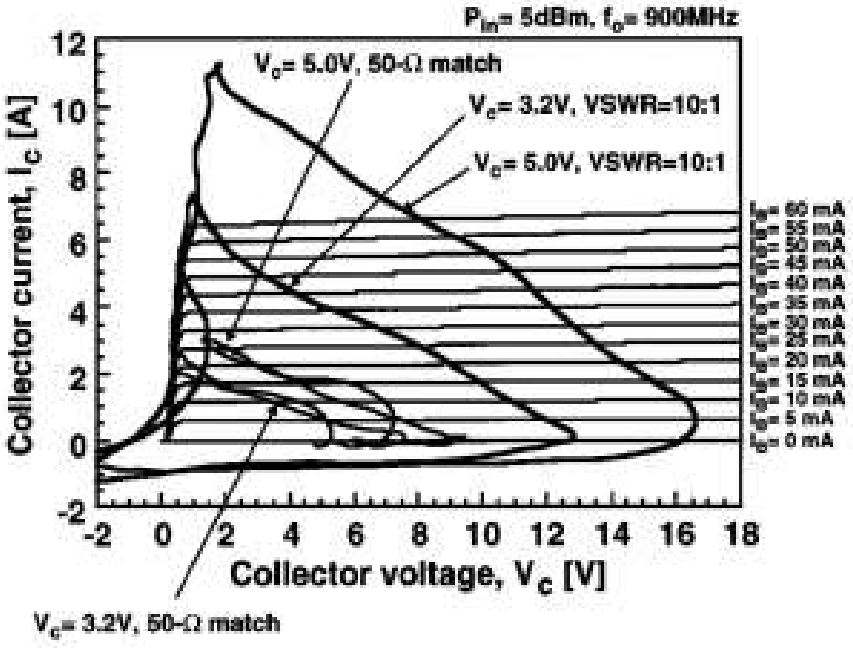


图 2.129: 一款 PA 的典型的动态负载线



Vcc (V)	VSWR	Vcc Max (V)	Icc Max (A)
3.2	1:1	7.5	2
5.0	1:1	9.5	3
3.2	10:1	12.5	7
5.0	10:1	16.5	11

图 2.130: 不同电压及 VSWR 下，PA 动态负载线的变化

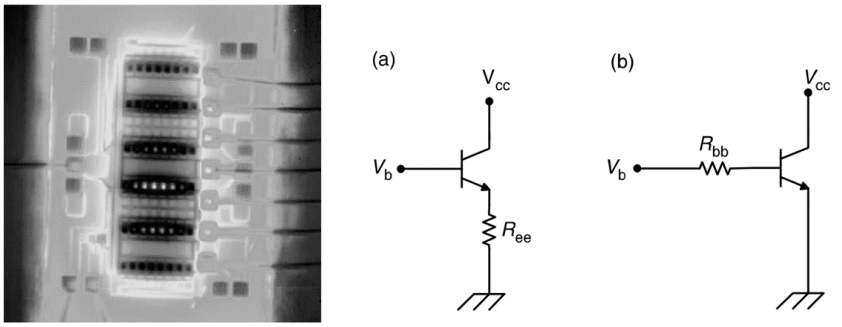


图 2.131: PA 的热分布不均（左），及 Ballast 电阻的设计

2.8.2 如何保护 PA 不被烧毁?

2.8.2.1 设计保障

PA 需要合理设计，确保满足 Ruggedness 需求。

电流设计保障

需要合理设计器件的尺寸，确保在各个条件下，器件所通过的最大电流，小于器件的最大耐受电流。

在对通流的设计中，需要着重注意的是，PA 末级并联了多个晶体管器件，需要保证电流均分在整个器件中，而不是所有电流集中于某一个器件，将器件烧毁。

由于 HBT 器件开启电压随温度升高而降低，过大的电流会降低开启电压，同时使电流进一步增大，直到器件烧毁。这种效应叫做 Thermal Run-away，是电流烧毁中一种常见形式。

为了防止 Thermal Run-way 的发生，需要在晶体管的 Base 端或者 Emitter 端加入 Ballast 电阻。Ballast 电阻的存在，使得电流变大的过程中，Vbe 电压会减小，防止电流的进一步增大 [3]。如图2.131所示。

电压设计保障

对于电压防护，一般采用在末级晶体管 Collector 并联放置二极管串的方式进行稳压，使得输出摆幅稳定在二极管串的开启电压。

在电压防护电路设计中需要注意，一定要保证防护电路放置位置的对称性，确保所有器件的电压摆幅得到保护。

电路设计保障

为了保护射频 PA 工作再合理的工作范围，设计公司还会对 PA 加入过压保护（OVP）、过流保护（OCP）、过驱动保护（ODP）等保护电路。在电压、电流及输入信号超过额定范围后，对以上参数进行钳位保护。

2.8.2.2 Ruggedness 测试

由于 PA 可靠性难以靠仿真来准确设计，PA 设计完成后，必须通过完整的 Ruggedness 测试，来确保 PA 的可靠性。

完整的 Ruggedness 测试环境如图2.132所示。

Ruggedness 测试需要涵盖如图2.133测试条件。

以上测试项需要交叉组合，确保任意条件下，PA 均不会有 Ruggedness 问题。

由于半导体器件的最大击穿电压 BVCEO 随温度降低而降低，PA 增益随温度降低而升高，一般 Ruggedness 最恶劣的点发生在低温。所以，一般低温下，最大输入功率、最高电压、最大 VSWR 下，为 Ruggedness 最差条件。



图 2.132: Ruggedness 测试环境

测试项	条件
Pin	应用中最大的输入功率
Vcc	应用环境中最大的电源电压范围
VSWR	应用环境中天线失配引起的最大驻波比变化
温度	涵盖应用中所有温度范围
波形	涵盖应用中恶劣的应用波形

图 2.133: Ruggedness 测试需要涵盖的测试条件

2.8.2.3 应用保障

虽然合格的 PA 在出厂前，进行了完整的 Ruggedness 测试，但在应用中仍然需要对应用环境加以注意，确保 Ruggedness 在应用中得到保障。应用中需要的主要保障如下：

1. 适当控制电源电压
   由图 3 所示，PA 在低电压应用时，有较小的电压及电流摆幅，PA 的 Ruggedness 将得到较好的保障。所以在应用中，适当控制电压电压，尽可能使用较低的电源电压，有助于提升器件的 Ruggedness。
2. 适当控制输出功率
   大功率输出时，PA 输出将有更大的电压电流摆幅。在应用允许范围内，适应控制输出功率，将有助于 Ruggedness 提升。
3. 注意电源完整性及信号时序
   手机是一个相当复杂的系统，涉及到多个模块之间的联动。在应用中，需要着重注意电源完整性（是否有过高的电压脉冲）、偏置控制信号的时序、输入信号的大小及时序，来确保 PA 是被工作在正常的状态。

2.8.3 总结

PA Ruggedness 设计是一个复杂工程，与器件物理、电路设计、系统应用均相关。在 PA 设计中，一定要对 Ruggedness 仔细设计，才可以确保手机在各种环境应用中，均不会出现“烧片”。

由于篇幅所限，无法对“烧 PA”的底层原理及规避展开详细讨论。欢迎扫描文章末尾二维码，参与更多在线讨论。

参考文献

1. M.-D. Tsai et al., ”10.6 a 4G/5G cellular transmitter in 12nm FinFET with harmonic rejection”, IEEE ISSCC Dig. Tech. Papers, pp. 182-184, Feb. 2020.
2. K. Yamamoto, S. Suzuki, K. Mori, T. Asada, T. Okuda, A. Inoue, et al., ”A 3.2-V operation single-chip dual-band

AlGaAs/GaAs HBT MMIC power amplifier with active feedback circuit technique", IEEE Journal of Solid-State Circuits, vol. 35, no. 8, pp. 1109-1120, 2000.

3. Golio, Mike, Golio, Janet , RF and Microwave Circuits, Measurements, and Modeling (The RF and Microwave Handbook, Second Edition)

在线阅读

本文摘自慧智微微信公众号，扫描以下二维码，在线阅读

2.9 PA“自激”问题的分析与解决

作者：彭洋洋 博士

作为高频、高增益及高功率器件，PA（Power Amplifier，功率放大器）的“自激”是必须要认真设计与考虑的问题。

“自激”译自“Self-Oscillation”，是指电路在非激励的频率下，自我产生周期性信号并维持的现象 [1]。“自激”的发生一般是因为电路环路出现不稳定，在某个频率产生“震荡”。所以“自激”问题又叫“稳定性（Stability）”问题、或“震荡（Oscillation）”问题。

PA 作为射频系统里输出功率最大、增益较高、应用环境最复杂的器件，是系统中最易发生“自激”的电路模块。在大功率 PA 中，一些自激现象有可能产生不可控的大功率及大电流，进而烧毁 PA，造成 Ruggedness 问题；即使一些自激现象轻微，不至于使器件损坏，但这些杂散也会恶化射频系统收发性能，也需要避免。图2.134为发生自激的典型频谱。

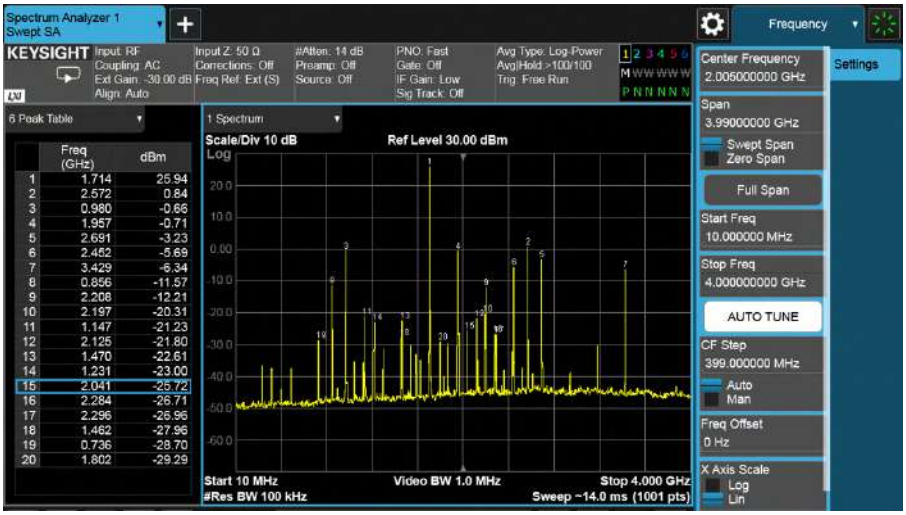


图 2.134: PA 的自激现象（主频信号为 1.714GHz，其他杂散信号为自激产生）

PA 是一个带有非线性的功率输出器件，其稳定性问题来源复杂，较难分析。本文尝试将稳定性的主要来源，以及分析、解决思路做一个梳理，讨论如何规避 PA 的自激。

2.9.1 “自激”的分类

“自激”并不是当前半导体 PA 的特有现象，而是伴随着有源电路的产生一直存在，早在 100 年前的真空管时代，稳定性问题就已经得到重视并开展研究 [2]。对于射频电路设计者来说，“自激”也并非一直是“噩梦”，在 VCO（Voltage Control Oscillator，压控震荡器）设计中，就需要建立起稳定的自激来产生需要的本振信号。

根据产生的原因，主要可以将“自激”分为如下两类：

- 线性自激
- 非线性自激

线性自激

线性自激是指由于耦合、正反馈等线性耦合回路引起的环路自激，一般的低频、高频震荡均属于线性自激。

典型的线性自激震荡电路如 Hartley Oscillator 和 Colpitts Oscillator。在射频 PA 电路中，带有高通反馈特性的 Hartley Oscillator 与 PA Collector/Base 间反馈引起的震荡成因更为接近，Hartley Oscillator 的分析理解方法可用于部分线性自激分析中。如图2.135所示。

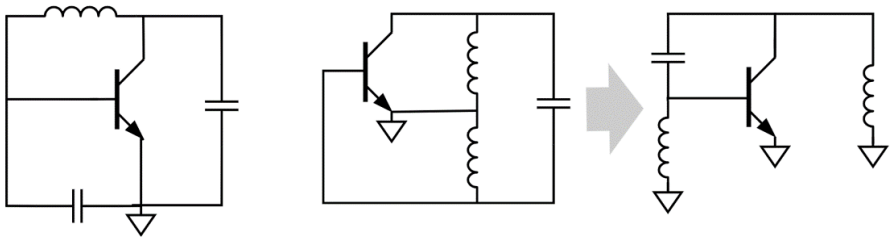


图 2.135: （a）Colpitts Oscillator （b）Hartley Oscillator （c）Hartley Oscillator 电路针对于 PA 的电路等效

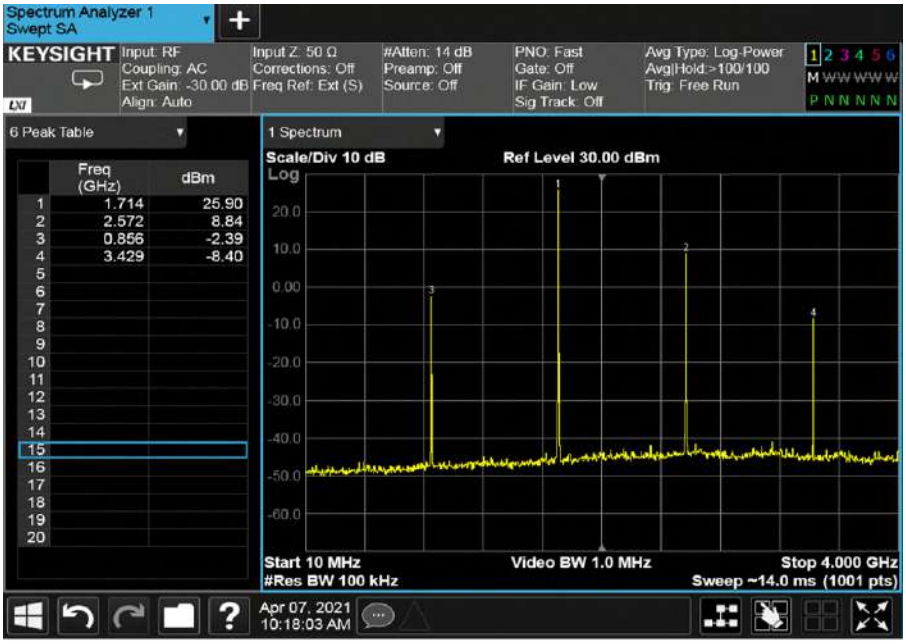


图 2.136: 非线性自激产生半频震荡的典型频谱

非线性自激

非线性自激是指由于器件的非线性引起的自激，一般分为谐波自激与次谐波（半频、1/3 频等）自激。这些自激的发生与器件特性随参数变化而变化有关，所以又被叫做参数震荡（Parametric Oscillation），参数震荡在过去 50 年里也被广泛研究。如图2.136所示。

以下将对两种“自激”的分析与解决展开讨论。

2.9.2 “线性自激”的分析与解决

2.9.2.1 “线性自激”的分析

线性自激产生的原因是环路发生了正反馈，一般的分析方法有：环路分析法、负阻分析法、S 参数网络分析法等。

- 环路分析法是分析反馈环路，看是否存在环路增益大于 1 的现象；
- 负阻分析法是指分析发生自激部分电路的阻抗，如果是负阻且负载阻抗小于负阻阻抗的绝对值，则将可能发生自激；
- S 参数网络分析法是利用 S 参数的方法，进行稳定性分析。

以上分析方法在不同的电路设计中都有采用，不同方法殊途同归，可相互转化，均可以对“自激”现象进行分析。以下采用环路分析法对自激问题进行说明。

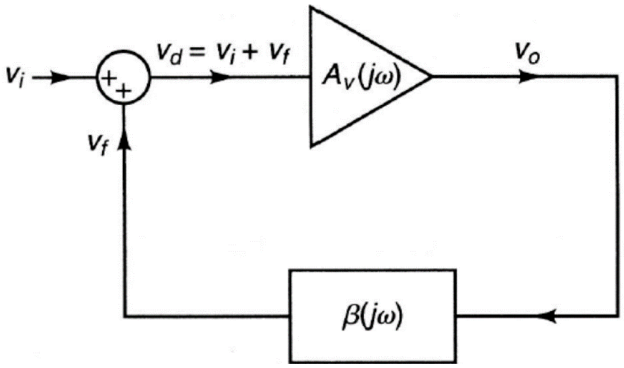


图 2.137: 带有反馈的射频环路

对于带有反馈的射频环路，可表示为如图2.137所示。在此环路中：

$$v_o = A_v(j\omega)v_d$$
$$v_f = \beta(j\omega)v_o$$
$$v_d = v_i + v_f$$

环路增益可表示为：

$$A_{vf}(j\omega) = \frac{v_o}{v_i} = \frac{A_v(j\omega)}{1 - \beta(j\omega)A_v(j\omega)}$$

$\beta(j\omega)A_v(j\omega) = 1$ 为震荡发生的边界条件。Hartley 及 Colpitts 结构形成的震荡均可以用环路分析法进行分析。

当满足震荡条件时，白噪声在环路中不断被放大强化，产生自激。因为白噪声在所有频率范围均存在，所以任何频率点满足震荡条件，均会发生震荡。

2.9.2.2 线性自激”的规避

PA 设计侧的规避

在 PA 设计时，需要对自激问题做仔细排查与规避。射频 PA 一般通过 S 参数网络分析法进行分析，S 参数网络分析法在教科书中均有详细讨论，在此不做过多讨论。

需要注意的是，通过 S 参数进行稳定性分析有以下限制：

- S 参数是基于小信号的参数，所分析出来的稳定性是小信号状态下的稳定性。
- S 参数分析法依赖于模型的准确性和完备性，如果耦合路径在模型中没考虑，分析结果将产生偏差。

S 参数网络分析后，会得到稳定系数和稳定圆，一般设计中需要保证全频段稳定系数大于 1，即绝对稳定。如果在设计中发生稳定性系数有小于 1 的现象，则需要在设计中进行规避。在链路中增加损耗性器件是一种常见的设计方法。在增加损耗性器件时，需要根据不稳定的特性进行设计，有效解决稳定性问题的同时，尽量小的影响射频性能。

图2.138为典型的稳定性改善电路，不稳定区域在 Smith 圆图左侧，此时在输入端串联电阻最为有效，而并联电阻并不能有效改善此电路的稳定性 [3]。

损耗型网络也可以设计成带有频率响应，来针对性的改善某个频点的稳定性。图2.139分别为改善低频和高频稳定性的典型电路。

PA 应用侧的规避

“线性自激”一般是由于信号的耦合产生正反馈引起，如果耦合发生在应用侧，则可以通过应用侧的方法进行规避。

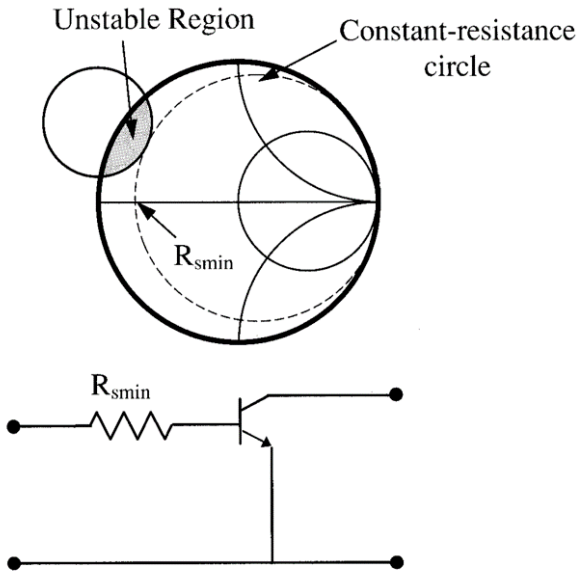


图 2.138: 典型稳定性的改善：增加链路损耗 [3]

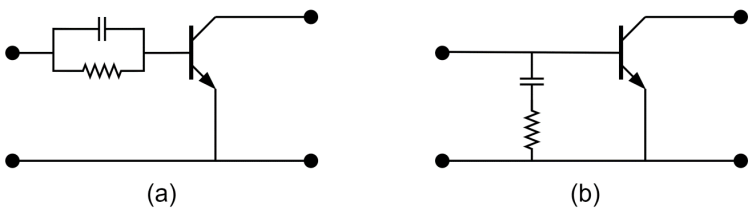


图 2.139: （a）改善低频稳定性的电路（b）改善高频稳定性的电路

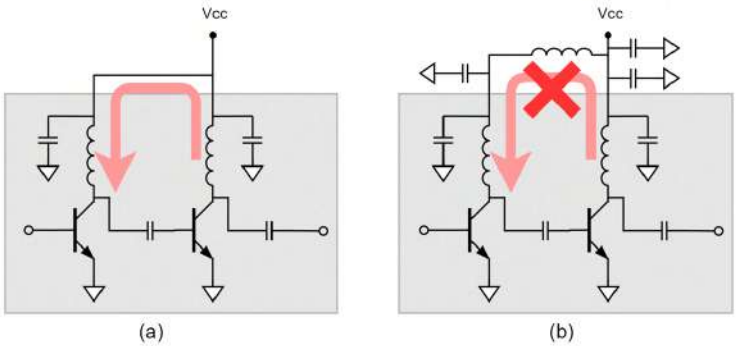


图 2.140: 电源引起的低频耦合及规避（a）通过电源的低频耦合路径（b）通过电源的低频耦合路径规避

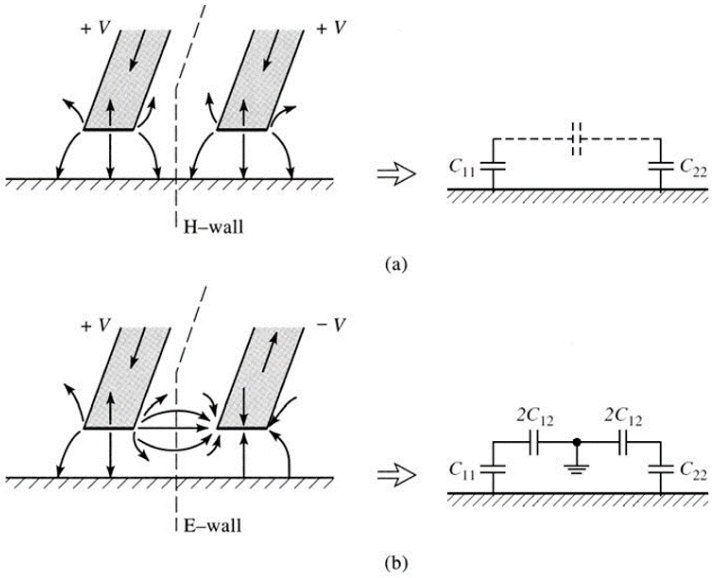


图 2.141: 耦合传输线的等效模型 [4]

在应用侧常见的信号耦合路径有：

- 通过电源线引起的低频耦合。
- 通过信号线引起的高频耦合。
- 由于接地不良引起的共模耦合。

通过电源线引起的低频耦合

在低频段，一般每降低 10 倍的频率，晶体管的最大可用增益会提高 20dB，晶体管在低频段有极高的增益，如果对低频的信号处理不好，则会在低频引起震荡。

低频信号耦合的最主要通路是电源线。由于电源的馈电和去耦网络的低通特性，如果去耦不充分，则电源处有大量被放大的低频噪声信号，当这些信号反馈到前级时，将有可能发生震荡。

如果发生低频震荡，需要仔细检查电源处的去耦电容是否使用得当，是否存在将多级电源拉在一起的现象。规避的方法是将去耦电容仔细设计，必要时可利用去耦电容的自谐振特性，构建某个频率点的陷波；并且将不同级的电源线尽量区隔，可在不同级电源线间串接电感增加隔离。如图2.140所示。

通过信号线引起的高频耦合

对于非接触射频走线，一般频率越高耦合越明显。对于高频信号，需要避免因为空间耦合引起的信号正反馈。对于应用中常用的微带线结构，耦合传输线的等效模型如图2.141所示 [4]。不同传输线间的耦合等效为并联电容。

在规避上，需要仔细检查信号走线，由于输入和输出间放大器增益较大，需要尤其注意输出和输入之间的信号反馈。

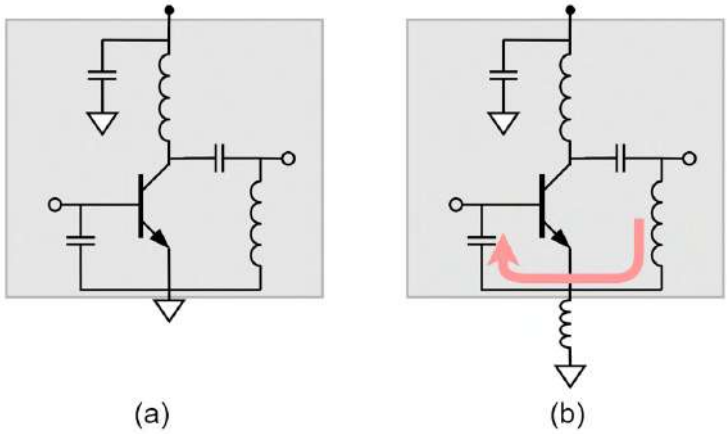


图 2.142: （a）良好接地（b）非良好接地引起的共模耦合

由于接地不良引起的共模耦合

在手机 PA 设计中，一般是采用多颗晶圆在基板上进行系统级集成实现。基板上接地通孔的存在，使得 PA 模组电路存在共模电感。如果在应用中发生接地不良，将会增加共模电感的感值，进而增加模组内部的信号耦合。如图2.142所示。

在规避上，需要确保模组芯片接地良好，减少共模电感值。

2.9.3 “非线性自激”的分析与解决

PA 是一个非线性器件，其非线性主要表现为两个方面，首先 PA 一般工作于 Class AB 状态，本身会产生非线性；其次是用于设计 PA 的 HBT 器件本身的 CBC、CBE 电容等存在非线性。PA 的非线性可以用 Volterra 级数来表示。

$$v_o = a_1 v_i + a_2 v_i^2 + a_3 v_i^3 + a_4 v_i^4 + a_5 v_i^5 + \ldots$$

由于器件的非线性引起的自激称为“非线性自激”，非线性自激分为两类，分别是：

- 谐波自激
- 次谐波自激

谐波自激是指发生“自激”的输出信号处于输入信号的谐波频率。由于谐波能量产生稳定，频率可控，可通过滤波器进行滤除，并且不会对主信号能量产生明显影响。谐波自激一般并不作为通常的“自激问题”进行分析和处理。

次谐波（Sub-harmonic）自激是指发生于信号次谐波（半频、1/3 频等分数频）频率的自激。由于次谐波发生的频段可能是有用频段，并且一旦发生可能对主信号质量造成影响，是需要在设计中规避的自激问题。以下将对次谐波自激进行详细讨论。

2.9.3.1 “次谐波自激”产生的机理

次谐波自激原因复杂，较难分析，一直没有清晰简单的模型进行讨论。目前对次谐波自激的分析主要从两个角度进行，分别是器件角度和系统角度。

从器件角度分析，Imbornone 等人在 1997 年于 JSSCC 上发表的文章认为，PA 设计中半频震荡与 BJT 器件的 Base Charge Storage 相关 [5]。该方法源自于对 PN 结器件的非线性分析，Penfield 等人在 1962 年对变容二极管的次谐波产生做过讨论 [6]。不过由于器件侧机理复杂，这种分析方法还没有在 PA 设计中得到广泛应用。欢迎对此分析方法有研究的专家扫码留言，详细讨论。

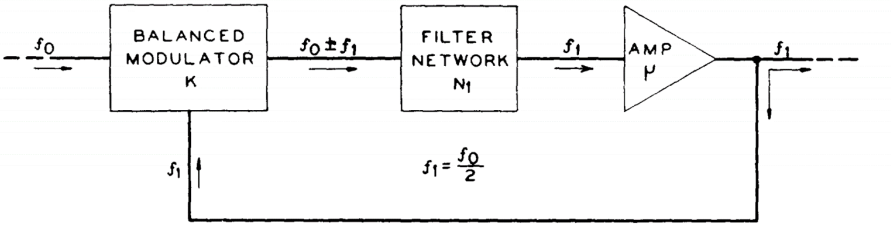


图 2.143: R. L. Miller 于 1939 年对半频信号的产生进行分析 [7]

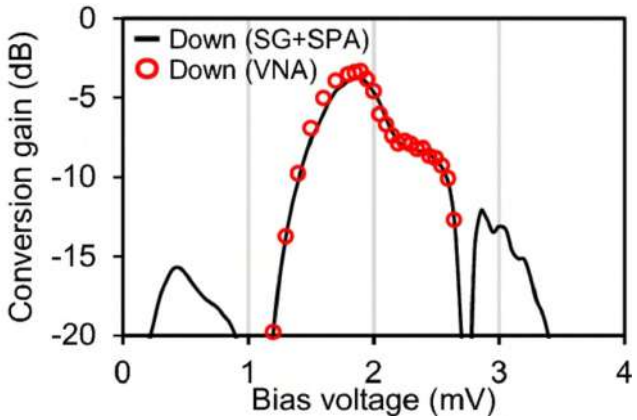


图 2.144: 一个典型的混频器混频增益与偏置间的关系 [8]

在系统角度分析方法上，R. L. Miller 等人于 1939 年发表论文，讨论再生分频器的设计 [7]，如图2.143所示。文章将主频信号与半频信号放在混频环路中进行分析，非线性器件产生混频增益，当混频增益与反馈回路共同引起的环路增益大于 1 时，将产生半频信号。

R. L. Miller 利用这种特性，设计了产生半频信号的分频器。业界称此种分频器为“再生分频器（Regenerative Divider）”，同时也叫 Miller 分频器。

R. L. Miller 应用于 Miller 分频器件的方法同样可用于分析半频震荡：当 PA 中的非线性元件 HBT 在半频混频增益过大，同时存在较大增益的半频的反馈回路时，将可能发生半频震荡。

2.9.3.2 “次谐波自激”的规避

了解次谐波自激产生的机理之后，可对次谐波自激进行规避。

方法 1：降低混频增益

次谐波自激由于次谐波混频引起，可减小混频增益，切断次谐波与主频的混频。

PA 作为非线性器件，其混频增益与偏置点相关。一个典型的混频器混频增益与偏置的关系与图2.144所示 [8]。混频增益在某个区间存在最大值，若半频震荡发生，可以适当增加或减小偏置，改变 HBT 器件的混频增益。

方法 2：增加反馈损耗

另外一个改善次谐波自激的方法是增加次谐波点的路径损耗，从而减少环路增益。

对于半频震荡，可以针对半频频点在链路中针对性的加入损耗性网络，打破半频的起震条件。典型的针对半频的损耗性网络如图2.145所示：

方法 3：检查耦合路径

检查模组内是否有其他耦合路径，如果有半频或其他次谐波频率信号的耦合路径，需要加以规避。

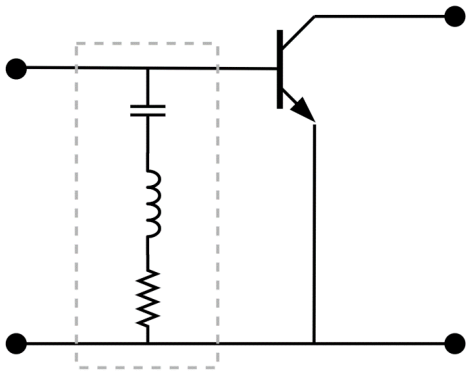


图 2.145: 谐振于半频频率的损耗网络

2.9.3.3 “次谐波自激”的说明

一般次谐波自激的发生是由于芯片内部器件的非线性引起，大多数情况下，半频或者其他次谐波频率的耦合路径也存在于芯片内部，若内部配置不改变，一般较难通过应用侧来解决。

若在应用中发生如半频震荡、1/3 频震荡的次谐波自激现象，建议联系 PA 原厂寻求解决方案。

2.9.4 实战中的“自激”问题

由于一旦确认是震荡问题，则需要花费大量的资源进行实验。而发生杂散不一定是震荡，还有可能是带外干扰信号，所以一定要准确清晰定位是否发生震荡：

- 如果看到的只是干净、稳定的少量毛刺，且不随 PA 的功率、Bias 变化而变化，则有可能是空间干扰信号或者 Transceiver 干扰信号，不是发生震荡。
- 如果看到噪底整体抬高，出现不高的鼓包而不是毛刺，也有可能不是震荡，需要检查是不是经 PA 放大的输入噪底。
- 为了减少干扰，PA 输入端需要串接滤波器，滤除输入带来的带外杂波；测试也需要在屏蔽环境进行。

在实际应用中，自激问题并不如理论分析中那样容易定位。而且自激发生后，通常是产生如文章首图中的众多毛刺，让人分不清究竟是低频震荡、高频震荡还是半频震荡。不过从经验上，总是可以找到最主要的震荡来源。一般对震荡定位时可以采用以下方法：

- 如果是众多毛刺，则从最高的 5-10 根毛刺看起；
- 以 MHz 为单位，变换 5 次以上主频的频率，记录几根毛刺的频率变化；
- 根据毛刺频率变化关系，分析哪根毛刺是震荡产生，哪根毛刺是混频产生；
- 确定震荡毛刺之后，根据其震荡频率，进行改善规避；

自激问题成因复杂，不易分析，在分析过程中一定要大胆假设，小心求证。

2.9.5 总结

自激问题与 Ruggedness 问题一样，是 PA 工程师逃不开的常见问题。

严重的自激会导致整机杂散超标，出现合规问题，并且自激还可能会影响通带内的信号质量。不可控的自激还可能导致功率及电流过大，引起 PA 烧毁。自激问题一定要在设计和应用中高度关注，加以规避。

您在设计和应用中有什么好分析和解决方法，欢在线迎留言讨论。

本文整理过程中得到众多同事及同行的帮助，在此表示感谢。

参考文献

1. https://en.wikipedia.org/wiki/Self-oscillation
2. Duncan Jr RD. Stability Conditions in VacuumTube Circuits. Physical Review. 1921 Mar 1; 17(3):302.
3. Gonzalez G. Microwave transistor amplifiers analysis and design. Prentice-Hall, Inc.; 1996 Sep 1.
4. Pozar DM. Microwave engineering. John wiley & sons; 2011 Nov 22.
5. Imbornone JF, Murphy MT, Donahue RS, Heaney E.New insight into subharmonic oscillation mode of GaAs power amplifiers undersevere output mismatch condition. IEEE Journal of Solid-state circuits. 1997Sep;32(9):1319-25.
6. Penfield P, Rafuse RP. Varactor applications.MIT Press (MA); 1962.
7. Miller RL. Fractional- frequency generators utilizing regenerative modulation. Proceedings of the IRE. 1939 Jul; 27(7): 446-57.
8. Kojima T, Uzawa Y, Shan W, Kozuki Y. On-Wafer Cryogenic Characterization Technique of an SIS-Based Frequency Up and Down Converter. Journal of Low Temperature Physics. 2020 Apr; 199(1):219-24.

在线阅读

本文摘自慧智微微信公众号，扫描以下二维码，在线阅读

2.10 射频贴片电感的选型与使用

作者：皓皓子

在射频电路中，电感是非常重要的无源器件，在射频电路中可以实现匹配、滤波、谐振及去耦等作用。下图为射频电感的四种典型应用，分别为阻抗匹配、谐振负载、串联负反馈以及滤波电路 [1]，如图2.147所示。

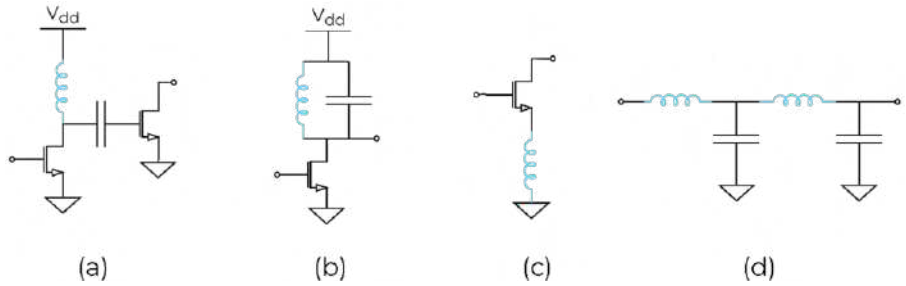


图 2.147: 电感器件的典型应用

在手机应用中，目前蜂窝移动通信设备工作的频率一般在 6GHz 以下频段，此频段常用的电感类型有基板绕线电感、键合线电感和表贴电感器件。

- **基板绕线电感**

  基板绕线电感是通过基板 Layout 走线形成电感，这类电感的缺点是占用基板面积较大，不适用于设计大电感，一般基板绕线感值范围在 5nH 以下；

- **键合线电感**

  键合线电感是利用芯片封装时的键合线来形成电感，这类电感的品质因数很高（可以到 50 甚至以上），但电感值比较小，一般在 1nH 以下，并且电感量受键合线精度影响，难以精确控制。经验上，1mm 长键合线的感值以 0.6nH 简单估计；

- **表贴电感器件**

  表贴电感器表是将电感表贴器件（Surface Mounted Devices，SMD）使用引脚直接焊到 PCB 上，这类电感使用方便，感值范围大，标准尺寸下有不同感值、Q 值及 DC 电阻值的器件可供选择，在射频系统中使用广泛。下图为表贴电感器件在射频前端模组 PAMiD 及手机 PCBA 的使用 [2]，图2.148中带有蓝色的表贴器件均为电感。

在射频电路中，电感一旦被使用，均是在对射频性能影响明显的重要位置。而表贴器件的性能参数与封装、工艺有很大关联，电感的 Q 值也有较大差异（10～40）。所以需要对表贴电感元件的特性有深入的认识，确保正常使用。

本文首先介绍射频表贴电感的分类和特征，然后结合电感的典型应用，对射频表贴电感的选型中需要考虑

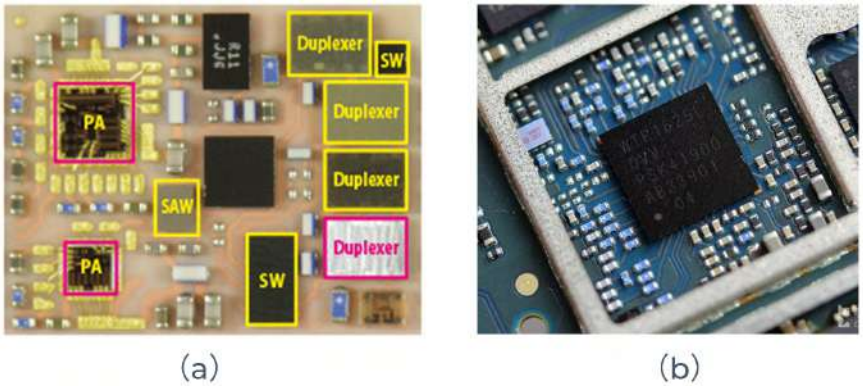


图 2.148: 表贴电感在射频前端模组 PAMiD 及手机 PCBA 中的使用

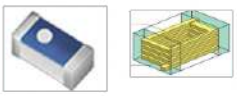
	实现方式	特征	结构图示
多层型 Multilayer Method	多层陶瓷等实现	• 精度低、Q值低 • 尺寸小、成本低	
薄膜型 Film Method	薄膜上刻蚀实现	• 精度高、Q值高 • 尺寸小、成本高	
绕线型 Wire Wound Method	铜线绕线上实现	• 精度高、Q值高 • 尺寸大、成本低	

图 2.149: 射频表贴电感三种实现类型

的主要因素进行介绍。

2.10.1 射频表贴电感的分类和特征

射频表贴电感根据工艺方法不同，可以分为多层型、薄膜型和绕线型 3 类 [3]，其特点与构成方式如图2.149所示。

从应用的角度，射频电感的选型主要从以下几个因素考虑：

1. 尺寸选择
2. 电感量
3. 自谐振频率
4. Q 值
5. 额定电流
6. DC 电阻

2.10.1.1 尺寸选择

表贴器件有标准的尺寸和封装形式，确保在使用中可以灵活选取使用。对于无源器件，常用的标称名称与尺寸间的对照关系如下表。

需要注意的是，日常使用中的“0201”、“01005”名称是以 inch 为单位的 EIA 名称（Electronic Industries Alliance 电子工业协会），而不是以公制为单位的 IEC 名称（International Electrical Commission，国际电工委员会）。图2.150为尺寸选择表格。

对于电感器件，一般情况下尺寸越小成本越高、Q 值更低。电感性能和尺寸之间是一对折中。

2.10.1.2 电感量

电感量是电感器的最主要参数，也是射频工程师选型的首要参数。工程师会通过仿真及在板验证确认最终电感值。

在表贴电感感值提供上，电感厂商一般提供 E 系列优先数系（E series of preferred numbers）下的离散数值 [4]，感值从小到大呈指数关系。使用中需要根据需求，选择接近的合理感值电感。

在选型过程中，除了关注感值外，还需重点关注电感量的容差值。电感的容差值在数据手册中有标称，常见的规格有：

1. L≤4.2nH，电感容差值 ±0.1nH 或 ±0.2nH；

EIA名称 (inch)	尺寸 (inch²)	IEC名称 (公制)	公制尺寸 (mm²)
01005	0.0157 x 0.0079	0402	0.4 x 0.2
0201	0.024 x 0.012	0603	0.6 x 0.3
0402	0.039 x 0.020	1005	1.0 x 0.5
0603	0.063 x 0.031	1608	1.6 x 0.8
0805	0.079 x 0.049	2012	2.0 x 1.25
1008	0.098 x 0.079	2520	2.5 x 2.0
1206	0.126 x 0.063	3216	3.2 x 1.6
1210	0.126 x 0.098	3225	3.2 x 2.5
1806	0.177 x 0.063	4516	4.5 x 1.6
1812	0.180 x 0.130	4532	4.5 x 3.2
2010	0.197 x 0.098	5025	5.0 x 2.5
2512	0.250 x 0.130	6332	6.4 x 3.2
2920	0.290 x 0.200	7451	7.4 x 5.1

图 2.150: 表贴器件的尺寸选择

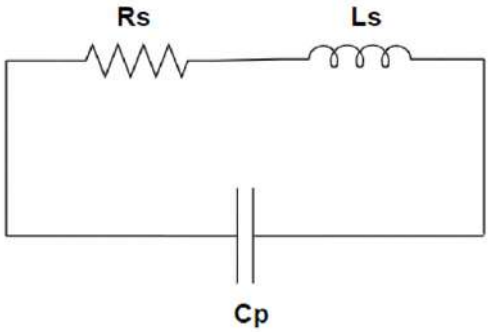


图 2.151: 电感的等效电路 [5]

2. L ＞ 4.2nH，电感容差值 ±3% 或 ±5%。

需要注意的是，不同厂商的电感器实现的技术方式不同，造成即使电感量相同，电感量的容差值可能不同，不可做直接替换，在替换后需要实测验证。

2.10.1.3 自谐振频率

在实际电感使用中，由于寄生电容的存在，寄生电容与本征电感会发生自谐振。发生自谐振的频率叫自谐振频率（Self-Resonant Frequency，SRF）。实际电感的等效电路如图2.151所示，寄生电容和电阻分别用 C_p 和 R_s 表示，电感器的自谐振频率计算公式为：

$$f \approx \frac{1}{2\pi\sqrt{L_s C_p}}$$

由于自谐振特性的存在，当工作频率低于谐振频率时，电感器件表现出电感性，阻抗随着频率的升高而增大；当工作频率高于谐振频率时，电感器件表现出电容性，阻抗随着频率的升高而减小。因此在实际应用中，应选择电感谐振频率点远高于工作频率的电感。电感的自谐振特性如图2.152所示。

按照经验值，电感的工作频率一般选择为 SRF 1/10 以下，此时的电感受寄生电容影响较小，电感值相对来说更精确。

2.10.1.4 Q 值

Q 值即电感的品质因数 Quality Factor，是电感储存功率与损耗功率之比。电感 Q 值的计算公式为：

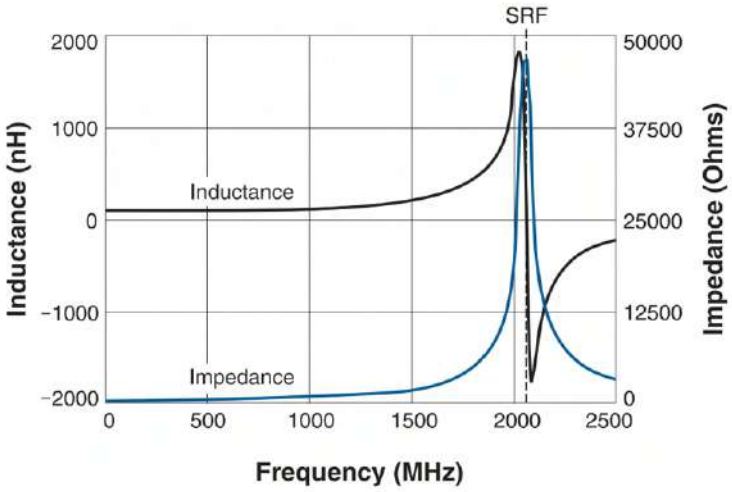


图 2.152: 电感的自谐振特性 [5]

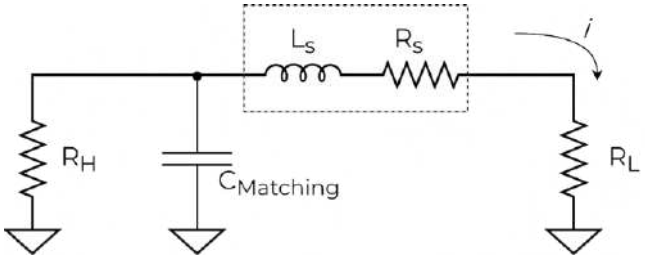


图 2.153: L 型电感电容匹配网络的阻抗变换

$$Q_L = \frac{\omega L_s}{R_s}$$

其中，ωL_s 为电感器的感抗，R_s 为电感器的实阻抗。

在射频电感使用中，最常见的应用是参与阻抗匹配，电感的 Q 值对匹配网络的损耗有直接影响。图2.153为简单的 L 型匹配网络示意图，图中串联电感 L_s 与并联电容 $C_{Matching}$ 共同完成阻抗自 R_H 到 R_L 的变换。

在匹配网络中，一般电容的 Q 值较高（>200），而电感的 Q 值较低（约 30），所以在匹配网络损耗计算中，主要考虑电感的影响。对于上图匹配网络，计算传输中的损耗如下：

如 [6]，定义网络 Q 值 $Q_{Network}$ 为：

$$Q_{Network} = \sqrt{\frac{R_H}{R_L} - 1} = \frac{\omega L_s}{R_L}$$

能量的有效传输效率为：

$$\eta = \frac{i^2 R_L}{i^2(R_s + R_L)} = \frac{\frac{1}{Q_{Network}}}{\frac{1}{Q_{Networks}} + \frac{1}{Q_L}} = \frac{Q_L}{Q_{Network} + Q_L} = 1 - \frac{Q_{Network}}{Q_{Network} + Q_L} \approx 1 - \frac{Q_{Network}}{Q_L}$$

即网络的损耗为：

$$Loss = \frac{Q_{Network}}{Q_L}$$

根据上式，匹配网络损耗与网络 Q 成正比，与器件 Q 成反比。即在网络变换比确定（网络 Q 值一定）的情况下，器件 Q 值越低，网络损耗越大。

根据上式，若将 50 Ohm 阻抗匹配至 5 Ohm，不同 Q 值电感器件带来的网络损耗如图2.154。可以看到，随着电感 Q 值的降低，匹配网络的损耗逐渐增加。

Q_L	Loss (%)	Loss (dB)
100	3%	-0.13
50	6%	-0.27
40	8%	-0.34
30	10%	-0.46
20	15%	-0.71
10	30%	-1.55
5	60%	-3.98

图 2.154: 50 Ohm 到 5 Ohm 变换时不同 Q 值电感对应损耗的变化

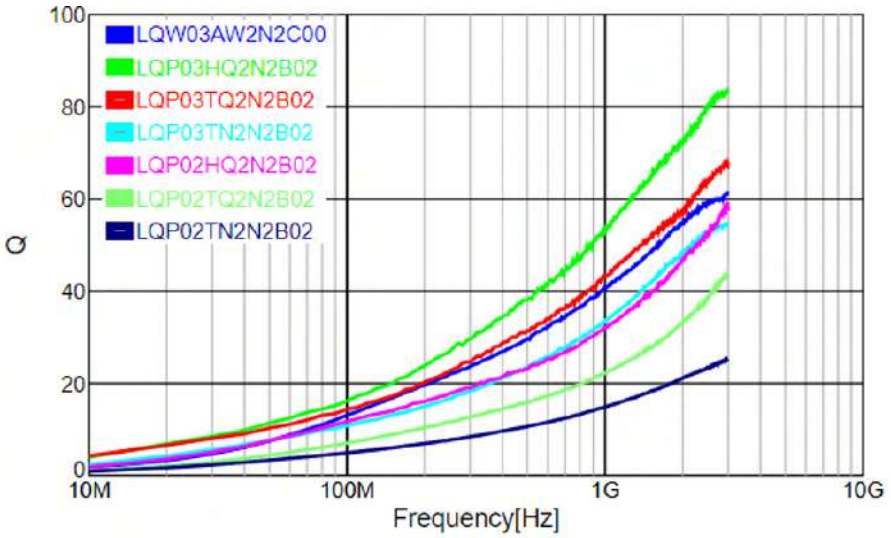


图 2.155: 不同类型电感 Q 值随频率变化的关系

2.2nH 电感 Q值@1GHz	0201封装		01005封装	
	型号	Q值	型号	Q值
TN系列	LQP03TN2N2B02	33	LQP02TN2N2B02	15
TQ系列	LQP03TQ2N2B02	43	LQP02TQ2N2B02	22
High Q系列	LQP03HQ2N2B02	53	LQP02HQ2N2B02	31
绕线型	LQW03AW2N2C00	41		

图 2.156: 不同系列电感的 Q 值对比

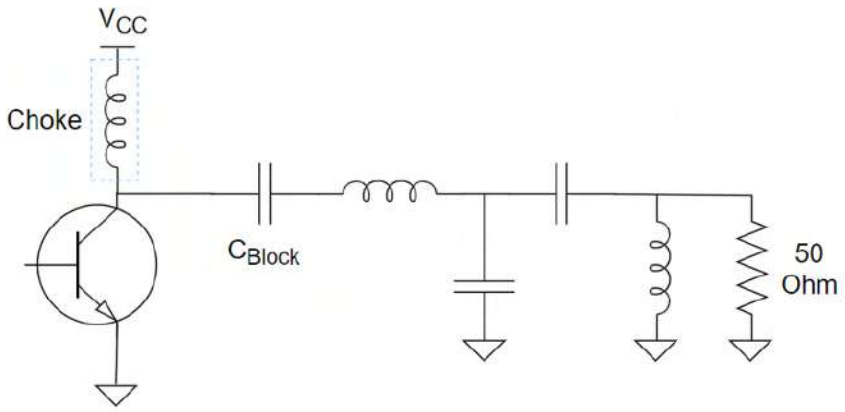


图 2.157: 5G PA 的输出匹配网络

2.2nH 电感 Q值@1GHz	0201封装		01005封装	
	型号	Q值	型号	Q值
TN系列	LQP03TN2N2B02	33	LQP02TN2N2B02	15
TQ系列	LQP03TQ2N2B02	43	LQP02TQ2N2B02	22
High Q系列	LQP03HQ2N2B02	53	LQP02HQ2N2B02	31
绕线型	LQW03AW2N2C00	41		

图 2.158: 村田不同系列 4.7nH 电感的额定电流与 DCR 值对比

电感 Q 值的大小取决于元件的制作工艺、制作材料等，电感的寄生电阻越大，Q 值越小。图2.155及图2.156为村田 0201 和 01005 系列电感的 Q 值对比 [8]。

从图中可以看出：

1. 封装越小，Q 值越低；
2. 随着频率变高，电感的 Q 值越高；
3. High Q 系列比 TN/TQ 系列电感的 Q 值高；
4. 绕线型电感的 Q 值较 TN/TQ 电感有一定优势，但随着器件厂家 High Q 技术的提升，绕线结构上的优势已经被薄膜电感替代。

一般 High Q 电感价格也更高，所以电感的 Q 值和价格之间也是一对折中。

2.10.1.5 额定电流与 DC 电阻

器件厂家数据手册中标注的额定电流一般指在室温下通电流，逐步提升电流至产品表面温度上升 20℃ 的电流，超出该电流值使用时可能会导致元件破损及组件故障。

在 PA 等有源射频电路中，射频电感可用于隔离交流（Choke），将射频信号与直流偏置和直流电源隔离，此时 Choke 电感将通过较大电流，对此电感的通流能力要求较高。通常的 PA 设计中该电感的电感值较小（约为 nH 量级），用基板绕线电感是首选，若布局布线受限，表贴电感的选取需要注意电感的额定电流是否满足设计要求。图2.157为 5G PA 的输出匹配网络示意图。

电感器件额定电流的大小与其 DCR（DC Resistance）直流电阻呈反比关系，DCR 越高，额定电流越小。DCR 指在无交流信号下测得的电阻 [10]，主要由内电极的电阻决定，电阻的计算公式：

$$R = \frac{\rho l}{S}$$

其中 ρ 为电阻材料的电阻率，l 为电阻的长度，S 为电阻的截面积。

图2.158为村田不同系列 4.7nH 电感的额定电流与 DCR 值对比：

可以看到：

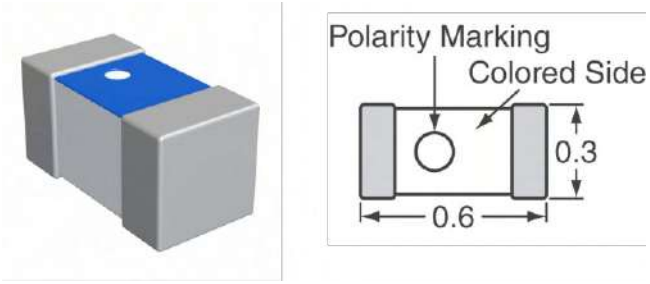


图 2.159: 村田电感数据手册中对于电感方向性的标识

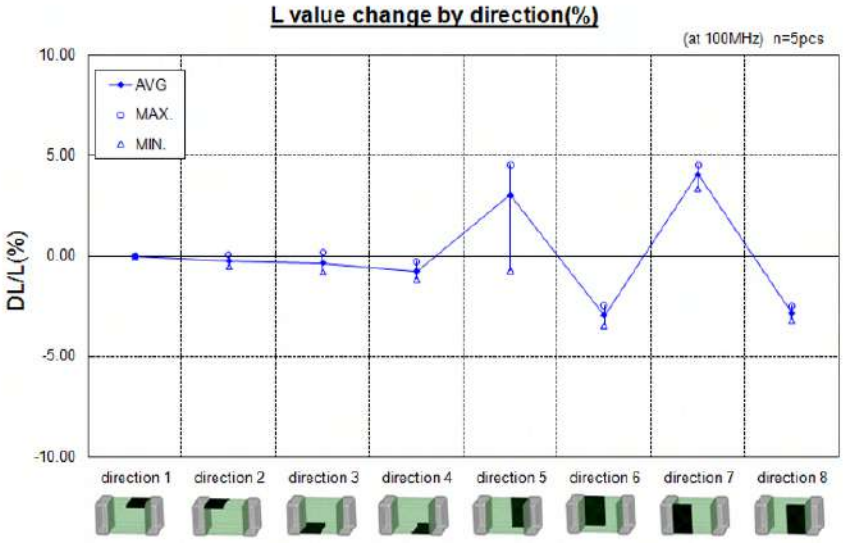


图 2.160: 贴片电感不同方向放置时感值的变化

- DCR 越大，额定电流越小；
- High Q 电感额定电流比 TN/TQ 电感大；
- 大尺寸电感有更大额定电流及更小 DCR。

实际选型中，额定电流应是电路中最大输出电流的 1.3 倍以上，需要留有一定的余量降额使用。

2.10.2 射频表贴电感的极性

2.10.2.1 表贴电感在使用中必须要考虑极性（Polarity）

与表贴电容不同，大部分表贴电感带有可辨别方向性的标记。图2.159为村田电感数据手册中对贴片电感方向的标识。

在使用中，必须要严格按照电感的方向性进行贴装使用，电感方向性的不同会造成感值的不同。图2.160为村田将电感按 8 个不同方向进行放置，测试电感值的变化。可以看到，在不同方向放置时，感值存在 5% 的变化[9]。

产生感值变化的原因主要有两个

- **线圈磁场方向的改变**

  电感一般通过绕线线圈实现，电感依靠线圈产生的磁场来储存能量。线圈放置方向的不同，造成线圈产生的磁场发生变化，进而使感值发生变化。图2.161为不同方向放置电感时磁场变化示意图。

  另外需要注意的是，电感产生的磁场的变化不仅会影响到自身电感的变化，还会影响到周围器件的电感变化。所以在使用中要注意电感器件与其他器件之间的距离，并对电感的摆放方向做严格定义。

- **电感的非对称特性**

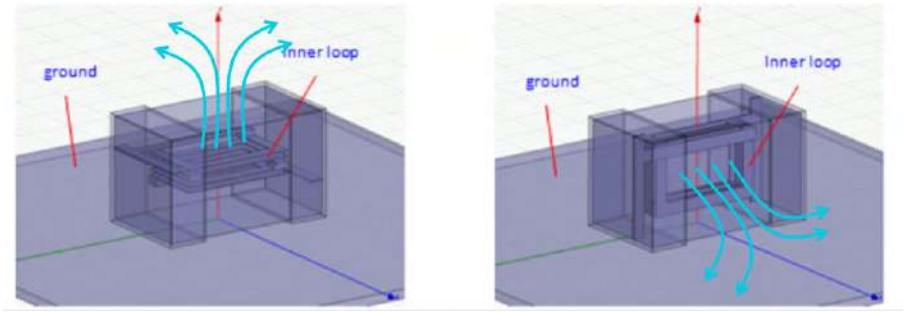


图 2.161: 不同方向放置电感时磁场的变化

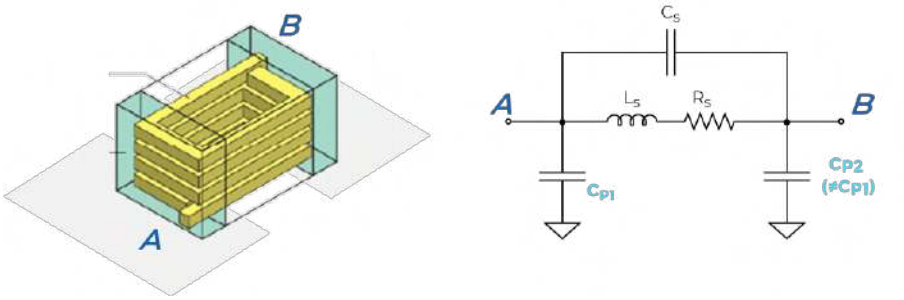


图 2.162: 表贴电感的非对称特性

由于表贴电感采用绕线方式实现，在物理上不可能完全对称。这种不对称特性造成了电感器件在一些寄生参数上的左右不互易。如图2.162所示，电感在 A 点与 B 点有不同的到地寄生电容，造成电感器件不能左右互换，在使用中必须规定极性。

2.10.3 总结

电感是射频电路中经常使用的无源器件，小小的电感元件，如果使用不当将会对射频性能造成极大的影响。以上就是射频贴片电感在选型与使用中的注意要点。

您在实际使用中觉得电感器件有哪些需要重点考虑的要点？欢迎扫码在线留言讨论。

参考文献

1. 池保勇余志平石秉学著.CMOS 射频集成电路分析与设计. 清华大学出版社
2. https://corporate.murata.com/ more_murata /techmag /metamorphosis20/ productsmarket /pamid
3. https://article.murata.com/zh-cn/article/ types-and-characteristics -of-radio-frequency-inductors
4. https://handwiki.org/wiki/E_series_of_preferred_numbers
5. Key Parameters for Selecting RF Inductors, Coilcraft.
6. Razavi, B. . ”RF Microelectronics, 2ndEdition.” (2012).
7. 朱晓明等著. 硬件十万个为什么-无源器件. 西安: 北京大学出版社
8. https://ds2.murata.co.jp/simsurfing
9. https://article.murata.com/en-us/article/basic-facts-about-inductors-lesson-6
10. https://www.sunlordinc.com/category.aspx?NodeID=239
11. Yue C.P.S.S Wong. Physical modeling of spiralinductors on silicon. IEEE Tans. Electron devices, 2000,47(3):560-567

在线阅读

本文摘自慧智微微信公众号，扫描以下二维码，在线阅读

2.11 射频开关，简约不简单

作者：彭洋洋 博士

在你拿着手机打电话或者上网时，你可能不会想到，在手机里有种小小的射频器件，在以每秒钟近千次的速度快速切换，帮助你享受到高速、低延迟、高清晰度的通信服务。这个小器件就是射频开关（RF Switch）。

射频开关是一种能够让手机在不同射频信号通路之间切换的器件，它就好像是信号的“桥梁”，不断将信号通路连接或切换。射频开关应用的范围也很广泛，在 2G/3G/4G/5G 蜂窝通信系统、Wi-Fi、蓝牙、GPS 等系统中，均是不可或缺的器件。

和家里电灯开关一样，对于射频开关来说，只有“打开”、“关闭”两种状态，功能非常简约。但这种简约的功能背后，却隐藏着复杂而精妙的设计思路，以及不简单的技术挑战。

本文将从目前手机中最常用的开关：SOI 开关出发，讨论“简约而不简单”的射频开关设计。

2.11.1 射频开关：功能简约

2.11.1.1 功能简约

正如其名称一样，开关的功能就是“开”和“关”。射频开关也不例外，射频开关是工作在射频频段的开关，其功能就是控制射频信号的“通”与“断”。

目前手机中应用最广泛的是半导体器件开关，如 RF-SOI 开关，pHEMT 开关等。这些半导体开关的功能与普通的电气开关相同，符号表示也一致。如图2.164所示。

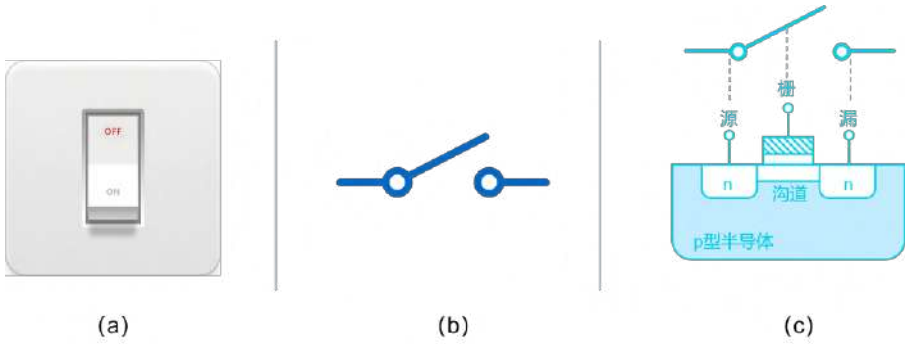


图 2.164: （a）电气开关（b）开关符号（c）半导体开关

2.11.1.2 应用简约

在射频系统中，射频开关扮演着重要的角色。射频开关的主要功能有：

- 频段选择：使信号在多个不同的射频通路间切换
- 收发切换：在 TDD（时分双工）系统中，完成接收与发射的切换
- 天线切换：在多天线系统中，使信号在不同天线间切换

在《除了调匹配，射频人还要掌握系统知识》中，可以看到，射频开关器件大量分布在射频前端系统中，在天线切换层、频段开关层及子路径实现层中需要用到多个射频开关器件。如图2.165所示。

2.11.1.3 评价简约

开关的评价也非常简约：开启的时候能量尽可能多的传过去；关闭的时候能量尽可能少的漏过来。

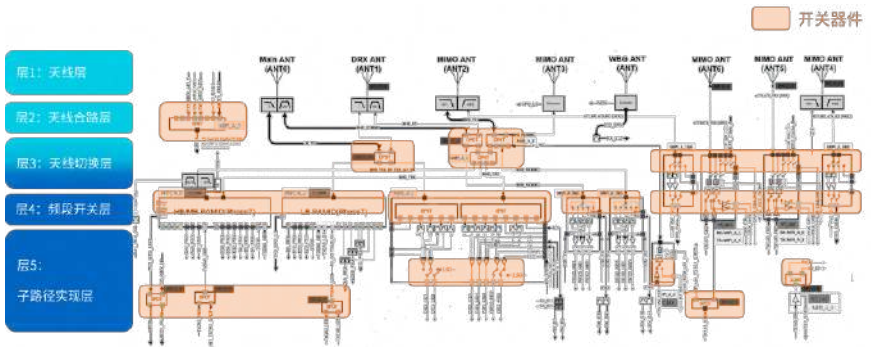


图 2.165: 射频开关器件在射频前端系统中的分布

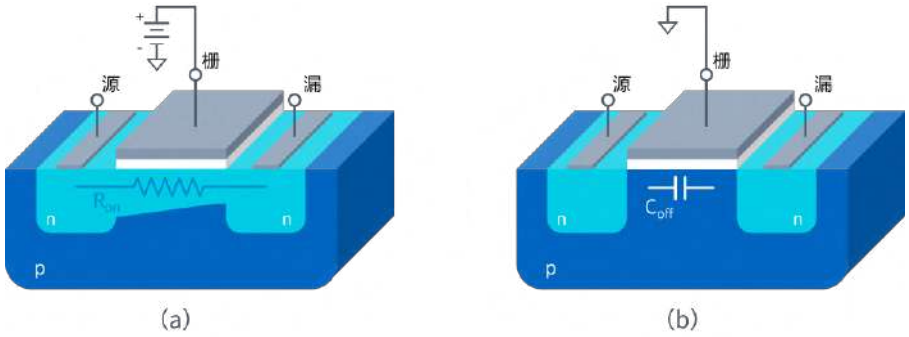


图 2.166: （a）on 状态的 Ron；（b）off 状态的 Coff

开关有两种工作状态，分别是“开启 on”和“关闭 off”。对于半导体晶体管制作的开关，晶体管在 on 状态可近似等效为一个电阻，这个电阻的阻值就被定义为 Ron；晶体管在 off 状态时可近似等效为一个电容，这个电容就被定义成 Coff。这就是衡量开关本征性能的两个最重要参数：Ron 与 Coff 参数的来源。如图2.166所示。

在 on 状态下，由于 Ron 是串接在射频通路中，所以 Ron 就决定了开关损耗的大小；而在 off 状态下，Coff 的存在会造成信号的泄露，所以 Coff 决定了开关隔离的大小。Ron 与 Coff 两个参数都是越小越好。

在晶体管作为开关的设计中，还有一个非常有意思的特性，即 Ron 与 Coff 的乘积是定值。理解起来也比较直观：当设计中想要用多个晶体管并联来实现低的串联电阻时，Ron 会呈倍数的减小，但因为是晶体管的并联，这时 Coff 却会呈倍数的增加。二者的乘积始终不变。

这一特性也使得 Ron*Coff 成为衡量开关特性的简约衡量指标，评价一个工艺作为开关使用的优劣，不论取什么尺寸的晶体管，只需要将其 Ron 与 Coff 相乘，就可以得到其特性参数。如图2.167为文章 [1] 中不同 RF-SOI 工艺的对比，可以看到，不同工艺的 Ron 与 Coff 乘积会有不同，大致在 115fs 至 165fs 之间，但相同工艺下的多种器件得到的乘积基本相同。

由于 Ron 与 Coff 乘积决定了开关插损与泄露能量的大小，所以在设计低插损、高隔离开关时，应尽量选择低 Ron 与 Coff 乘积的半导体工艺。

2.11.2 射频开关：做好不简单

虽然对于一个射频来说，功能、应用评价都非常简约、直观，但想要将开关做好却并不容易。主要原因在于射频系统越来越复杂，对于“架桥铺路”功能的射频开关也提出了越来越高的要求：

- 发射通路上的开关必须要承载大功率的通过
- 切换速度要足够的快
- 复杂系统下，开关的“刀”和“掷”数目激增

这就像是将原来只负责通过 40 公里时速小车的一分二岔路口，变成大型立交桥，并且还需要保证重型大车以 120 公里时速的快速通过。这对“立交桥”设计提出了极高的要求。如图2.168所示。

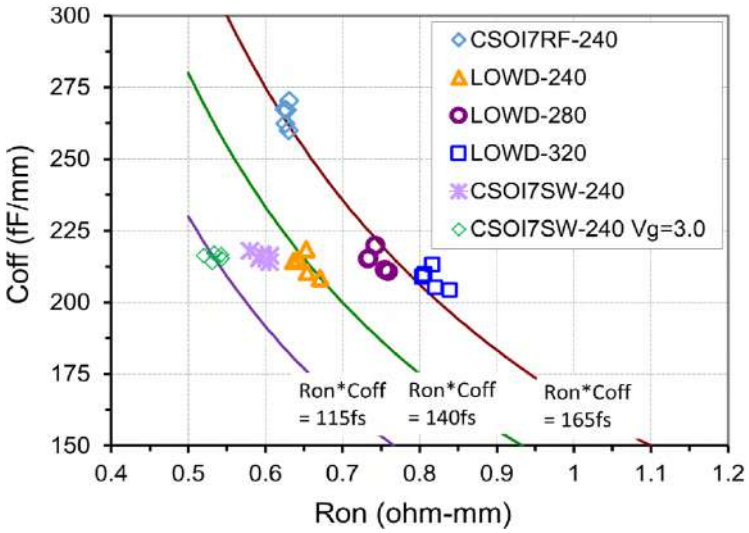


图 2.167: 不同工艺的 Ron 与 Coff

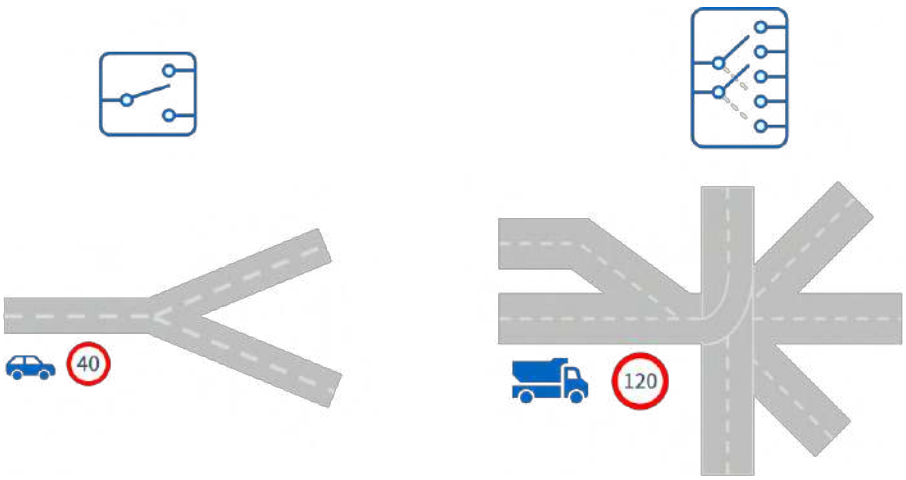


图 2.168: （a）简单系统中开关要求（b）复杂系统中开关要求

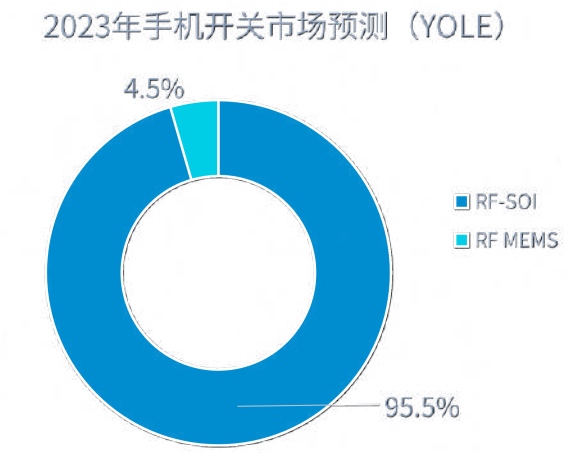


图 2.169: 手机开关所使用的工艺

2.11.2.1 大功率的处理

根据 YOLE 的预测 [2]，至 2023 年，全球手机市场所用到的开关主要集中于 SOI 与 MEMS 两种工艺，并且 SOI 会是绝对的主流工艺。如图2.169。

除了 MEMS 之外，可以用于手机的工艺还有很多，比如半导体工艺中的 GaAs pHEMT、GaAs FET 等，这些工艺在微波毫米波、卫星通信等有广泛应用。另外，普通 CMOS 工艺也可以用于低频、高插损的开关设计。SOI 之所以能够从这些工艺中竞争胜出，还是因为其作为开关使用时的重要优势：

- 低插入损耗：SOI 开关具有低阻抗和低电容，可以减少 RF 路径中的信号衰减和功耗；
- 宽带宽：SOI 开关可以在很宽的频率范围内工作，从 9 kHz 到 44 GHz 甚至更高，这使得它们能够支持多种标准和频段
- CMOS 兼容的正极控制接口：SOI 开关可以与 CMOS 逻辑电路轻松集成，并由正电压信号控制，这简化了设计并降低了成本；
- 坚固的 ESD 保护：SOI 开关在所有引脚上设计具有高 ESD 耐受性电路，这增强了 RF 系统的可靠性和耐用性。

与 SOI 相比，GaAs pHEMT 开关虽然具有良好的线性度和隔离度，以及低的通态电阻和截止电容。但它们需要负的栅极电压、有限的集成能力和低的 ESD 保护，并且 GaAs pHEMT 工艺价格高，不利于低成本规模应用。这些都使得 SOI 可以实现射频开关后，迅速将 GaAs pHEMT 等工艺取代。

但 SOI 进军开关市场的路径也并不顺利, SOI 需要解决的首要问题就是功率问题。

对于手机来说，发射通路中的功率一般在 1W 量级附近，对于 50 Ohm 系统来说，射频摆幅会到 10V 以上，考虑到负载变化带来的影响，这个电压甚至可能会超过 20V。而 SOI 中的 MOSFET 器件击穿电压只有 2V 左右，功率耐受在最初的时候成为 SOI 在手机开关应用中的最大问题。

但这并不是技术不可以解决的问题，一些 SOI 的先驱厂商的工程师做了聪明的尝试。2015 年，Peregrine 公司的工程师 Dylan Kelly 等人成功利用 SOS（Silicon on Sapphire，蓝宝石上硅）工艺，采用叠管技术，支持高的耐压摆幅，设计制造出可以满足手机 GSM 应用的 6T 射频开关，并且性能与 GaAs pHEMT 媲美 [3]，如图2.170所示。由此拉开了 SOI 设计手机射频开关的序幕。

采用叠管之后，虽然串联使用时，on 状态的电阻 Ron 会倍数级的增加，但 off 态的 Coff 寄生电容会同倍数的减小。所以可以采用增大晶体管面积的方式，使 Ron 减小至原来值，这里 Coff 也增加至原来值。Ron 与 Coff 的乘积依然保持不变，开关的损耗与隔离特性并没有受到影响。耐压增加又不影响射频性能，叠管设计这一方法瞬间在射频开关应用中普及开来。

图2.171为 Peregrine 公司 2005 年设计的 GSM 手机开关 [3]。开关采用 8 个叠管的设计，共有 6 个支路。支路 1 和 2 对开关插损有高要求，所以选了较大的晶体管尺寸，而 3/4/5/6 之路插损和尺寸之间做了折衷，晶体管

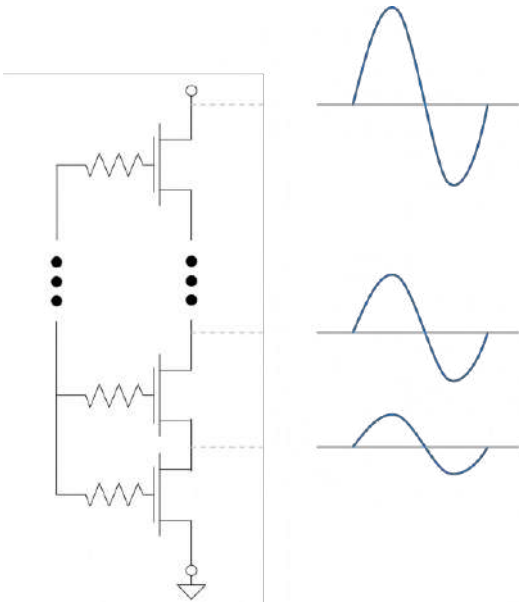


图 2.170: SOI 工艺中，依靠叠管来支撑高的电压摆幅

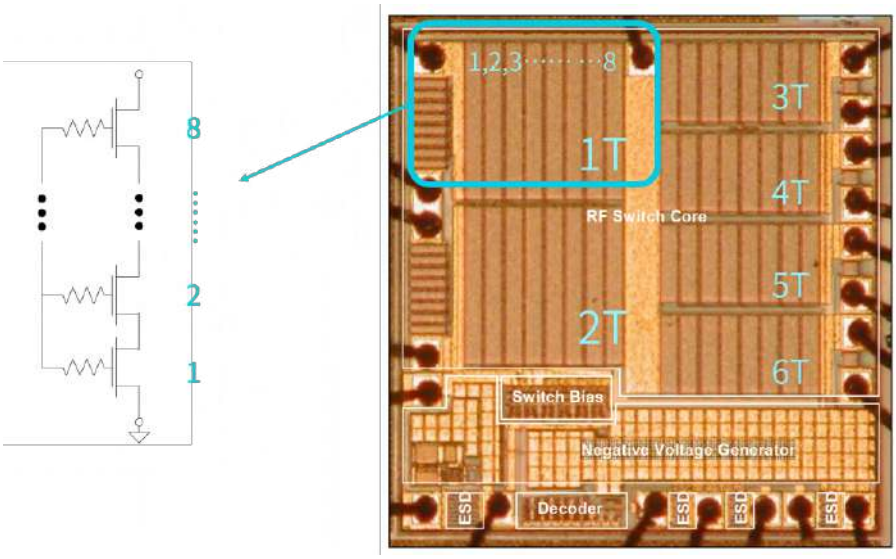


图 2.171: Peregrine 2005 年设计的 SOS CMOS 开关

尺寸比 1 和 2 之路小了一半，所以 Ron 电阻也会增大一倍。

需要说明的是，以上分析均是假定开关几个叠管之间可以完美的将功率进行均分，但在实际设计中，想要控制电压均匀分布也不是易事，需要仔细设计偏置电路，并将寄生效应完整考虑进来。否则虽然设计上进行了叠管设计，但很有可能对于应对大功率无济于事。

如果在大功率方面处理不好，很好可能造成开关烧毁，造成不可恢复的可靠性问题。所以开关的大功率问题，是开关设计不简单的首要注意问题。

2.11.2.2 切换速度的处理

为了实现更优的射频体验，5G 引入了很多新特性。比如更灵活的子载波配置、天线轮发系统等，这些功能都对开关的切换时间有了进一步要求。在 4G 时代，开关切换时间的设计目标一般在 2μs 左右，但在 5G 系统中，这个设计目标已经降低到 0.5μs 以下。如图2.172所示。

快速切换给开关设计提出了很高的要求。在开关设计中，切换速度的提升主要通过在偏置电路中的优化实现。比如可以通过以下方式来提升开关切换速度：

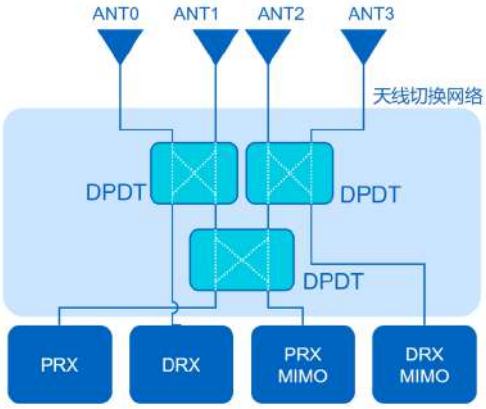


图 2.172: 5G 系统中复杂而快速的切换网络

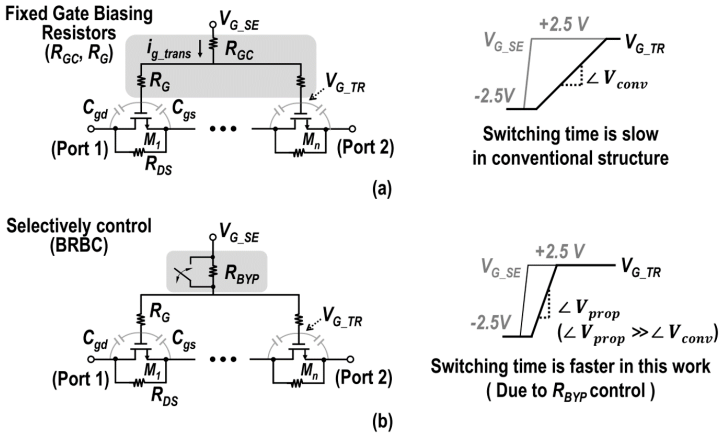


图 2.173: 文献提出的快速开关切换电路（a）传统方式（b）切换偏置电阻方式

- 优化控制逻辑拓扑，简化控制路径
- 提升控制电路驱动能力，快速实现控制信号的切换
- 优化偏置电路，缩短控制信号充放电时间

为了实现更高的切换速度，一些创新的方法也被引入进来。比如文献中提出，可以在开关切换的过程中，可以将用于隔离射频与偏置的偏置电阻暂时切除，以达到快速控制切换的目的。在切换完成后，再将偏置电阻补充回来，保证射频性能不受影响，如图2.173所示。采用这种方法，文章完成了 0.35μs 切换时间的 5G 开关设计。

偏置电路的优化需要结合射频性能进行，速度的提升需要建立在射频性能尽量不受影响的基础上。开关速度的处理是开关设计中另外一个不简单的问题。

2.11.2.3 复杂架构的处理

在从“岔路口”到“立交桥”的演进过程中，开关的拓扑结构越来越复杂，由此也带来一系列的设计问题。比如其他支路的寄生处理、多支路之间的耦合、多通道同时开启的相互影响等。这些问题都给 5G 手机开关带来挑战。

开关一般用“刀”和“掷”来定义架构。刀的英文是的英文名称是 Pole，简称 P，指开关中的活动刀轴；掷的英文名称是 Throw，简称 T，指开关的活动刀头可以通向的触点数目。

比如，1P2T 开关，指的就是开关有 1 个活动刀轴，可以通向 2 个通路；而 2P6T，指的就是开关有 2 个活动刀轴，可以通向 6 个通路。在日常开关使用中，“1”也被称为 Single，“2”也被称为 Double，以 1P2T 与 2P6T 开关为例子，日常也被称为 SPDT，与 DP6T，如图2.174所示。

在复杂射频系统中，开关拓扑主要会被扩展为多 T、多 P 与多通三种类型，如图2.175所示。

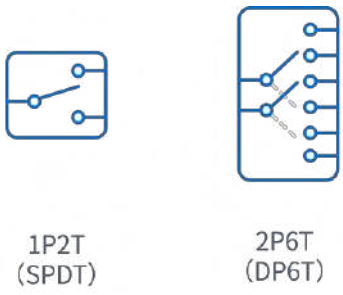


图 2.174: SP2T 与 DP6T 开关

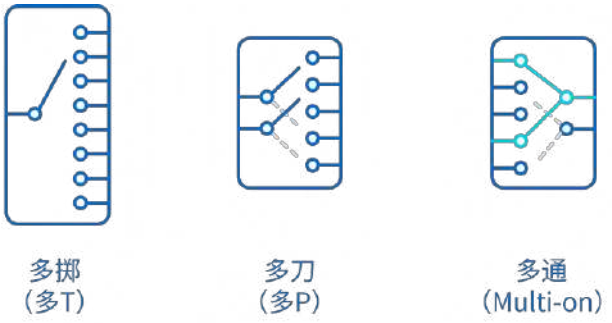


图 2.175: 复杂的开关架构

多 T 处理

多 T 指的就是开关有非常多输出口。在使用中，同一时刻只会有一个开关开启，但其他关闭状态的开关都是寄生的电容负载。

在多 T 开关的设计中，为了减少过多关断开关对导通支路电容寄生的影响，同时增加隔离度，可以采用将多 T 开关分组的方式进行设计。文献 [5] 中 16T 开关分为四组，分别为 GSM、LTE1、LTE2 与 Rx，在减小互相之间影响的同时，可以针对性的对不同需求进行设计，如图2.176所示。

多 P 处理

多 P 指的是开关有多个活动刀轴。和机械开关多放置几个活动刀头就可以解决不同，半导体开关必须要通过通路矩阵的方式依靠拓扑来实现。

以 DPDT 开关为例，文献 [6] 中给出了其中一种实现方式。可以看出，与 SPDT 开关相比，其通路数目增加一倍。如果开关的 P 数与 T 数进一步增加，开关设计复杂度也会指数级增加，如图2.177所示。

多通处理

多通（Multi-on）是指开关的两个或者两个以上支路可以同时打开。

多通需求的原因是因为手机载波聚合（Carrier Aggregation，CA）以及 4G/5G 双连接（LTE/NR-Dual Connection，EN-DC）的需求。这些需求中需要两个射频通道同时工作，所以就需要开关支持多通功能。

在 MTK 所发布的射频前端规划中，也提出过对多通开关的需求。例如在 MTK 提出的支持 CA 的架构中，就依赖于天线开关的两通道同时打开 [7]，如图2.178。

开关的双通道打开也对开关提出了新的需求，首先需要开关可以处理好两个频段之间的干扰问题，同时还需要使两频段间的工作状态尽量少的相互影响。

以上框图为互相独立的两个射频通路间同时打开，在 5G 的系统架构中，有时还需要将同一个开关口同时连接到不同的输出口，这对开关设计提出了新的要求。

以上开关的复杂架构演进，也让开关设计不再简单。

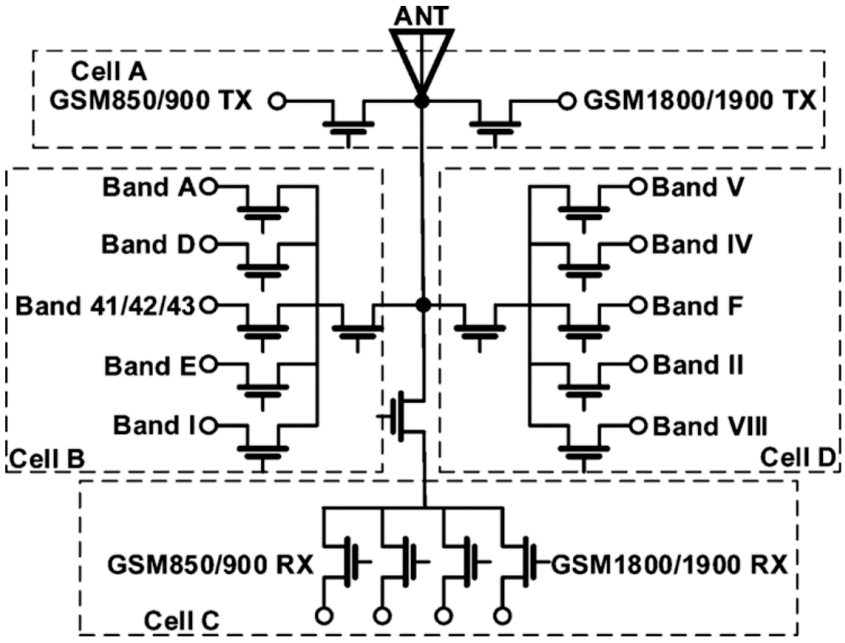


图 2.176: SP16T 开关设计

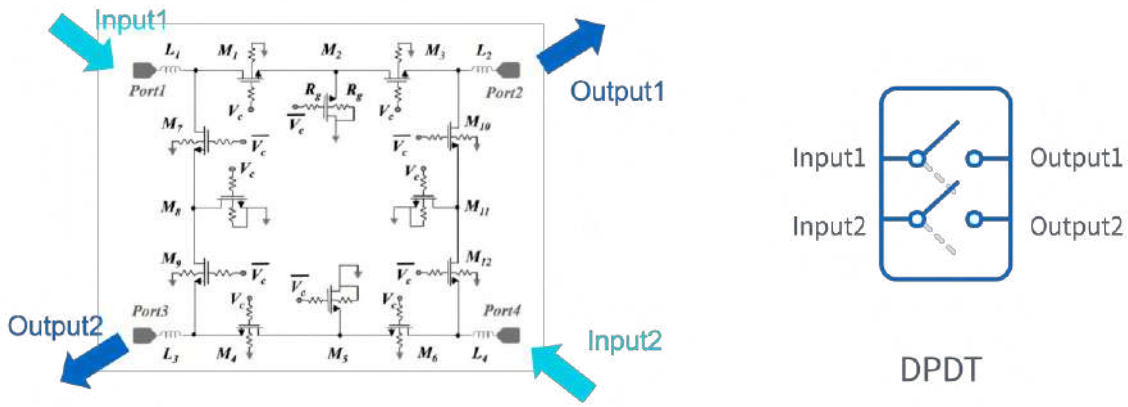


图 2.177: DPDT 开关的实现

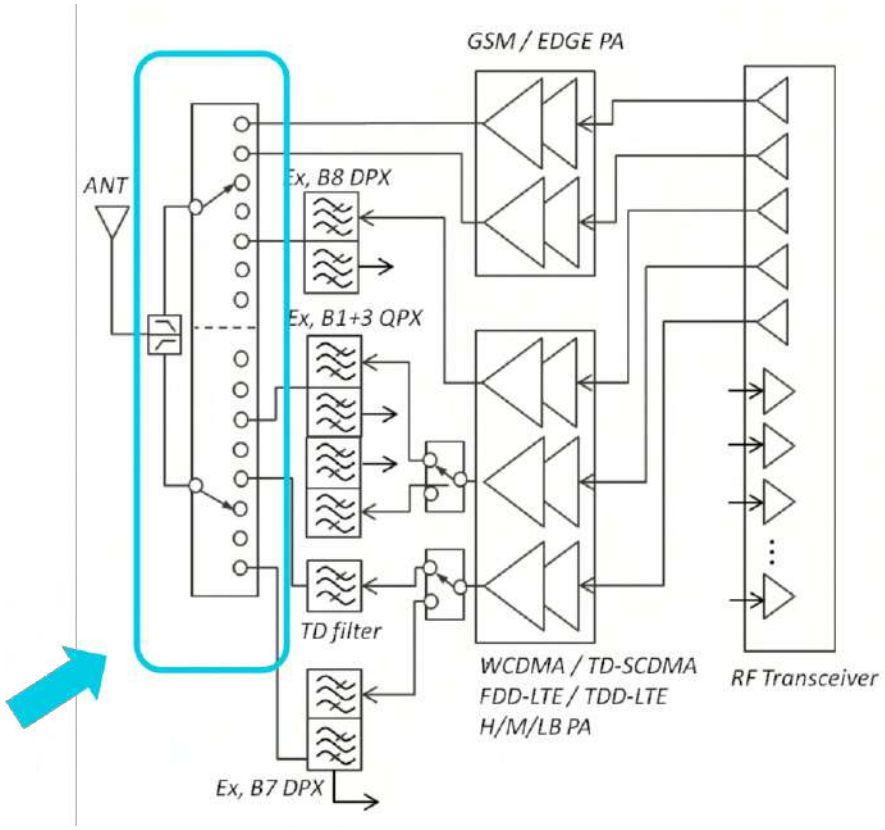


图 2.178: 射频前端中的多通开关

2.11.3 总结

在射频前端的四大件中，和 PA、LNA、滤波器比起来，射频开关看起来是最简约、最常见的器件，也经常被人们认为是最简单的射频器件。

但射频开关也是应用场景最为复杂的器件，不管是在复杂射频系统构建、还是在射频通路切换中，都可以看到射频开关的身影。不同系统需求下，对开关的要求也是千差万别。要应对好不同场景下的需求，开关设计并不简单。

随着 5G 的到来，射频前端系统越来越复杂，射频开关也在 5G 系统构建中大显身手。随着对射频性能要求的提升，以及未来 6G 的到来，开关应用会更加广泛。

在开关设计中，了解开关的应用场景，有的放矢，才可以选择合适的开关设计。欢迎扫描文章末尾二维码留言讨论，设计好“简约不简单”的射频开关。

参考文献

1. M. Jaffe et al. ，“Improvements in SOI technology for RF switches,”2015 IEEE 15th Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems (SiRF), San Diego, CA, USA, 2015, pp. 1-3
2. YOLE，5G’s impact on RF front-end industry, 2017
3. D. Kelly et al.,“The state-of-the-art of silicon-on-sapphire CMOS RF switches”, IEEE Compound Semiconductor Integrated Circuit Symposium, 2005. CSIC ’05.
4. W. Seo et al.，“High-Powered RF SOI Switch With Fast Switching Time for TDD Mobile Applications”, IEEE Access, Volume: 11, 2023
5. Jie. Cui et al.,“Monolithic Single-Pole Sixteen-Throw T/R Switch for Next-Generation Front-End Module”, IEEE

Microwave and Wireless Components Letters, Vol. 24, No. 5, 2014

6. Bo. Yu et al.,“DC 30-GHz DPDT Switch Matrix Design in High Resistivity Trap-Rich SOI”, IEEE Transactions on Electron Devices, 2017
7. MediaTek,“LTE RF Front-End Architecture for Mass Market”, MediaTek White Paper, 2014

在线阅读

本文摘自慧智微微信公众号，扫描以下二维码，在线阅读



2.12 高可靠芯片的设计与评估

作者：彭洋洋 博士

作为芯片行业从业者，估计很多人最怕听到的就是一些话是：芯片“可靠性”出问题了。对于 5G 射频芯片，评估基本性能都已经非常复杂，出了“可靠性”问题更是会让人胆战心惊。

如果“可靠性”评估不彻底，器件在终端使用中出现失效，则会影响产品功能，严重时可能造成不可挽回的惨痛后果。

但如果对“可靠性”的评估不科学，层层加码、过度苛责，又会严重影响项目进度，使成本、时间急剧上升。

为了科学的解释和评估电子元器件的可靠性，伴随着 1947 年半导体晶体管的发明和大规模应用，逐渐兴起了一个学科：**可靠性工程**。

可靠性工程是提高元器件在整个生命周期内可靠性的一门工程技术学科，涉及设计、分析、试验等各个产品开发过程。可靠性工程于 1950 年前后在美国兴起，伴随着半导体技术的进步，已经发展了 70 年。本文借可靠性工程的相关理念和方法，和大家就芯片的可靠性设计与评估理念做一个讨论。

2.12.1 可靠性工程中的几个重要理念

2.12.1.1 失效时间：性能退化是基本物理规律

正如人有生老病死，自然界中物体的特性也会逐渐退化。这也符合热力学第二定律，熵增定律的表现：孤立不可逆系统的熵（无序程度），会随着时间增加而增加。

既然每个期间都会性能退化，那器件可以使用的时间究竟是多少呢？于是，失效时间（Time to Failure，TF）的概念就被引入了进来。

在可靠性工程中，认为器件参数 S 是与时间 t 相关的函数。性能参数 S 与初始性能 S_0、时间 t 的关系，可以用级数的方式表示出来（m 为幂律指数）：

$$S = S_0 \cdot (1 \pm A_0 t^m)$$

通过以上公式可以看到，参数 S 会随时间变化而变化，变大与变小受正负号 ± 控制，变化的幅度受 A_0 控制。

当器件参数 S 随时间变化而减小时，假设可以在不同时间将器件参数均采集出来，就可以得到以下器件参数 S 随时间变化的曲线。如图2.180所示。

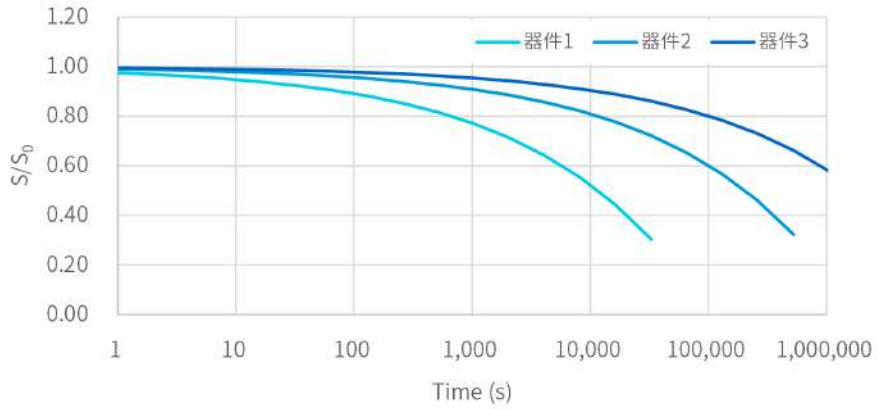


图 2.180: 器件参数 S 随时间变化而减小

得到以上参数后，如果将 S 变化原有值的 80% 作为临界值，就可以得到不同器件失效时对应的时间。这个时间就定位为失效时间（Time to Failure，TF）。如图2.181所示。

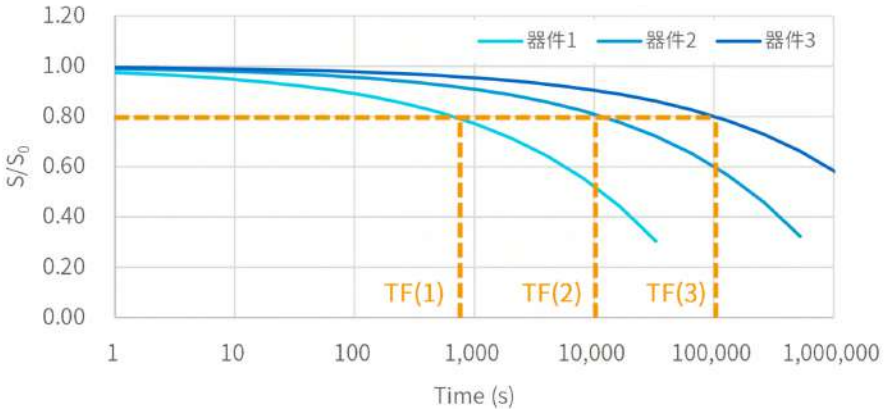


图 2.181: 器件失效时间（TF）的定义

2.12.1.2 加速退化：可控时间内测试出器件寿命

如前节所述，器件性能的退化是随时间演进而发生的。对于消费类产品来说，一般生命周期长达数年。在产品推出来之前，先做完整个生命周期的可靠性验证是不现实的，更不用说一些工业级、车规级的芯片需要 10 年以上的可靠性寿命。

于是，加速退化的概念就被引入进来。

加速退化（Accelerated Degradation）是指通过提高器件测试条件的应力、或者升高湿度，来加快器件退化的过程。其目的是在不改变失效发生物理机理情况下，缩短失效测试的周期。

加速因子：确定加速的速度

加速测试确实可大大减少可靠性测试所需要的时间，但究竟可以加速到多快呢？加速需要采用的测试条件应力、温度与加速时间之间又是什么关系？这里就有了加速因子的概念。

加速因子（Accelerated Factor）是加速测试理论中的重要概念，是指在加速应力条件下快速采集器件失效时间数据，并将这些数据外推到器件正常使用环境，得到正常使用下的失效时间。

需要说明的是，采用加速模型进行加速老化实验时，一定要确保两点：

- 加速必须是均匀的
- 不能改变失效的物理机制

第一点中，加速均匀性的要求是加速模型计算中的需求。第二点“不能改变失效物理机制”同样需要注意，需要确保不因为参数应力过大，而出现其他失效问题（如瞬间失效烧毁等）。

设计合理的加速测试环境需要对器件的失效机理有着深入的分析。在集成电路产品中，典型的失效机理包含 [2]：

1. 电迁移
2. 应力迁移
3. 腐蚀
4. 热循环疲劳
5. 时间相关介电击穿
6. 热载流子注入
7. 离子键断裂

加速测试方法需要根据以上失效机理进行合适设计。

常见的加速退化方法

加速因子与加速时间有对应关系是较为容易理解的，但要精准推导出二者的数学联系就较为复杂了。

于是，行业组织这时就发挥出了较大的作用。

	JEDEC标准	加速因子	目的
Pre-Condition	JESD22-A113	温度和湿度	可靠性试验前的预处理
HTOL	JESD22-A108	温度和电压	验证器件工作寿命
温度循环(TC)	JESD22-A104	温度和温度变化率	验证器件极端高低温下可靠性
uHAST	JESD22-A118	温度和湿度	加速腐蚀，验证高温高湿下的可靠性
bHAST	JESD22-A110	电压、温度和湿度	偏压下加速腐蚀，验证器件高温高湿下的可靠性
THB	JESD22-A101	电压、温度和湿度	偏压下加速腐蚀，验证器件高温高湿下的可靠性
高温存储(HTS)	JESD22-A103	温度	验证器件高温下的长期可靠性

图 2.182: 常见的加速测试方法

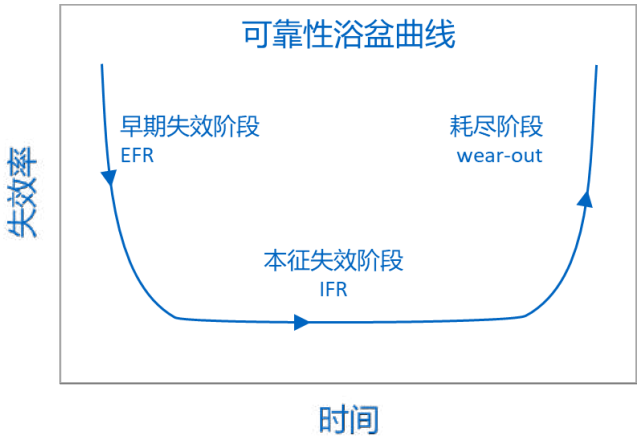


图 2.183: 可靠性浴盆曲线

1958 年，由半导体器件制造厂商、设计厂商以及终端应用厂商等，共同成立了行业标准制定组织 JEDEC（Joint Electron Device Engineering Council，联合电子器件委员会），用于制定统一的产业标准制定。

根据器件的老化机理，JEDEC 定义了系列的加速测试方法。常见的加速测试方法如图2.182所示。

JEDEC 所定义的加速测试方法已成为电子器件的标准测试方法。在实际产品交付中，需要将以上可靠性报告，伴随产品一起向客户交付。

浴盆曲线模型：将失效分时期看待

为了理解不同时期内器件失效发生的物理机理，可靠性工程中引入著名的“浴盆曲线”概念。

浴盆曲线（Bathtub Curve）认为一般器件失效有三个明显不同的阶段：

- **早夭期**（Early Failure Rate，EFR）：在器件早期使用阶段发生，失效率较高。这个时期主要的失效原因是器件生产时的严重缺陷造成；
- **本征失效期**（Intrinsic Failure Rate，IFR）：这个阶段是由于器件内部材料内部存在微小缺陷造成的；
- **耗尽期**（Wear-out）：这个阶段器件的性能已经退化，达到使用的末期。此时失效由于器件的正常退化造成。

可靠性浴盆曲线反应了失效的一般规律，甚至可以用生物的生命周期来类比：一般新生儿在出生的 24 小时内，需要医生格外关注，如果新生儿有一些出生时的缺陷，则在这个时期很容易出现生命危险；而经过一年之后一直到 70 岁左右，这个时期内死亡率较低并且趋于平缓；而到 70 岁之后，人类的器官开始出现消耗殆尽的现象。

对于集成电路来说，早夭期（EFR）可能会延续一年的时间。为了剔除存在缺陷的产品，使之不至于流向客

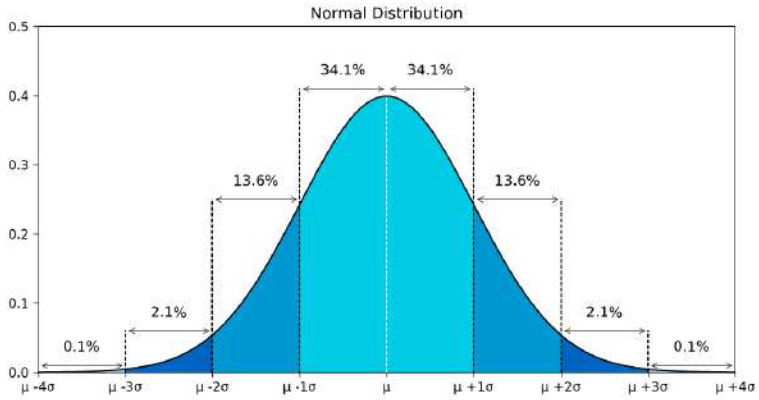


图 2.184: 正态分布及其对应区间的概率

户造成高的失效率，集成电路会采用让产品试运行一段时间，来清除有先天缺陷产品，这个清除的过程，被称为 Burn-in（老化）。

将器件做长达一年的 Burn-in 是不现实的。所以，一般会用到前一节提到的“加速退化”的方法，提升器件工作时的电压或其他工作条件的应力，使失效时间缩短到几小时、几分钟，甚至几秒钟。与加速退化的分析类似，具体所加应力大小与对应时间的关系，需要根据器件的失效各类来确定。

Burn-in 是一种对产出产品进行 100% 筛选的非破坏性实验，目的是将早夭产品剔除。由于实验对应芯片工作的时长是早期失效阶段（比如一年正常工作时间），所以并没有使器件进入耗尽期而发生性能的退化失效，也不会明显减少器件使用寿命（器件的寿命通常为数年）。

另外需要说明的是，Burn-in 是减少早期失效的一种方式，需要在每颗产品出厂前进行全数测试。如果工艺和设计电路成熟，有数据表明器件早期失效率稳定，也可以在测试中去掉 Burn-in 的测试。

2.12.1.3 统计概念的引入：评估性能波动的影响

失效时间定义了器件性能随时间的变化。在大批量使用时，还会存在不同器件间性能波动的问题。

既然不同器件之间的性能是不完全一致的，那如何衡量波动和确定卡控门限（Limit）呢？于是，统计的概念就被引入了进来。

正态分布（Normal Distribution）是数学家们在 18 世纪所发现的一种统计规律，著名高斯在 1809 年对其进行了理论推导与完善，所以正态分布又被称为高斯分布。正态分布的作用不止是揭示了一个数学现象，而是在实际工程应用中有着重要的作用。正态分布及其对应区间的概率如图2.184所示。

正态分布中两个重要参数是均值 μ 和标准差 σ。均值决定了数据的平均数，标准差决定了数据的分散程度。在正态分布中：

- 68% 的数据在平均值 1 个标准差内；
- 95% 的数据在平均值 2 个标准差内；
- 99.7% 的数据在平均值 3 个标准差内

一旦根据统计分布得出某一变量的均值和标准差后，就可以根据高斯分布，计算出任意给定区间内的概率分布。所以，此种统计的方法可以用来计算给定上下限的良率，或者推测性能波动带来的 ppm 失效率。

2.12.2 高可靠性芯片的设计和评估

基于以上可靠性工程理念，在芯片产品开发过程中加入高可靠性设计与评估。

根据芯片是否带电测试，芯片的可靠性主要分为 Electric 类与 Mechanic 类两类。射频前端产品常见的可靠性测试内容如图2.185所示。

分类	评估项目	英文名称	英文缩写
Electric	健壮性	Ruggedness	-
	稳定性	Stability	-
	静电防护	Electro Static Discharge	ESD (HBM/CDM/MM/LU)
	浪涌	Surge	-
	加压高温高湿加速	Biased High Accelerated Temperature and Humidity Stress Test	bHAST
	高温工作寿命	High Temperature Operating Life	HTOL
Mechanic	跌落	Drop Test	DT
	机械冲击	Mechanical Shock	MS
	可焊性	Solderability	SD
	湿度敏感等级	Moisture Sensitivity levels	MSL
	温度循环	Temperature Cycling	TC
	高温存储	High Temperature Storage	HTS
	高温高湿加速	Unbiased High Accelerated Temperature and Humidity Stress Test	uHAST
	加压高温高湿加速	Temperature Humidity Bias	THB

图 2.185: 常见的可靠性测试项目

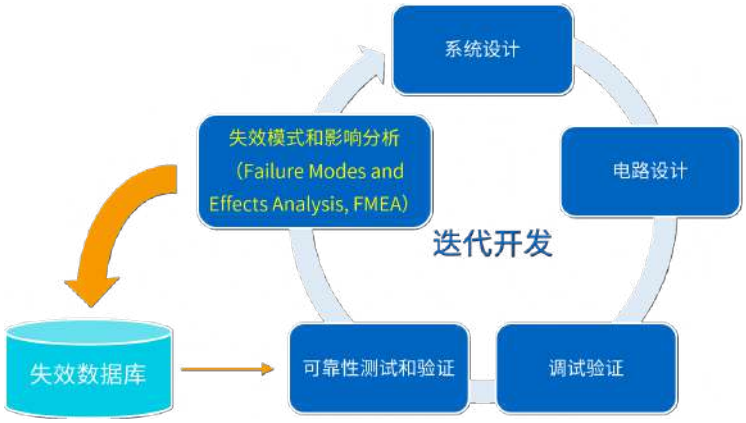


图 2.186: 包含 FMEA 的迭代开发流程

以上可靠性测试项根据产品的不同需求，选择进行评估。部分测试目的相同的测试项（如 THB 与 bHAST），可以二选一进行测试。

2.12.2.1 高可靠芯片的设计

“可靠性是设计出来的，不是测试出来的”是高可靠芯片设计中的重要理念。实现满足以上可靠性需求的高可靠模组芯片，必须要在设计之初就对芯片的高可靠性加以考虑。

常见的高可靠性芯片设计思路是 FMEA（Failure Modes and Effects Analysis）方法，图2.186为典型的包含 FMEA 的迭代开发流程。

2.12.2.2 高可靠芯片的验证

芯片的可靠性验证一般遵循一定的测试标准，常见的测试标准为前一章节提到的 JEDEC 标准，另外还有车规产品验证的 AEC-Q 系列标准。

这些标准中建立起标准的测试方法。同时，在器件选取数量、判定标准上也做了明确规定。图2.187为 JEDEC 相关测试项及测试标准。高可靠的芯片设计必须完全通过以下测试。

缩写	项目	参考标准	试验条件	取样数目
TC	Temperature cycle	JESD22-A104	-65℃~150℃, 500 cycles	3 Lots * 77 pcs
HTOL	High Temperature Operating Life	JESD22-A108	125℃, 1000hrs	3 Lots * 77 pcs
CDM	Electronstatic Discharge Charge Device Mode	JESD22-C101	1kV	3 Lots * 9 pcs
HBM	Electronstatic Discharge Human Body Mode	JS-001-2014	1kV / 2kV	3 Lots * 9 pcs
LU	Latch Up	JESD78	±100mA/ 1.5*Vmax/MSV	3 Lots * 3 pcs
HTS	High Temperature Storage	JESD22-A103	150℃, 1000hours	3 Lots * 77 pcs
DT	Drop Test	JESD22-B111	1500Gs, 0.5ms	20 pcs
SD	Solderability	JESD22-B102	-	3 lots * 22 pcs
bHAST	High Accelerated Temperature and Humidity Stress Test	JESD22-A110	96 hrs @ 130℃, 85%RH	3 Lots * 25 pcs
uHAST	Unbiased Temperature/Humidity (Unbiased HAST)	JESD22-A118	130 °C, 85% RH, 96hours	3 Lots * 77 pcs

图 2.187: JEDEC 所规定的测试项及条件（部分）

2.12.3 总结

经过近年来的发展，国产芯片已经逐渐占稳中低端市场，并向中高端进发。中高端市场需要的芯片不仅仅需要有高性能、高集成度，更重要的是有过硬的质量。高质量、高可靠性的芯片是中国芯片设计企业从“以价取胜”走上“以质取胜”的必经之路。

高质量芯片是先进的设计能力、优质的流程管理、完整的可靠性评估带来的综合结果。开发高质量芯片，对设计企业的综合能力和开发投入提出了高的要求。慧智微期待利用高质量的射频前端芯片，助力提升“中国制造”的综合竞争力。

本文撰写过程中，得到了慧智微质量团队的专业讨论与梳理，在此表示感谢！

您在工作中有哪些提升芯片可靠性的经验，欢迎扫码留言分享讨论。

参考文献

1. 章晓文等, 2015, 半导体集成电路的可靠性及评价方法
2. 姚立真，2004，可靠性物理
3. J. W. McPherson, 2019, Reliability Physics and Engineering
4. https://www.jedec.org/

在线阅读

本文摘自慧智微微信公众号，扫描以下二维码，在线阅读

第三篇 射频前端模组

本篇内容

- 从 4G 到 5G，手机射频芯片十年之路
- 什么是 Phase8？它将给 5G 带来什么变化？
- 5G 射频前端模组中的滤波器
- 5G L-PAMiD 的前世与今生
- 尺寸缩小 40%：国产工规级 5G n77/n79 L-PAMiF 领先量产

随着通信技术的快速演进，射频系统也变得越来越复杂。在实际应用中，射频芯片不再是单一功能的器件，而是由多个子电路紧密结合而成的复杂系统。这种复杂性不仅体现在芯片内部的结构设计上，还体现在芯片与外部系统的连接和交互上。为了应对这种复杂性，射频前端的设计趋势逐渐向着模组化方向发展。

模组化设计是射频前端领域的一项重要演进趋势。它将多个射频子电路集成在一个模组中，通过优化布局和连接方式，提高整体性能，降低系统复杂性和成本。了解模组化设计的演进背景和技术逻辑，对于我们深入理解射频前端技术的发展至关重要。

在本章中，我们将首先回顾手机射频芯片过去 10 年的发展脉络。从最初的单一功能芯片，到后来的多芯片组合，再到如今的模组化设计，手机射频芯片经历了翻天覆地的变化。我们将详细阐述这一发展过程中的关键技术节点，介绍 Phase2 直到 Phase8 的技术架构及方案，帮助大家把握射频前端技术的演进轨迹。

我们还将对 L-PAMiD 的过去与未来发展做一个梳理。L-PAMiD 是射频前端模组中的一个重要组成部分，它负责处理低频段的信号放大和传输。随着通信技术的不断发展，L-PAMiD 也在不断演进，以适应更高的性能要求和更小的尺寸限制。我们将分析 L-PAMiD 的技术特点、应用场景以及未来发展趋势，为大家提供全面的了解。

通过本章的讨论，我们希望能和大家深入讨论射频前端模组化设计的演进背景和技术逻辑，理清手机射频芯片的发展脉络和现状，以及 L-PAMiD 的关键技术和未来趋势。了解这些背景信息，将有助于我们在实际应用中更好地选择和使用射频前端模组。

3.1 从 4G 到 5G，手机射频芯片十年之路

作者：彭洋洋 博士

2011 年巴塞罗那举行的世界移动大会（MWC 2011）中，时任中国移动董事长王建宙表示，2011 年将是 4G TD-LTE 商用元年。4G 两大通信标准 TD-LTE 和 FDD-LTE 在 2011 年全面正式商用。

到了 2021 年，通信协议 10 年一变，5G 成为代表未来的通信标准。在正式商用两年之后，5G 即达到了超过 4 亿的连接数。5G 生态逐渐形成，蓬勃发展。

在过去十年中，无论对于已经基本固化的 4G 通信协议，还是仍在演进中的 5G 协议，射频前端方案都在不断演进。射频前端的“十年一大变，两年一小变”，演变出看似复杂的多种方案。MMMB PA、TxM、L-PAMiD、L-FEM……各种简写名词层出不穷。本文对过去 10 年手机射频前端方案做一个整理，和大家一起回顾射频前端方案的过去，讨论射频前端方案可能的未来。

3.1.1 射频前端方案演进

3.1.1.1 “Phase X”系列的出现

现在谈到手机射频前端方案，无不例外都会提到“Phase2 方案”、“Phase7 方案”、“Phase7LE 方案”等名词，这些方案是怎么出现的呢？

一般射频前端方案由器件厂商、平台厂商及终端厂商三方共同定义开发完成。如图3.1所示。

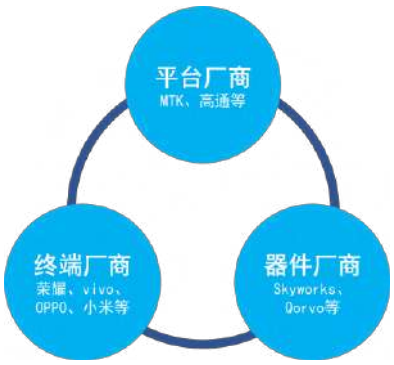


图 3.1: 参与射频前端方案定义的厂商

在 2010 年左右之前，手机平台方案并不像当前这么集中。除了 MTK、高通及展讯外，ADI、TI、Agere、Infineon、Philips、Freescale、Renesas、Skyworks 等公司，都提供过手机平台解决方案。由于当时平台方案分散，射频方案这个细分方向对射频技术要求高，所以并没有平台厂商可以将射频方案统一起来。在这个时期，射频方案定义的主导权主要在射频器件厂商，由射频器件厂商发起定义、平台适配和客户推广。

在 2010 年之后，MTK、高通、展讯及海思平台的崛起，手机平台方案越来越集中。越来越集中的还不止是平台厂商，随着山寨机的没落，终端厂商也逐渐向头部聚集。平台厂商、终端厂商及器件厂商，都对射频前端器件“生态”的形成更加重视，能否形成器件统一的“生态”是新方案定义中非常重要的考虑点：

- **平台厂商：**
  - 平台厂商不断积累射频前端的定义能力，并将射频前端方案统一纳入到方案规划中，统一方案，向客户提供一站式“turn-key”解决方案。
  - 射频前端愈加复杂，新的射频前端方案必须要完成复杂的平台适配才能完成方案应用，这也使得平台厂商在射频前端方案定义的话语权进一步增加。
- **终端厂商：**
  - 需要不同供应商提供的器件之间可以灵活替换，降低供应风险。
  - 方案统一可以降低应用时的适配难度。

编号	芯片名称	芯片功能
1	Qorvo某型号	ASM，Antenna Switch Module。天线切换开关，完成不同频段的天线切换
2	Sky77807	4G高频（2.3-2.7GHz）多频多模PA
3	Sky77648-11	4G中低频（0.7-0.9GHz，1.7-1.9GHz）多频多模PA；2G PA

图 3.2: 典型的射频前端方案

- 器件厂商：
  - 收敛统一的方案可降低多个方案带来的巨额开发和维护费用，降低生产成本。
  - 方案的统一降可低新器件的开发风险。

由于以上考虑，MTK 发起定义的“Phase X”系列射频前端方案受到终端厂商、器件厂商的支持，成为公开市场过去近 10 年主流的射频前端方案。MTK 的“Phase X”系列方案伴随了整个 4G 的发展，占据整个 4G 市场约 80% 的市场份额，并且在 5G 时代依然是公开市场最为主流的方案。以下将对 Phase 系列方案进行详细讨论。

以上为公开市场方案定义的变化，还有一部分射频前端方案是直接由终端厂商发起定义的，这种方案出现在头部终端厂商中。如苹果 iPhone 手机中所用的射频前端方案均为苹果公司所自定义的方案。这种自定义的方案也出现在华为及三星手机之中。自定义射频前端方案全年出货比例约占整体射频前端方案的 20% 以下，并且比较闭环，不在本文讨论范围。本文将主要讨论公开市场中通用的射频前端方案演进。

3.1.1.2 Phase1：史前时代

严格来说，并没有 Phase1 方案的定义。

MTK 的射频前端定义是从 2014 年的 Phase2 开始的。在 Phase2 推出之前，TD-LTE/FDD-LTE 已经全面商用 3 年了，在这三年出现的方案一般称为 Phase1 方案。

Phase1 方案并不统一，一般来说是最大程度的复用射频前端厂商 3G 时代的产品定义：与原来 2G/3G 重合的频段复用原来的 pin 脚；4G 的新频段用单独分立的通路进行覆盖；再用天线天关将所有频段切换到同一根天线上。

图3.2为典型的射频前端方案，发射部分主要由三款芯片构成。图3.3为典型的手机方案框图。

除了以上型号，在同一时期的 Sky77621、Sky77753、RF7378 等，都属于“Phase1”时代的芯片方案。这些方案主要由 Skyworks、RFMD（现 Qorvo）厂商定义开发。

3.1.1.3 Phase2：顺应时代，成就经典

Phase2 方案是 MTK 由 2014 年定义的第一代归一化 4G 射频前端方案，7 年过去回头来看，Phase2 方案的定义依然经典。用现在时髦的一句话说：Phase2 方案，YYDS（永远滴神）。

如前文所述，在 2G/3G 时代，射频前端的方案并不统一，Skyworks、RFMD（现 Qorvo）等公司时常会有缺货发生。不少国内创业公司在 2011 年前后，依靠 RF9810、Sky77590 等芯片缺货挖掘到了第一桶金。缺货对国内创业公司是机会，但对终端及平台厂商却是灾难：射频前端的缺货会影响到平台出货和终端生产。于是，MTK 在 2012 到 2013 年左右开始着手定义 Phase2 方案。

Phase2 方案的定义不仅仅考虑到了当前方案的统一，还考虑到了方案生态的可达成性、未来协议的演进、4G 三模/五模的共存等等。Phase2 对于 Phase1 的改动主要如图3.4所示。

Phase2 将 Phase1 的 2G PA，与 ASM（Antenna Switch Module，天线开关模组）整合，形成 TxM（Transmitter Module，发生模组），将 4G 频段的 PA 整合，形成完整的 4G MMMB PA（Multi-Mode, Multi Band Power Amplifier Module，MMMB PAM，习惯简称为 MMMB PA 或者 MMMB）。

经过改进，Phase2 方案有以下优势：

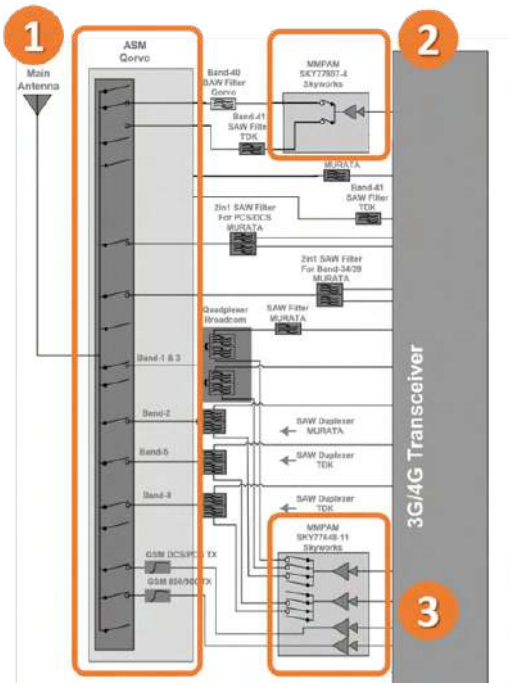


图 3.3: 荣耀 6X 射频前端方案框图（来源：Navian RF Devices / Modules For Cellular, 2016）

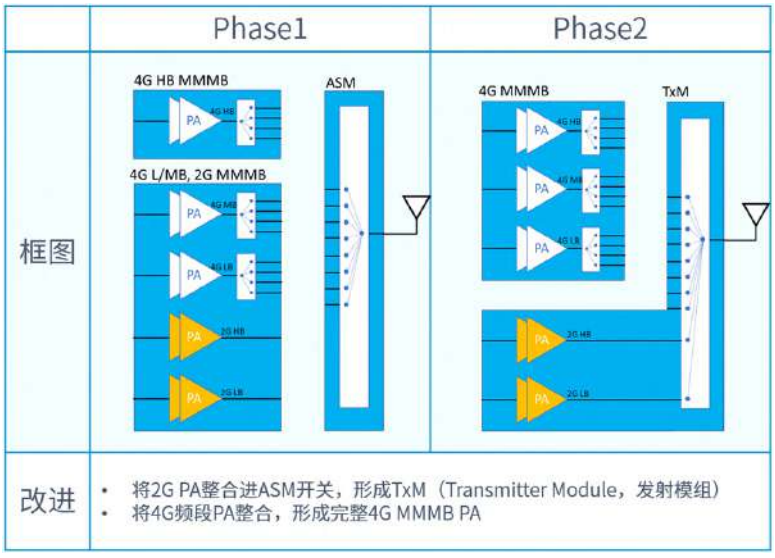


图 3.4: Phase2 对于 Phase1 的改动

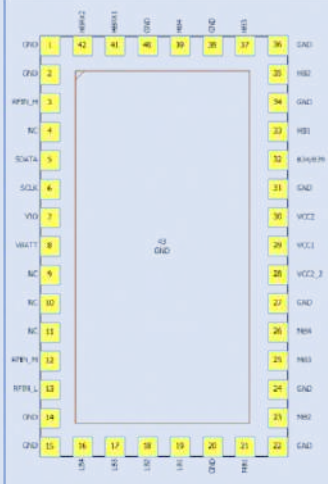
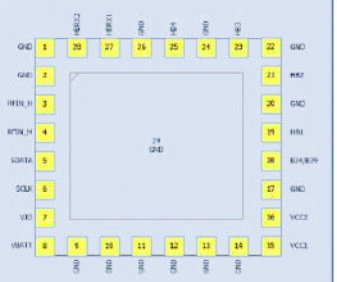
	Phase2 5模MMMB PA	Phase2 3模MMMB PA
框图		
尺寸	• 4x6.8 mm²	• 4x3.65 mm²
支持模式	• WCDMA、TDS-CDMA、FDD-LTE、TDD-LTE	• TDS-CDMA、TDD-LTE
代表产品	• Sky77643 (Skyworks) • S318、S5643 (慧智微)	• Sky77824 (Skyworks) • S816 (慧智微)

图 3.5: Phase2 的定义

- 灵活性增强：由于 2G PA 的设计方法与 3G/4G PA 有大的不同，2G PA 与 4G PA 的分离可以带来大的设计灵活性。
- 2G PA 一般采用 SAW-less 方案，输出不需要经过额外的 SAW 滤波器等，2G PA 与 ASM 集成为 TxM 可以降低 2G PA 后端插损。
- 供应商可针对性发挥优势：不同供应商在 2G 与 3G/4G 的技术积累与能力不同，分离后可以极大化的发挥不同供应商的优势。
- 2G PA 的分离，可为后续 2G 退网做好准备。
- 4G 频段的整合，为日后 4G 乃至 5G 频段的发展做好准备。

Phase2 在定义时，还考虑到了不同运营商的兼容。除了定义全网通的 4G MMMB PA 芯片，支持 GSM、WCDMA、TDS-CDMA、TD-LTE 及 FDD-LTE 的 5 模方案外，还定义了只支持中国移动频段的 GSM、TDS-CDMA、TD-LTE 的 3 模方案芯片，这两颗芯片尺寸大小不同，但却可以共板替换，定义的相当巧妙，如图3.5。不过由于 3 模市场很快过去，全网通手机成为大势，MTK 巧妙的兼容定义并未被大规模使用起来。

国际厂商的 Phase2 的代表产品是 Skyworks 的 Sky77916 + Sky77643，以及 RFMD（现 Qorvo）的 RF521X + RF5422。国内厂商也在这个时代进行追赶。方案归一化后，国内厂商开始加大投入，开发与 Skyworks/Qorvo 等国际厂商相同方案的产品。虽然产品定义和目标产品都是清晰的，但过去几年国内厂商和 Skyworks/Qorvo 等厂商的竞争中并不占优。直到 2020 年前，国内市场在 4G Phase2 市场的市占率仍小于 10%。Skyworks/Qorvo 等公司在国内竞争对手杀入后，仍然保持优势的原因是：

- 国内厂商大多数采用“跟随战略”，即在 Skyworks/Qorvo 推出产品之后，快速进行拷贝和复制，推出功能类似的产品，然后靠低价格杀入市场。
- Skyworks/Qorvo 等国际厂商性能更优：有近 20 年的技术积累，采用相同的方案，国内厂商无法在性能上超越（在相同方案中，目前综合性能仍然是国际厂商最优）。
- Skyworks/Qorvo 成本更低：IDM 模式、年 20 亿美金以上的销售额，支撑国际厂商拥有更低成本（国内均为 Fabless 模式，并且销售额小 1 到 2 个数量级）。

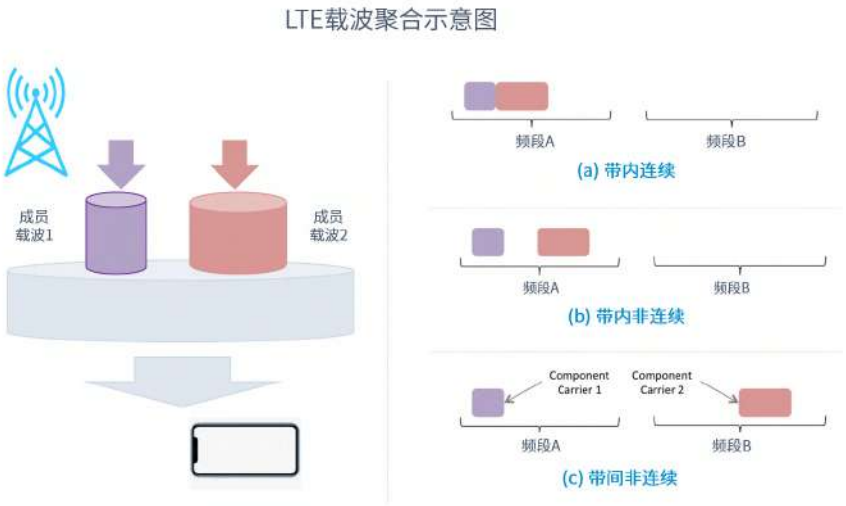


图 3.6: LTE 载波聚合示意图

- Skyworks/Qorvo 掌握知识产权：同质方案的知识产权风险，让头部客户使用有顾虑。

2020 年以后，随着 4G 毛利逐渐降低，5G 市场逐渐起来，国际厂商开始放弃对 4G 市场的占据，国内厂商得以机会在 4G 市场取得份额。但由于定价权依然在国际厂商手中，若采用相同方案，国内厂商依然无法取得可观的毛利。

慧智微在 2011 年成立时即尝试采用创新的方案进行设计，取得性能、成本及知识产权优势。慧智微采用自有知识产权软件定义可重构方案实现的 Phase2 射频前端拥有更少硬件，并可以实现在频段内的性能调谐。目前已实现数亿片芯片出货。

3.1.1.4 Phase3/5：完善方案，支持 CA

Phase3 及 Phase5 的定义在 2015 年到 2016 年，也是全球 4G 建设最为火热的时候。除中国外，大部分运营商获得的频谱都是通过拍卖的方式获得，频谱资源珍贵，运营商一般无法获得连续较宽的频谱。相比于中国移动在 4G 时代 B41 获得的 2575-2635MHz 的 70MHz 带宽（进入 5G 后，中移动在 B41/n41 带宽将拓展至 160MHz），国际运营商通常只有几 MHz 或者 10 几 MHz 信号带宽。为了提升用户体验，CA（Carrier Aggregation, 载波聚合）方案开始被大家关注。

载波聚合是 LTE-A 中的关键技术，可以将 2～5 个 LTE 成员载波（Component Carrier，CC）聚合在一起，实现最大 100MHz 的传输带宽，有效提高了上下行传输速率。

按照上下行 CA 的功能不同，CA 可分为下行 CA（DL-CA，Down Link CA）及上行 CA（UL-CA，Up Link CA），下行 CA 只完成下行载波聚合，提升下行速率；上行 CA 一般可同时完成上行及下行载波聚合，上下行速率均可得到提升。

按照载波频段的不同，CA 可分为带间 CA（Inter Band CA），及带内 CA（Intra Band CA）。同时，带内 CA 又有连续与非连续之分。如图3.6所示。

CA 方案较为复杂，不同细分场景和不同的 CA 组合需要有不同的方案来响应。MTK 先后定义了 Phase3 及 Phase5 来支持不同的 CA 场景。Phase3 可以支持到 2 下行 CA 及带内上行 CA；Phase5 利用三工器、多工器的引入，又将 CA 能力提升到了 3 下行 CA 及带间上行 CA，不过 PA 后端插损增加，对 PA 输出功率的要求提升。如图3.7所示。

由于分立方案实现 CA 较为复杂，Phase3 及 Phase5 作为完整射频前端方案并未形成大规模生态。CA 市场并非全球市场，所以 Phase3/Phase5 也没有形成对 Phase2 的取代，对 CA 能力有强需求的主要是海外高端手机，在 Phase6 PAMiD 方案定义完成后，这些手机快速转向了 PAMiD 方案。

在 MTK 的定义中，并没有“Phase4”方案，原因是华人社会对数字“4”的避讳。据说 MTK 对“Phase4”

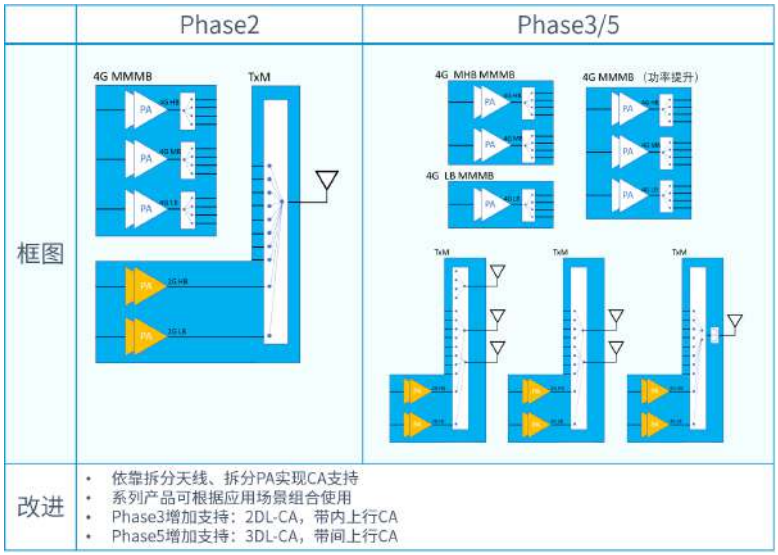


图 3.7: Phase3/5 的定义

的跳过，也让 Qorvo/Skyworks 等国际厂商了解到了“4”这个数字在中文发音的额外含义，使国际厂商在产品命名中的数字使用也更加慎重。

3.1.1.5 Phase6/Phase6L：进入 PAMiD，依然经典

在分立方案开发完成后，国际大厂开始向 PAMiD 深度布局：2014 年，Skyworks 宣布与松下组建合资公司；2015 年，RFMD 与 Triquint 合并，成立 Qorvo 公司；2016 年，高通宣布与 TDK 建立新的合资公司 RF360。

PAMiD 的全称是 PA Module integrated with Duplexer，PA 滤波器集成模组。在这个模组中，同时集成了 PA 模组与滤波器组，也集成了天线开关等。PAMiD 集成度高，链路插损小，使用简便，一直是高端手机的首选方案，iPhone 从 iPhone4 时代，即开始采用 PAMiD 方案，方案来自于 Avago（现 Broadcom）、Skyworks、Triquint/RFMD（现 Qorvo）等厂商。

虽然射频前端厂商在 2016 年之前就在 iPhone 等手机上应用 PAMiD 方案，每家厂商也都有自己的方案在推广，但公开市场一直缺少统一定义，PAMiD 方案在公开市场并没有很好的应用。

MTK 在 2016 年推出 PAMiD 方案 Phase6 定义，随后又进行成本优化，去掉冗余载波和滤波器，升级到更贴合中国市场的 Phase6L（Phase6 Lite），Phase6L 也在公开市场的 PAMiD 方案中取得成功。其框图如图3.8所示。

在 MTK Phase6/Phase6L PAMiD 方案成功定义的 2016 年前后，MTK 先后发布中高端 Helio P 系列以及旗舰 Helio X 系列 SoC，准备与高通在旗舰市场一决高下。但随着手机终端厂商将 MTK Helio X10 芯片应用于千元档位手机，MTK SoC 平台的旗舰之路遇阻，随后 MTK 宣布放弃旗舰平台 Helio X 的开发。MTK 所定义的 Phase6/Phase6L PAMiD 射频前端方案，与当时 MTK SoC 平台所处的千元机市场无法匹配。

虽然 MTK 平台冲击高端受限，也不妨碍 MTK 所定义的 PAMiD 射频前端方案成功。MTK 所定义的 Phase6 以及 Phase6L 射频前端方案，先后在海思、高通等其他平台方案中量产。即使当时 MTK 平台没有冲向高端，但 MTK 所定义的射频前端生态也在高端手机市场量产，MTK 射频前端的定义能力及号召能力可见一斑。

3.1.1.6 Phase7/Phase7L/Phase7LE：5G 的开门红

5G 对全世界来说都是新的。

5G 频段是新的，标准是新的，甚至需求也是不断变化的。在需求未清晰的情况下，5G 早期的方案也差别很大。高通、华为海思、村田、Qorvo 及 Skyworks 等厂商，都在 2018 年推出过不同形式的方案。

MTK 在对协议、运营商、终端客户及器件厂商的信息综合分析后，定义了 Phase7 方案。Phase7 方案的 Sub-3GHz 部分主要由 Phase6/Phase6L 继承而来。在 5G 新增加的 Sub-6GHz UHB 部分，重点定义了支持 n77/78/79

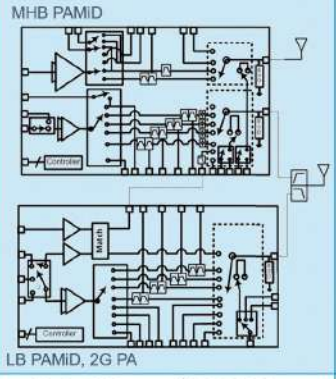
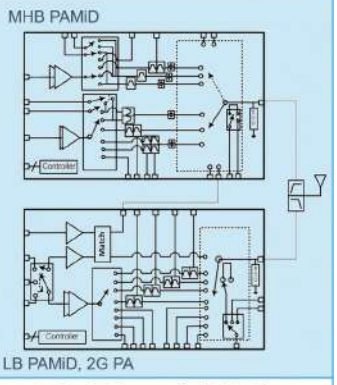
	Phase6	Phase6L
框图		
尺寸	• 8.6x6.5 mm² (MHB) • 7.6x6.0 mm² (LB)	• 8.6x6.5 mm² (MHB) • 7.6x6.0 mm² (LB)
集成器件	• PA、滤波器、双工器、ASM开关等	• PA、滤波器、双工器、ASM开关等
支持功能	• 2G GSM • 4G LTE Cat 16 • DL 4CC, UL 2CC • 3 Ant • ET	• 2G GSM • 4G LTE Cat13 • DL 3CC, UL 2CC • 2 Ant • APT
代表产品	• QM77031 (MHB) • Sky78185-11 (LB)	• QM77031 (MHB) • Sky78185-21 (LB)

图 3.8: Phase6 系列产品的定义

频段、集成 SRS 开关的双频高集成模组。

Phase7 方案的推出，很好的适应了 5G 的新需求，众多终端厂商的 5G 射频前端方案快速切换至 Phase7 方案。

MTK 将 5G 平台方案取名“天玑”，并发布 1000、800、700 系列，布局 5G 高、中、低端市场。由于 5G 完整方案的推出，MTK 平台在 5G 大有斩获，图3.9为 MTK 平台 2021 年市占率增长情况，在 2021 年 Q2，MTK 平台出货市占率达 43%，比其他第三方平台（高通、展锐、三星）之和还要多。

在推出第一代 Phase7 之后，MTK 快速定义 Phase7L（Phase7 Lite）、Phase7LE（Phase7L Enhancement，Phase7L 增强版），适应 5G 市场的快速变化需求。Phase7/Phase7L/Phase7LE 各代之间的演进关系如图3.10所示。

Phase7 主要应对初期的 5G 应用，基于 Phase6 增加了 5G 的支持，包括：

- Sub-3GHz PA 进一步提升功率及线性，以支持 5G 高功率、高阶调制的需求；
- 天线开关复杂度升级，以应对 5G 对 SRS 切换、MIMO、智能天线切换的需求；
- 方案提升了 eLNA 的重要性，定义了集成 eLNA、RX filter 和 Switch 的 L-FEM 产品形态。

Phase7L 基于快速发展的 5G 需求，进行了迭代，Sub-3GHz 进一步提高集成度，在 PAMiD 产品形态中加入主集接收 LNA，形成 L-PAMiD 产品形态。Phase7LE 随着 5G 需求趋于收敛应运而生，预计将成为未来主要方案：

- UHB 从 1T1R L-PAMiF 及 1R L-FEM 方案，演进至 1T2R/2R 的产品方案，进一步提升集成度；
- 继续优化模组内开关、ENDC 支持、双工器等功能，进一步减少模组外围器件需求，达到整体方案的高性能和简洁。

随着 5G 应用的推进，5G 射频前端方案也开始收敛。不止是 MTK 平台，高通及展锐平台的方案也统一至 Phase7 系列方案。值得一提的是，高通自有品牌的最新 5G UHB 射频前端产品，也逐渐向 MTK 所定义 Phase7 系列方案靠拢。Phase7 系列成为又一个经典方案。

得益于深厚的技术积累，慧智微快速响应了 MTK Phase7 系列方案的定义。慧智微于 2019 年 12 月，作为国内首家射频前端公司推出了兼容于 Phase7 的 UHB 高集成 L-PAMiF 及 L-FEM 模组，并在 2020 年实现包括国际

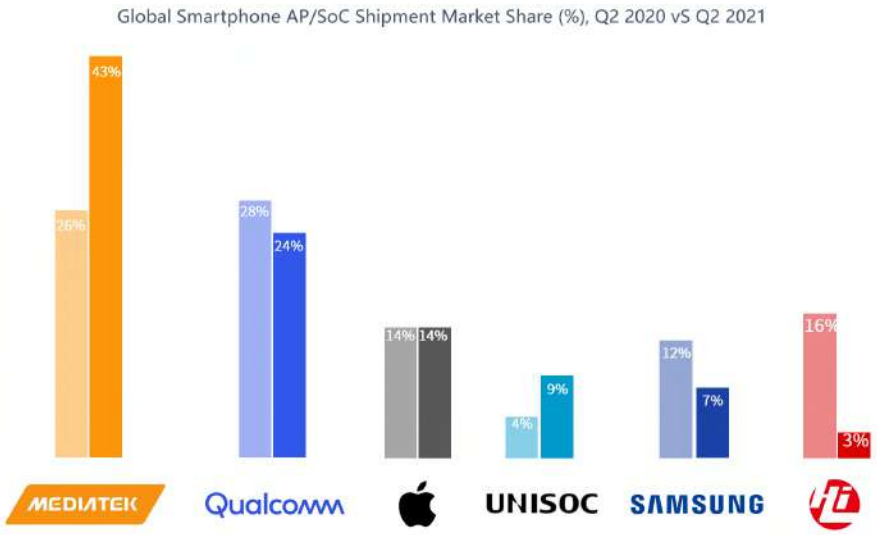


图 3.9: 智能手机 SoC 平台出货占比（2020 Q2 及 2021 Q2，来源：Counterpoint）

头部终端厂商在内的数十款 5G 终端量产。

慧智微在 5G 的突出成果也得到了国内及国际行业组织的认可，工信部“中国芯”将 2020 年重大创新突破产品奖颁发给了慧智微，5G 全球推动组织 GTI 将 2021 年度荣誉奖授予慧智微。慧智微分别作为国内的首家射频前端公司，得到了以上两个奖项。在 2021 年 9 月，慧智微被高通评选为“2021 5G 生态最值得关注的十大新锐创业企业”。

2021 年 4 月，慧智微 Phase7LE 高集成 1T2R UHB L-PAMiF 及 2R L-FEM 模组客户送样，2021 年 9 月产品正式量产，慧智微实现 5G 新方案产品持续领先推出。

3.1.1.7 Phase5N：虽非官方定义，但却顺理成章

虽然 Phase3、Phase5 作为完整方案并未成为全球性的大节点，但 Phase3、Phase5 定义下所产生的个别芯片在日后方案中有了举足轻重的作用：Phase3 时代定义的 TxM，可以很好的支持 5G 时代的多天线场景；Phase5 时代因为 CA 方案中后端引入四工器、Diplexer 等插损增加，将 Phase2 MMMB PA 的功率提升了 1dB，这提升的 1dB 受到了终端厂商的欢迎，可以用来抵消部分应用中 PA 后端过大的插入损耗，部分厂商直接将提升功率版的 MMMB PA 称为“Phase5 PA”。

5G 到来之后，头部终端厂商主导将 Phase5 MMMB PA 增加支持 5G NR 信号的定义，被业界称之为 Phase5N（“N”代表支持 5G NR）PA，基于这颗 MMMB PA 所构建起来的 5G 方案也称之为“Phase5N 方案”。由于大家对 Phase2/5 MMMB PA 相当熟悉，Phase5N PA 只是在原来的基础上增加了 5G NR 信号支持，pin 脚等未做修改，这颗物料也顺理成章的得到大家的接受。其架构如图3.11所示。

Phase5N 5G 方案中，Phase5N MMMB PA 只是对 MTK 5G 方案的 Sub-3GHz 部分做了修改，在 Sub-6GHz UHB 部分依然沿用 MTK 的定义。Phase5N 方案与 MTK 所定义 Phase7/7L/7LE 中 PAMiD/L-PAMiD 的替代关系如图3.12所示。

不过需要说明的是，Phase5N 并不是 MTK 的定义。到目前为止，MTK 还没有正式推出过分立 5G NR 的射频前端方案定义。MTK 官方的 5G 方案，依然是集成的 PAMiD/L-PAMiD 方案。Qorvo、Skyworks 也没有响应 Phase5N 方案，而是推广单价更高的 PAMiD/L-PAMiD 产品。

在行业的共同推进下，射频前端方案以如图3.13的“两年一个大节点”，进行了十年演进。

3.1.2 射频前端方案的未来

过去 10 年，平台深度参与射频前端方案的定义，安全保障了在 4G 的大规模商用以及 5G 新协议的快速部署。未来射频前端方案将演进到什么方向呢？

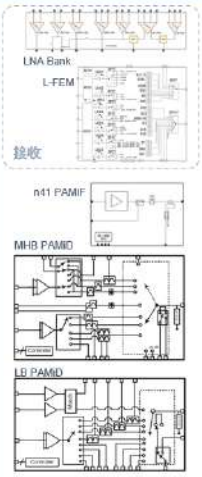
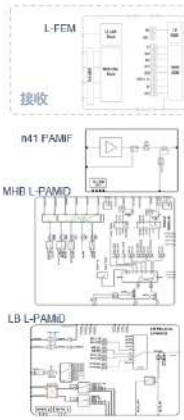
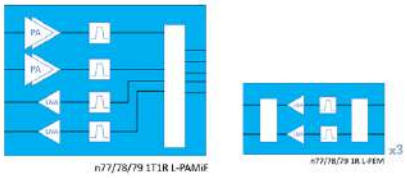
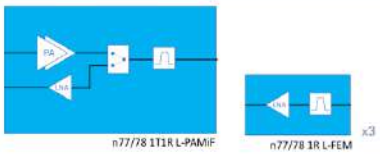
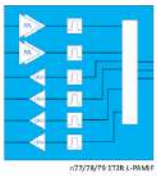
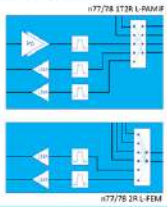
	Phase7	Phase7L	Phase7LE
Sub-3GHz			
说明	• 沿用Phase6定义，增加NR支持 • 新增L-FEM等定义	• PAMiD内增加LNA	• MH L-PAMiD增加 EN-DC
尺寸(仅PAMiD)	• 8.6x6.5 mm² (MHB) • 7.6x6.0 mm² (LB)	• 8.6x6.5 mm² (MHB) • 7.6x6.0 mm² (LB)	• 8.6x6.5 mm² (MHB) • 7.6x6.0 mm² (LB)
代表产品	• QM77040/38 • Sky58091	• QM77048 • Sky58080/1	• QM77058 • Sky58100/1
Sub-6GHz UHB (5G Sub-6GHz新增加频段)	双频  单频 	双频  单频 	
功能	• UHB n77/78/79 L-PAMiF及L-FEM (1T1R)		• UHB n77/78/79 L-PAMiF及L-FEM (1T2R)
代表产品	• 双频： • Sky58255+Sky53728 • S55255+S15728 （慧智微） • 单频 • QM78207+Murata-507 • S55207+S15720（慧智微）		• 双频： • Sky58281 • S55281（慧智微） • 单频 • Sky58277+Sky53759 • S55217+S15722（慧智微）

图 3.10: Phase7/Phase7L/Phase7LE 各代之间的演进关系

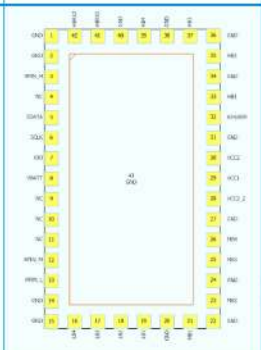
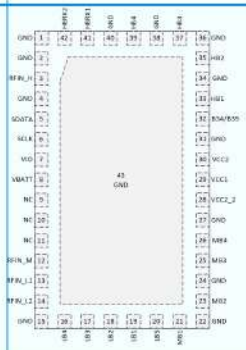
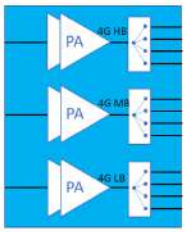
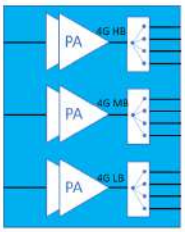
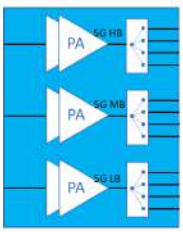
MMMB PA部分	Phase2	Phase5	Phase5N
Pin脚			
功能			
尺寸	• 4x6.8 mm²	• 4x6.8 mm²	• 4x6.8 mm²
支持模式	• 3G/4G	• 3G/4G	• 3G/4G/5G
输出功率	• PC3：27.5~28dBm	• PC3：28.5~29dBm • PC2：31~32dBm	• PC3：29~29.5dBm • PC2：31~32dBm
代表产品	• Sky77643-21(Skyworks) • S5643-51 (慧智微)	• Sky77643-61(Skyworks) • S5643-62/-66 (慧智微)	• S55643-11(慧智微)

图 3.11: Phase5N 的方案演进

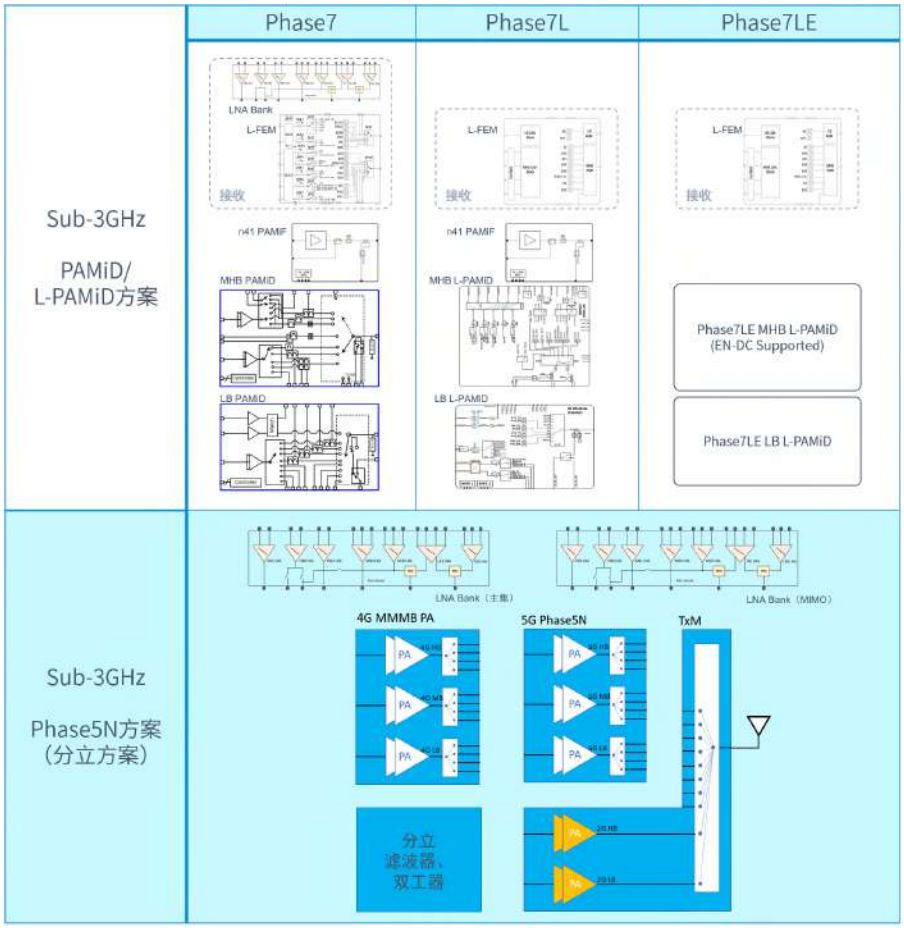


图 3.12: Phase5N 方案与 MTK 所定义 Phase7/7L/7LE 中 PAMiD/L-PAMiD 的替代关系

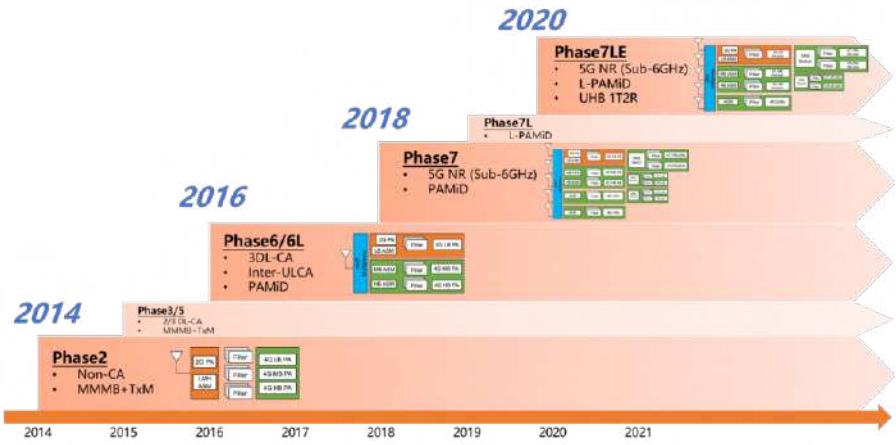


图 3.13: 射频前端方案的“两年一个大节点”

猜测有以下几个可能的趋势：

- **射频前端方案继续强调“生态”**

  生态的形成会带来良好的质量，合理的价格，安全的供给。有良好生态的射频前端方案将继续是终端厂商的优先选择。

- **头部终端厂商深度参与规格定义与产品定制**

  目前头部终端厂商越来越集中，并且对射频前端的理解能力也越来越强。除了苹果、三星、华为之外，国内的 OPPO、vivo、小米及荣耀也都已经具备射频前端方案的定义能力。未来头部终端厂商将深度参与到射频前端产品定义中来。

- **高集成模组化是大方向**

  受限于可集成化小型 SAW/BAW 滤波器及双工器，国内厂商现在还无法在 Sub-3GHz 提供 PAMiD 及 L-PAMiD 方案。随着越来越多优秀公司的投入，一旦滤波器及双工器供应解决，国内厂商有望实现 PAMiD 及 L-PAMiD 模组产品实现突破。

- **核心技术为王**

  只有掌握核心差异化的技术，才有机会在归一化生态的产品竞争中获胜，才有机会深度参与到头部客户的差异化定制中来。未来竞争将更加激烈，更考验厂商的核心技术能力。

3.1.3 总结

射频前端芯片行业是一个备受关注的行业，本文尝试回顾过去十年手机射频前端方案的发展，提供一些信息供参考讨论。射频前端芯片行业也是一个快速变化的行业，射频前端方案与通信协议息息相关，“十年一大变，两年一小变”是这个行业过去几年发展的规律。

观往知来，只有了解射频前端的过去，才能把握射频前端的未来。慧智微电子期待与你一起了解过去，开创未来。

本文整理过程中，得到众多同行、同事的帮助，在此一并表示感谢。

说明：部分图片来自网络及公开渠道。

附，简称名词解释及框图：

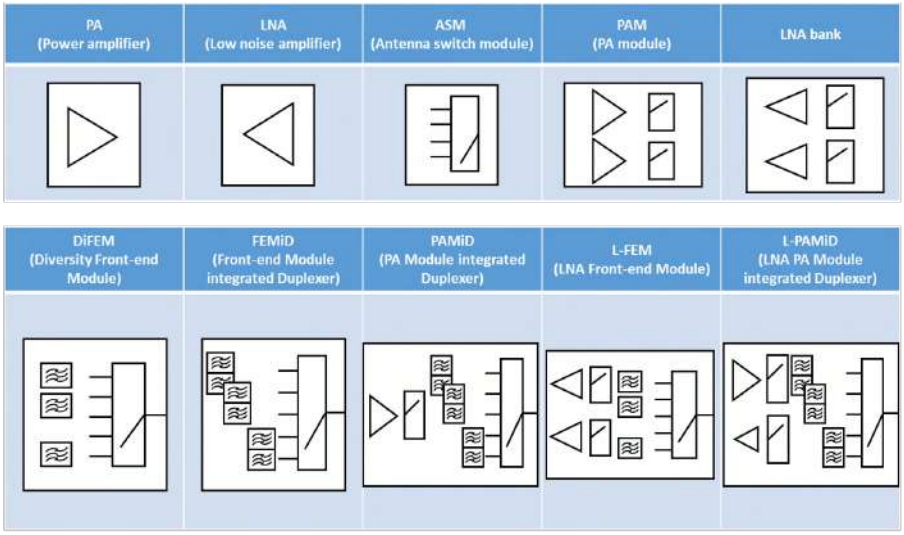


图 3.14: 简称名词解释及框图

在线阅读

本文摘自慧智微微信公众号，扫描以下二维码，在线阅读

3.2 什么是 Phase8？它将给 5G 带来什么变化？

作者：彭洋洋 博士

2023 年，5G 商用第四个年头。“能用就好”的阶段已过，5G 正在迈向“深度优化”。

中国 5G 手机出货在过去几年迅速提升，根据工信部统计数据，在中国市场，5G 手机出货比例从 2019 年的 4.7%，快速增长至 2020 年的 52.9%，2021 年增长至 76.3%，2022 年至今，这个比例稳定在 77% 以上。中国市场年出货手机中，已经有接近 8 成均为 5G 手机。

5G 在国外也开始迅速部署。除了北美、西欧、亚太等 5G 先发地区外，南亚、东欧、北非、中南美洲等地区也在快速推进 5G 部署和商用。根据 GSMA 的统计，截至 2023 年，全球 5G 连接数将达到 15 亿，这一数字在 2026 年将翻一番，达到 30 亿。到 2030 年，全球手机连接数将达到 50 亿 [1]。5G 将在全球将迎来快速增长。

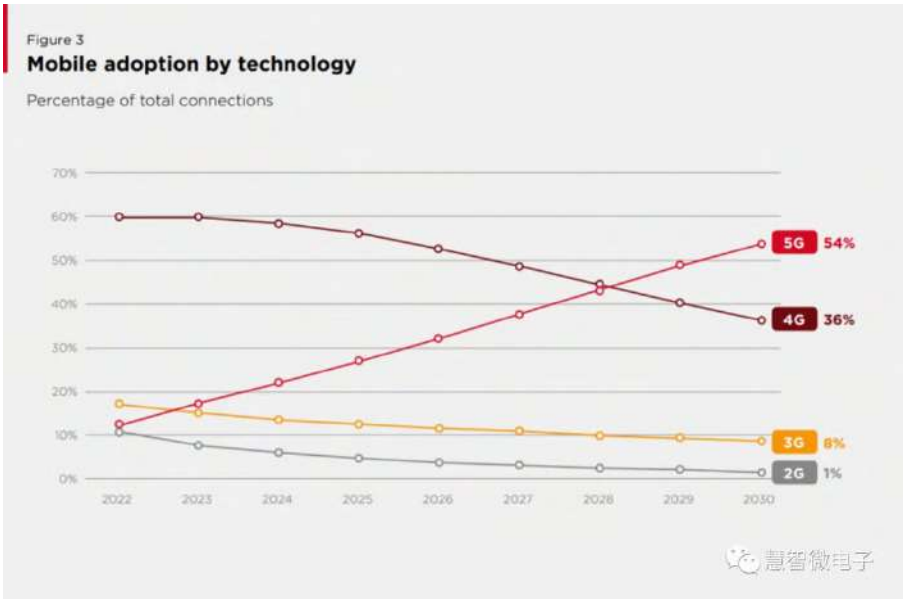


图 3.16: GSMA 所做的全球无线通信连接数预测 [

随着 5G 深度部署，5G 方案也在不断演进，以适应更广泛的网络应用和用户体验。作为对终端无线通信能力有主要影响的 5G 射频前端方案来说，也在不断寻求优化和创新。2023 年，5G 射频前端方案最明显的变化就是 Phase8 系列方案的提出和讨论了。

Phase8 方案究竟是什么？它与之前的 Phase7、Phase5N 方案有什么不同？它是未来的演进趋势吗？带着以上问题，本文就 Phase8 方案做一个梳理，与大家一起讨论近期热议之中的 Phase8 方案。

3.2.1 5G 射频前端方案演进

3.2.1.1 射频方案演进

过去 40 年，通信协议 10 年一个大节点演进，完成了从 2G 到 5G 的跨越变革。与之同时，射频前端方案也不断变化。通信协议与射频前端方案相互影响、相互促进，共同推动了无线通信业务的发展与普及。

进入 4G 时代，为了应对 4G 时代射频前端的复杂性和多样性，为了满足全球不同运营商、终端厂商、器件厂商的需求，作为芯片平台厂商的 MTK 联合终端厂商、器件厂商共同发起定义了规范化、兼容化的 4G 射频前端方案。Phase 系列射频前端从 2014 年开始定义，目前已经有 Phase2、Phase3、Phase5、Phase6、Phase6L、Phase7、Phase7L、Phase7LE 等多个版本，每个版本都根据不同的通信协议、载波聚合（CA）能力、集成度等因素进行了优化和创新。Phase 系列射频前端伴随了整个 4G 的发展，占据了整个 4G 市场约 80% 的市场份额。（有关 Phase 系列的发展，请参考慧智微文章《从 4G 到 5G，手机射频芯片十年之路》）[2]。



图 3.17: “Phase”系列射频前端的 10 年发展

Phase 系列射频前端的出现有着重要的意义，它对无线通信技术和业务的发展和普及有着积极的推动作用。Phase 系列射频前端的意义体现在：

1. 对于运营商来说，Phase 系列射频前端实现的方案归一，充分保障了运营商新协议网络部署的推进，并帮助运营商实现了终端网络建设成本及运维难度的降低；
2. 对于终端厂商来说，Phase 系列射频前端可以帮助终端厂商实现标准化的设计和生产，提高终端性能和效率，降低终端成本和风险；
3. 对于器件厂商来说，Phase 系列射频前端可以帮助器件厂快速收敛归一开发方案，提高器件的兼容性，增加器件的合作机会；

Phase 系列射频前端方案还给国内射频器件厂商与国际厂商“同台竞技”的机会，基于共同的方案定义，国内器件厂商可以快速的开发和推出符合市场需求的射频前端产品，缩短与国际厂商的差距，实现自身竞争力的提升。在过去 Phase 系列流行的 10 年里，中国射频器件厂商也在实现着自我增长与突破。

进入 5G 时代后，Phase 系列依然是公开市场的主流。自 Phase7 在 2019 年实现开门红以来，Phase7 系列方案在 2020 年及 2022 年，分别演进至 Phase7L 及 Phase7LE，实现着方案的优化迭代。

3.2.1.2 当前 5G 射频前端方案

Phase7 系列方案

尽管人们对 5G 的快速发展有所预期，但 5G 部署之迅速还是超过了许多人的预期。根据 GSMA 的数据，从发布到增长至 10 亿用户，3G 用了 11 年，4G 用了 8 年，而 5G 仅用了 4 年时间 [3]。

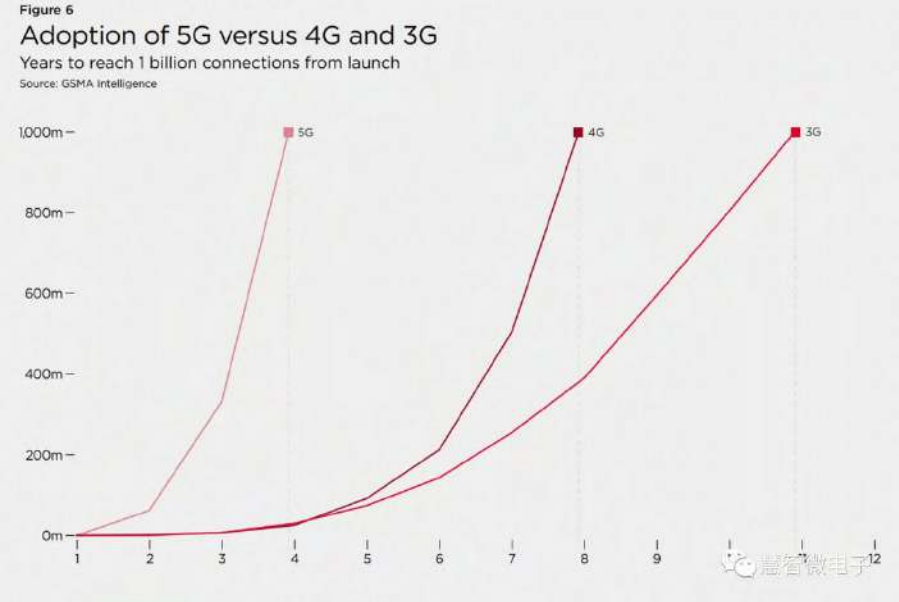


图 3.18: 增长至 10 亿用户所需时间

为了支持 5G 的快速部署，自 2018 年开始，MTK 联合器件与终端厂商，快速定义 Phase7 方案，应用于 2019 年发布的 5G 终端中。

对于平台、器件及终端厂商来说，开发新的方案存在着巨大的的风险与不确定性，更何况是 5G 这个对全球都还陌生的新兴事物。这种情况下，最稳妥的方案莫过于在原有方案上进行修改升级，最大化的复用原有方案。在这种思路下，Phase7 方案的 Sub-3GHz 部分，基本照搬了原来 4G 时代的 Phase6 方案。

2019 年商用之后，随着中国 5G 的快速部署，大家逐渐有了针对 5G 方案优化的思路。于是，在 Phase7 基础上，演进出 Phase7L（Phase7 Lite）方案，将外围主集部分的 LNA 整合进模组中；2022 年，又将 4G/5G 双连接支持特性加入进来，形成 Phase7LE（Phase7L Enhance）方案。图3.19为 Phase7 系列 Sub-3GHz 部分方案演进图示 [2]。

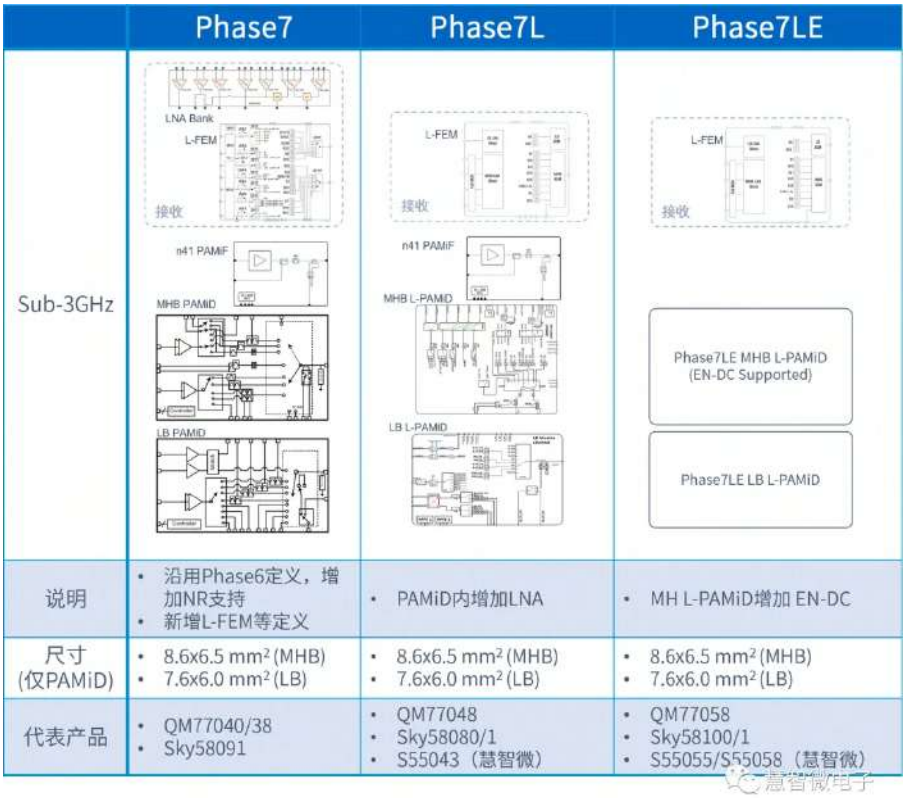


图 3.19: Phase7 系列射频方案演进

但不管怎么演进，Phase7 系列方案始终在 4G 时代已经定义好的 Phase6 方案的框架下。不论是整体功能、芯片尺寸、大体 Pin 脚位置，Phase7 始终保持 7.6x6.0 mm2 尺寸的 Low Band+GSM，与 8.6x6.5 mm2 尺寸的 Mid/High Band 的两颗 L-PAMiD 模组芯片方案。只靠功能的较小优化调整，来进行 5G 的支持。

Phase5N 方案

由于演进自 4G 时代的 PAMiD 方案，Phase7 系列方案有天生的劣势。比如成本高、冗余功能多、不够简洁。还有核心的一点，PAMiD 方案只有国际少数头部厂商可提供，也限制了方案的充分竞争，进而制约了 5G 方案的大规模商用。

在 2019 年至 2020 年初完成 5G 的第一轮部署之后，国内厂商就开始寻找 Phase7 方案之外的射频前端方案。于是 Phase5N 方案被开发出来。

与之前的 Phase 系列方案不同，Phase5N 并不是 MTK 官方定义的方案。Phase5N 方案出自终端厂商的需求：既然可以从 4G PAMiD 方案 Phase6 演进出 5G 方案 Phase7，那是否可以用 4G 分立方案的 Phase2，演进 5G 的分立方案呢？

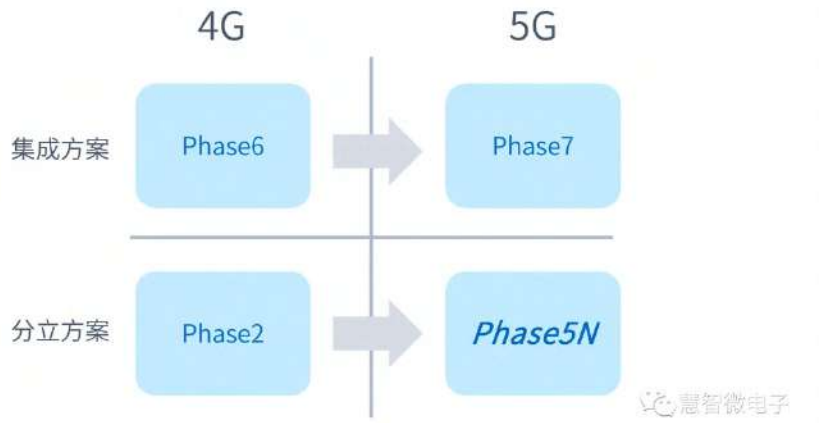


图 3.20: 5G 初期射频方案的演进思路

从接下来几年的发展来看，这个结论是肯定的。在国内终端厂商，与国内器件厂商的共同努力下，Phase5N 方案诞生了。Phase5N 方案是基于 4G Phase2 的方案，增加 5G NR 支持，所实现的射频前端方案。

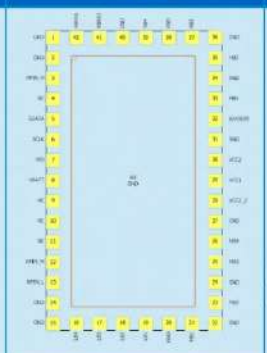
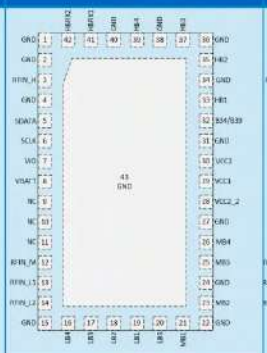
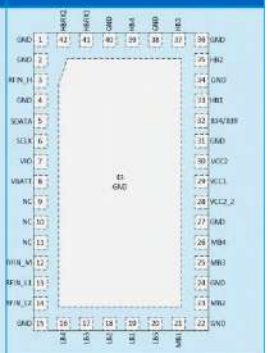
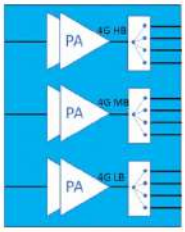
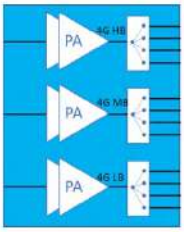
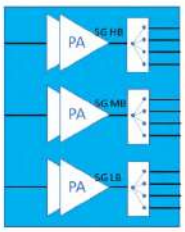
MMMB PA部分	Phase2	Phase5	Phase5N
Pin脚			
功能			
尺寸	• 4x6.8 mm²	• 4x6.8 mm²	• 4x6.8 mm²
支持模式	• 3G/4G	• 3G/4G	• 3G/4G/5G
输出功率	• PC3：27.5~28dBm	• PC3：28.5~29dBm • PC2：31~32dBm	• PC3：29~29.5dBm • PC2：31~32dBm
代表产品	• Sky77643-21(Skyworks) • S5643-51 (慧智微)	• Sky77643-61(Skyworks) • S5643-62/-66 (慧智微)	• S55643-11(慧智微)

图 3.21: 5G Phase5N 方案的出现

5G 的 Phase5N 方案继承了 4G Phase2 方案的所有特点，比如：

- 较低成本
- 较高灵活度
- 较大尺寸
- 较弱的射频性能
- 较复杂的调试难度

尽管 Phase5N 方案有诸多性能、调试的劣势，但因为足够便宜，“一白遮百丑”，Phase5N 方案广泛在对成本有诉求的终端方案中普及开来。据统计，2023 年在 2,000 元人民币以下的 5G 手机方案中，对于 Sub-3GHz 频段，90% 以上的手机采用 Phase5N 方案，而不是 MTK 所定义的 Phase7 系列方案。

Phase5N 方案经过国产射频前端厂商仔细迭代与充分竞争，从成本角度已经很难有方案与之匹敌。但在接下来 5G 演进中，Phase5N 方案仍存在问题：

- Phase5N 方案为分立方案，在支持海外 CA，以及更多 4G/5G 双连接，双卡双通等新功能时，需要用到大量分立物料，造成射频前端架构复杂；
- 以上场景下，射频前端方案需要借助大量的外围多 P 多 T 开关、多路器等进行方案实现，极大地加剧了 BOM 管控难度与管控成本；
- 方案复杂度的增加也使得方案总成本快速增加；
- 方案复杂度的增加使调试难度增加，并且增加后续性能管控风险。

从对当前使用的 5G 方案可以看到，虽然在过去 4 年时间里，5G 手机完成了快速的规模部署，但现有方案仍然是基于原有 4G 方案的框架优化而来。这些方案在 5G 应用中已经逐渐暴露出越来越多的缺点，成为 5G 未来深度部署的制约。

在这种背景下，专门为 5G 量身打造的 Phase8 方案应势而生。

3.2.2 Phase8 方案：面向 5G，应势而生

Phase8 方案是 MTK 联合器件厂商、终端厂商自 2021 年就着手定义的全新 5G 射频前端方案。在方案定义时，市面上 5G 终端已有方案可以使用，再也不需要手忙脚乱的推出一版方案了。卸下包袱的整个厂商们终于有时间，可以将 5G 的新需求和未来演进仔细的消化梳理，定义出一版真正适用于 5G 的射频前端。

另外值得一提的是，随着国内终端厂商和部分国产器件厂商技术能力和产品能力的提升，Phase8 还是国内终端厂商和器件厂商尝试合作，并参与定制的 5G 射频前端器件。

基于以上理念，Phase8 方案做了以下调整与优化：

- 重新定义 Pin 脚、尺寸，重新梳理芯片方案；
- 针对 5G 需求优化内部拓扑，将 5G 系统中所需要的 EN-DC、SRS 开关方案考虑进来；
- 针对 5G 应用场景，定义全新的通路隔离度、控制时序、开关架构；
- 去掉 4G 时代的冗余功能；
- 针对不同终端应用场景，定义 Phase8/8M/8L 方案，摆脱之前方案中的“一刀切”。

经过以上优化之后，Phase8 方案成为真正 5G 适用优化的射频前端方案。

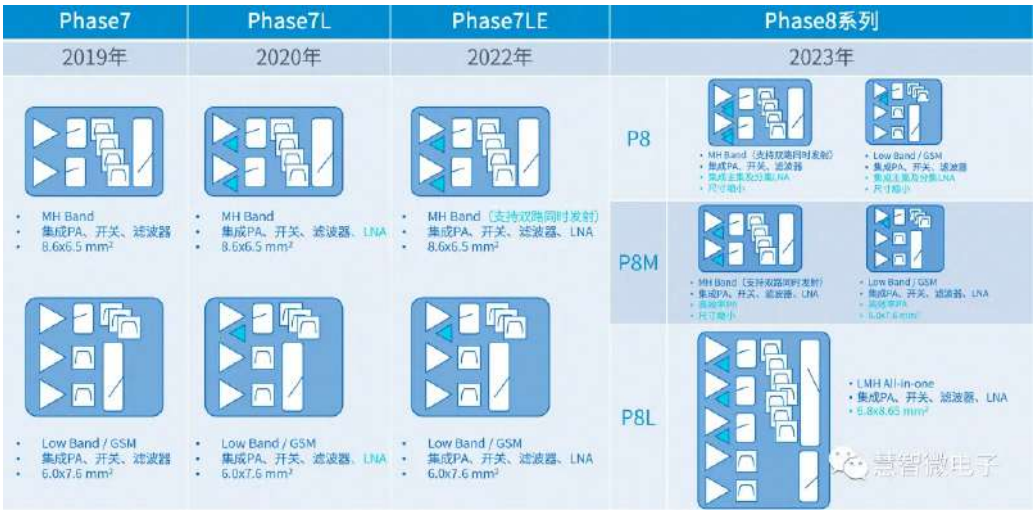


图 3.22: Sub-3GHz 部分射频方案演进

在 Phase8 系列方案中，国内器件及终端厂商需要重点关注的是 Phase8L 方案。

Phase8 与 Phase8M 方案的目标市场是高端及旗舰手机，方案强调强大的射频能力，以及完整的 CA、EN-DC 支持，采用 Low band 及 Mid/High Band 两颗 L-PAMiD 构成完整方案，并且采用如 DS-BGA 等更先进的封装，来实现更小的器件尺寸。与此相对应，成本并不是 Phase8 与 Phase8M 方案优化的主要目标。

Phase8L 方案考虑的是处于 2,000-4,000 人民币价位带手机的需求：支持合理的 5G CA 及 EN-DC 能力；采用 All-in-one 的方式进行设计，只需一颗就可以进行 Sub-3GHz 全频段覆盖。由此可以实现性能与成本的完美平衡。

在应用上，未来 5G 手机应用方案演进如图3.23所示。

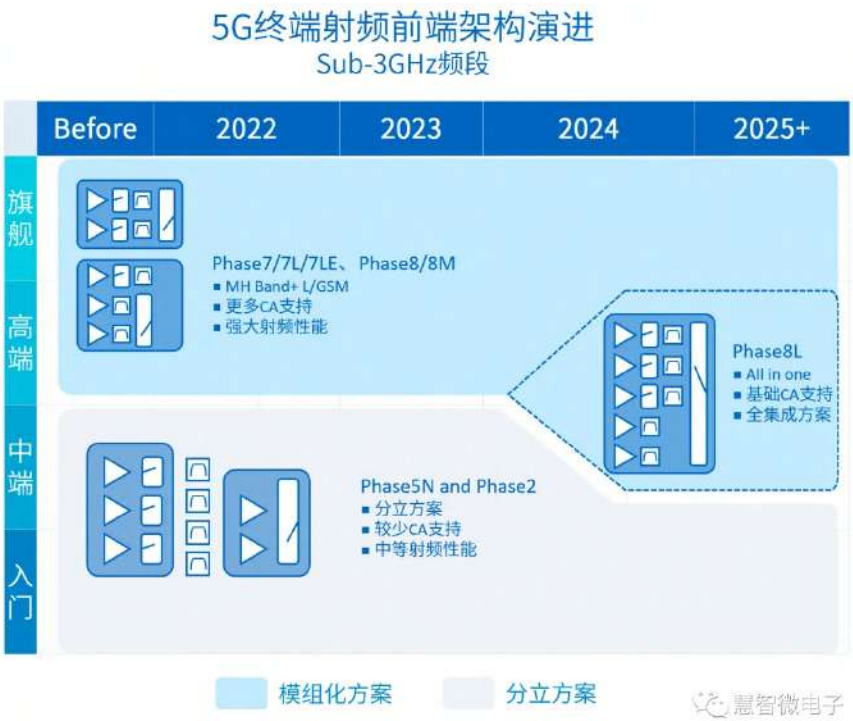


图 3.23: 5G 终端的射频方案

以下将对 Phase8L 方案进行讨论。

3.2.2.1 Phase8L 方案介绍

对于射频人来说，看到 Phase8L，自然而然想到 Phase7 系列里的 Phase7L，以及 Phase6 系列的里的 Phase6L。但 Phase8L 与 Phase8 的关系，却和 Phase7L 与 Phase7，Phase6L 与 Phase6 有所不同。

Phase7L/Phase6L 与 Phase7/Phase6 的关系，是在原版本上做演进和迭代修正，以在方案推出之后，根据实际情况实现方案的调整与优化。而 Phase8L 是 Phase8 系列中的全新定义，并不存在是在谁的基础上做“删减”。由此一来，Phase8L 就可以做针对性的定义，没有了原来的枷锁与限制。

Phase8L 方案形成以下优点：

- LMH+2G 真正的全集成 L-PAMiD 方案（All-in-one）；
- 面积是 Phase7LE 的 53%；
- 集成高性能/高效率 2G 射频通路；
- 单通道 4G/5G 性能与 Phase7LE 相当；
- 支持 M+H，L+H 双发 EN-DC，支持大多数多频 CA；
- 针对 5G 需求优化内部架构，适配 5G 手机最新需求；
- 整合周边器件，充分降低应用复杂度；
- 为未来更大规模集成奠定条件。

Phase8L 是 L-PAMiD 全集成方案

Phase8L 采用 L-PAMiD 全集成方案进行设计。

在射频前端方案中，一直以来有一个讨论：最优方案究竟是 L-PAMiD 全集成方案，还是分立方案？尤其是在分立方案竞争厂商越来越多，越来越充分，价格越来越低的情况下，大家不禁要问，L-PAMiD 方案还有必要吗？

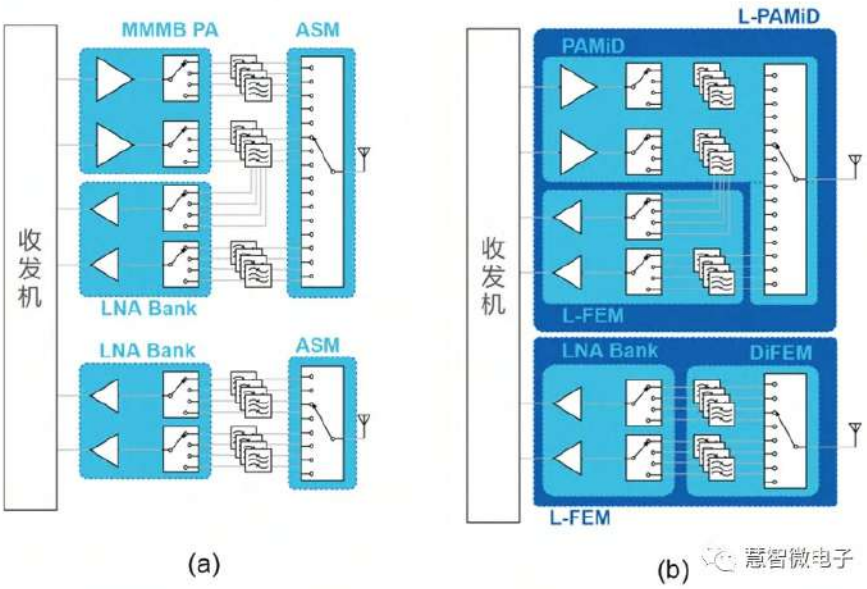


图 3.24: 分立方案（a）与模组方案（b）实现的射频前端系统

L-PAMiD 是一种集成化的方案，它将 PA、滤波器、开关、LNA 等器件集成在一个模组芯片中，实现收发通路的高集成度。相比于分立方案：

- L-PAMiD 可以减少器件数量与 PCB 面积，降低设计复杂度和成本，提高了系统的可靠性与稳定性；
- L-PAMiD 可以减少不同模块间的损耗，提升射频性能，降低射频功耗，提升用户射频体验；
- L-PAMiD 可以减少共用性硬件使用，减少分立器件的测试、生产成本，整体拥有更低制造成本。

综上所述，相比于分立方案，L-PAMiD 在集成度、性能、制造成本方面有更大优势，长远看，L-PAMiD 是终端射频方案的演进趋势 [3]。

但在市场竞争中确实可以看到，分立方案比 L-PAMiD 方案有更低的价格。造成这种现象的主要原因有：

- 分立方案可能采用更低成本的工艺来制造。比如部分分立方案采用低成本的 GaAs HBT 晶圆来制造，而 L-PAMiD 在设计之初，会选择采用较为成熟的工艺来进行设计；
- 可提供分立方案的厂商较多，竞争充分且激烈，近期甚至出现不健康竞争。造成分立方案售价过低；
- L-PAMiD 模组芯片中有功能冗余，造成 L-PAMiD 无法像分立方案快速针对应用减少冗余模块使用。

以上原因的产生，主要因为目前市面上可提供 L-PAMiD 的厂商较少。随着 L-PAMiD 供应商的增多，随着国内越来越多优秀射频前端厂商的加入，L-PAMiD 产品一定会形成优化的供应链、良好的竞争，以及应用导向的设计。L-PAMiD 方案与分立方案之间的成本差别，也将逐步趋于理性及合理。

Phase8L 采用 LMH All-in-one 设计

在 Phase6 及 Phase7 系列方案中，Sub-3GHz 一直采用 Low Band 与 Mid/High Band 分开的方案设计。

Phase6 定义是在 2016 年前后，在定义时，全球 4G 网络正在向频谱碎片化趋势演进。为了支持 4G 时代频谱碎片化需求，Phase6 在定义时，考虑到 Low Band（<1GHz）与 Mid/High Band（1-3GHz）放在同一个模组内的干扰问题，就将 Low Band 与 Mid/High Band 分成两个模组芯片。这一定义一直延续到 5G 时代 Phase7 系列方案中。



图 3.25: 全球主要地区频谱分布

进入 5G 时代后，更多 5G 新频谱被开辟出来，比如 n77 频段（3.3GHz-4.2GHz）。拥有大带宽的新频段资源后，移动终端对于频谱需求也从 4G 时代碎片化频谱 CA 需求中解脱出来，转向对重点 5G 频段，以及重点 4G/5G EN-DC 支持上。在这种思路下，将 Low Band 与 Mid/High Band 两颗芯片独立设计就变得必要性不大了。

于是，在 Phase8L 的定义中，Low Band 与 Mid/High Band 通路进行了合并，一颗芯片覆盖 Sub-3GHz 全频段。减少了方案面积，也降低了方案实现成本。

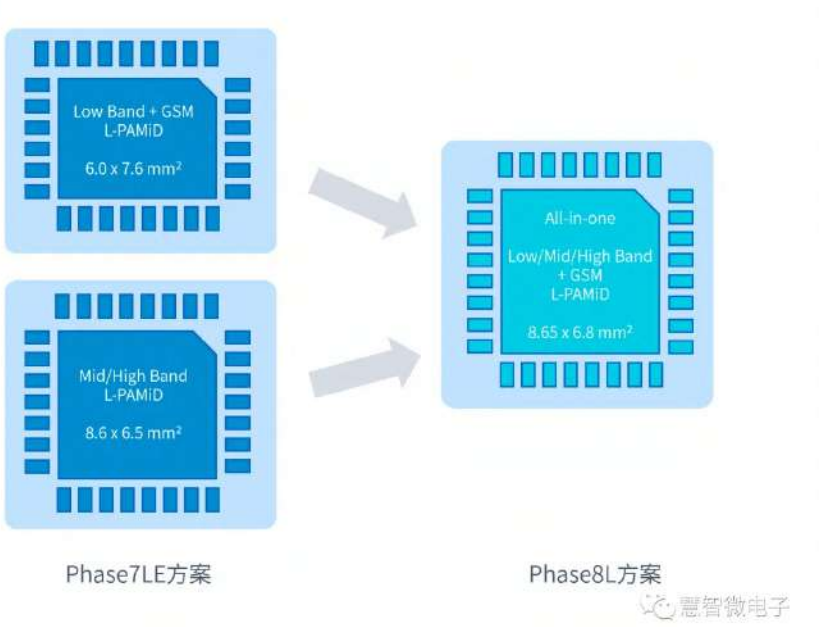


图 3.26: Phase8L 方案

Phase8L 针对 5G 需求优化了内部架构

Phase8L 方案不仅仅是将原来 Phase7 系列的两颗芯片进行功能合并，而是针对 5G 需求进行了架构优化。

在 EN-DC 支持上，Phase8L 方案支持 Mid Band 与 High Band、Low Band 与 High Band 的两通道同时发射，同一模组可支持国内的 B3 与 n41、B39 与 n41 的 EN-DC 组合，同时支持东南亚的低频与高频的 EN-DC 组合。

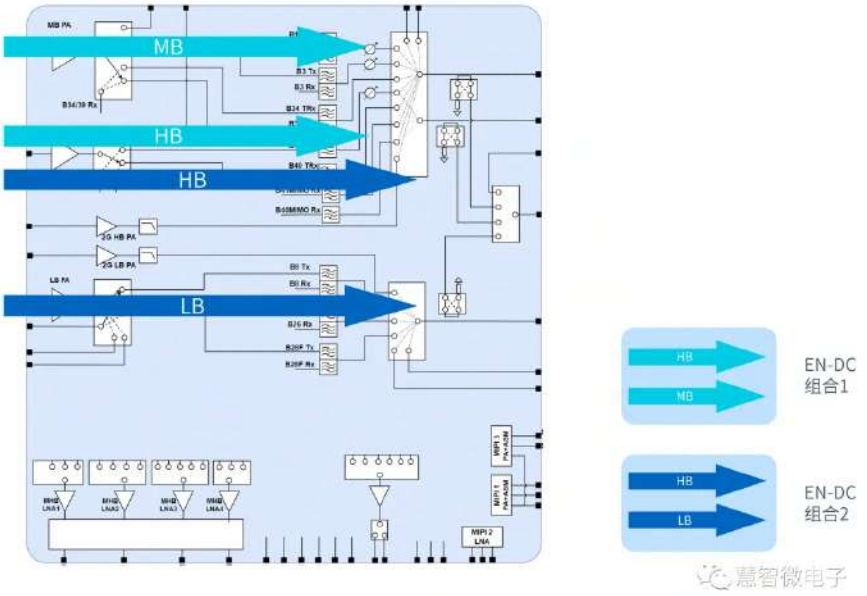


图 3.27: Phase8L 模组芯片中 EN-DC 组合的支持

对于 Low Band 与 Mid Band 的 EN-DC，Phase8L 也支持外接 Low Band 发射信号的输入，与模组内的 Mid Band 信号实现 4G/5G 双连接。

虽然已经进入 5G 时代，但 2G GSM 网络依然是终端必选支持的频段，以保障用户在 5G 网络尚不覆盖的地区使用。Phase8L 的定义中，在模组内需要支持 GSM 功能。

2G GSM 退网是最近讨论较为广泛的话题，从目前来看，较多运营商尚未有明确的退网时间表。对于终端来说，在运营商网络未完全退出前，需要保证终端 2G 功能的保留与可使用。Phase8L 对此需求做了考虑，在模组芯片内集成 2G GSM 功能。

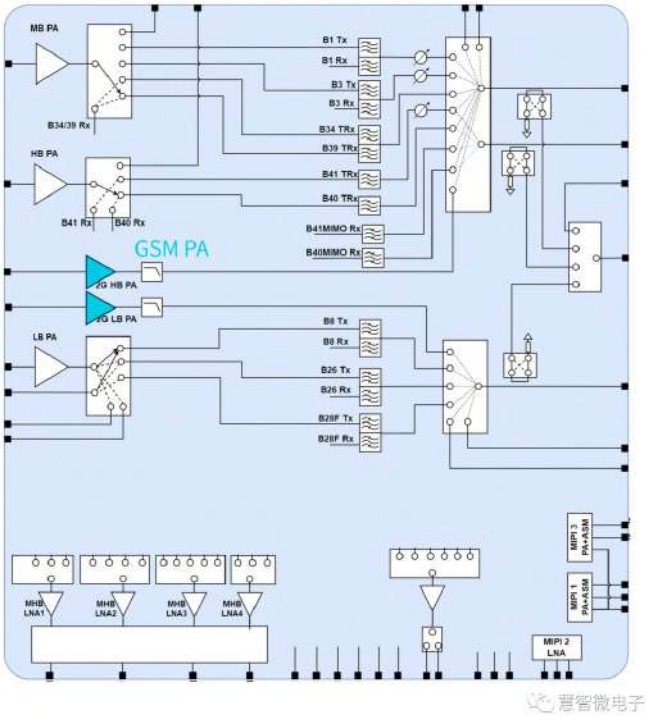


图 3.28: Phase8L 模组芯片架构

Phase8L 方案中还对开关功能进行了梳理和优化，针对 5G SRS 功能、CA 及 EN-DC 支持等功能做了重新的定义与性能优化。减少 Phase7 系列方案在应用中的干扰及 De-sense 问题。

5G 系统应用发展中，“低压应用”是降低成本的方案之一。通过设计实现 3.4V 即可工作的射频前端芯片，可摆脱射频前端对升压电源芯片的支持，减少射频前端成本。目前 Phase7 系列射频前端均需要升压电路，完成 PC2 的功率支持，部分产品甚至 PC3 功率也需要提升电源电压。在最新的 5G Phase8L 设计中，终端厂商倾向于模组芯片全频支持 3.4V 低压电源，以达到整体方案的最优设计。



图 3.29: 5G 终端的 3.4V PC2 功率输出需求

Phase8L 整合周边器件，降低方案复杂度

虽然在功能上，采用分立方案也可实现相应的 5G 方案，但方案复杂度将会呈数倍增加。

图3.30为采用 Phase8L 设计的 5G 射频前端方案，与采用分立架构设计的典型 5G 射频前端方案对比。

采用分立方案进行设计时，需要采购如 SP6T/SP2T 等开关、多工器、多 T 大功率天线开关等才能完成完整方案设计，仅 Sub-3GHz 部分 BOM 数量达到 20 颗以上，并且器件包含 PA、滤波器、发射开关、接收开关、LNA 等。对供应链管理，及商业议价提出了巨大挑战。

采用 Phase8L 方案，仅需一颗芯片就可以完成分立方案 20 颗 BOM 所完成的功能。并且实现更好的射频性能，与更简化的调试过程。实现 5G 终端最优射频解决方案。

3.2.3 Phase8L：充分拉通讨论，提供优化选项

Phase 系列射频前端方案经过了 10 多个年头，在这 10 多个年头里，国产厂商所处的位置分为三个阶段。

1. 阶段一：跟随替代阶段

   在 2019 年之前，国产射频前端厂商主要处于跟随阶段。在这个阶段，国产射频前端厂商没有定义的话语权，也没有对方案的理解与影响能力。

   在这个阶段，国内厂商所做的主要是将国际厂商的 4G 产品原样复制，再以更低的价格推向市场，在国际厂商身后，赚取微薄的利润。

2. 阶段二：并行开发阶段

   经过了长期的跟随之后，部分国内射频厂商对方案有了更深的认识，部分国内厂商也积累了自主的知识产权方案。虽然这个时候国内厂商仍然没有办法去影响平台以及终端厂商的定义，但这个时期部分射频厂商已经可以基于对方案定义的理解，快速的设计出相应的射频产品。

   此阶段代表型公司为慧智微。在 2019 年 5G 商用元年，慧智微基于自身的技术积累，与国际厂商同时同质推出 5G 新频段高集成 L-PAMiF 射频前端模组。使国内厂商首次在新协议推出时，与国际厂商同步推出新方案产品。

3. 阶段三：定义参与阶段

Phase8L L-PAMiD方案		分立方案	
BOM	数量	BOM	数量
Phase8L L-PAMiD	1	MMMB PA	1 or 2
		TxM	1
		发射滤波器及双工器	8
		接收滤波器	2
		大功率4P4T开关	1
		大功率DPDT开关	2
		大功率SP4T开关	1
		SP6T开关	1
		SP2T开关	2
		LNA Bank	1
合计	1	合计	20 or 21

慧智微电子

图 3.30: Phase8L L-PAMiD 方案与典型分立方案 BOM 对比

经过了 5G 时代的并行积累，国内射频公司、国内终端公司对于 5G 方案的演进有了更深刻的认识。在 5G 最新的 Phase8L 方案定义中，国内厂商开始参与到方案的定义讨论中来，并且参与逐渐深入。

在 Phase8L 的定义中，慧智微参与到与头部平台、头部厂商的 RFI 讨论制定中。针对对于 5G 射频前端的理解，慧智微提出了 Phase8L 相关定义的建议。相信这些建议可以使 5G Phase8L 的开发与应用更加优化。此外，针对 5G Phase8L 的应用，结合自主知识产权的射频技术方案，慧智微在 5G Phase8L 产品开发过程中也做了独特的架构及性能优化。慧智微也期待这些优化，可以使用户在使用中得到收益。



图 3.31: 图产厂商在射频前端定义中的不同阶段

由于 Phase8L 方案在定义之初就得到平台厂商、终端厂商、射频器件厂商充分讨论，结合用户的实际使用场景，在设计上、实现方案、终端应用中充分优化。Phase8L 方案的出现将给用户提供出集成度、性能与成本多赢的选项，帮助用户设计更优的 5G 终端。

3.2.4 总结

5G 商用已 4 个年头，转眼之间 5G 已经从代表未来的通信技术，变成“寻常百姓”家的通信服务。

随着 5G 的规模化应用，人们对于 5G 射频方案的理解也在不断提高，原有的射频方案要么价格太贵，要么性能太差。于是射频从业者们开始思考，除了当前的射频方案，我们真的没其他选项了吗？

Phase8 系列方案就是在这种场景下应势而生。尤其是 All-in-one 的 Phase8L 方案，它的出现将打破现有射频方案“成本”、“性能”二选一的纠结，它的出现将给出“性能”与“成本”更加完美平衡的解决方案。

5G 时代已顺利起步，但 5G 未来还依然长远。5G 所期待的万物互联、智能化/数据化的新时代还远没有到来，5G 的射频前端方案也依然需要演进。相信 5G 手机及 5G 物联网终端在 5G Phase8 和未来射频方案的加持下，会更加乘风破浪，拓展更多应用场景与行业需求，实现无线连接的无限可能。

参考文献

1. The Mobile Economy 2023，GSMA
2. 从 4G 到 5G，手机射频芯片十年之路，慧智微
3. 5G 射频前端模组的前世与今生，慧智微
4. 5G Connections Forecast, GSMA

在线阅读

本文摘自慧智微微信公众号，扫描以下二维码，在线阅读

3.3 5G 射频前端模组中的滤波器

作者：小小黑

滤波器是射频前端重要的模块之一。顾名思义，滤波器的主要功能是“滤波”，即通过有用信号，阻挡干扰信息。

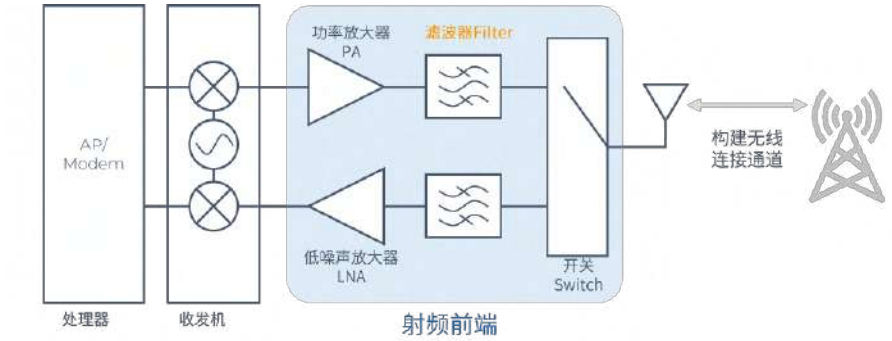


图 3.33: 射频前端中的滤波器

在射频通信系统中，“频谱”是非常宝贵且拥挤的资源。除了与我们生活息息相关的 5G、4G、Wi-Fi、GPS 及蓝牙信号外，还有通信卫星、军用卫星以及气象监测等信号。在实际生活中，无线信号无处不在，所以，就需要射频“滤波器”将无用信号处理干净。图3.34为拥挤的频谱资源示意图。

在射频前端系统中，滤波器的功能是：

- 用于发射时，滤波器可以将有用信号从众多噪声信号中过滤出来
- 用于接收时，滤波器可以有用信号之外的干扰信号过滤干净

滤波器在发射及接收通路的功能如图3.35及图3.36所示。

3.3.1 滤波器的主要衡量指标

滤波器的核心功能是通过特定频率信号，抑制带外信号，典型的滤波器频率响应示意图如图3.37。滤波器的主要指标有：

1. 中心频率
2. 带宽
3. 插损
4. 带外抑制

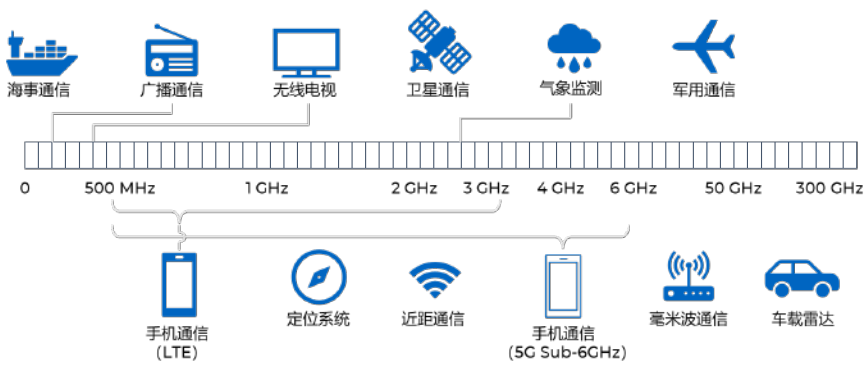


图 3.34: 拥挤的频谱资源

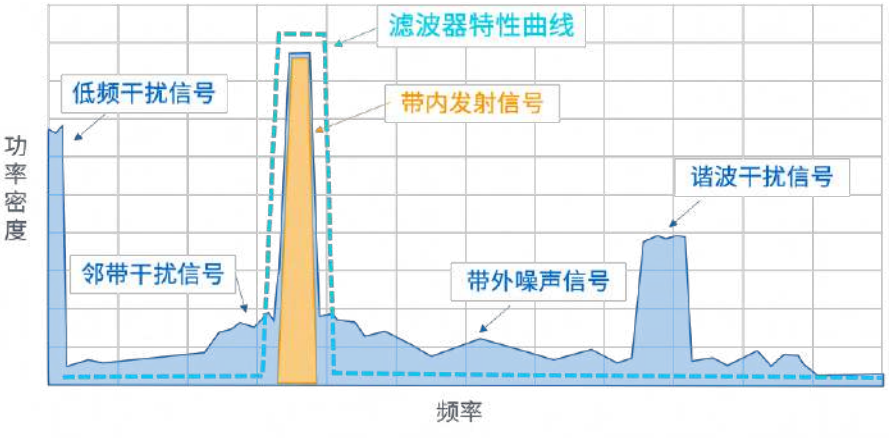


图 **3.35:** 发射通路的滤波器特性

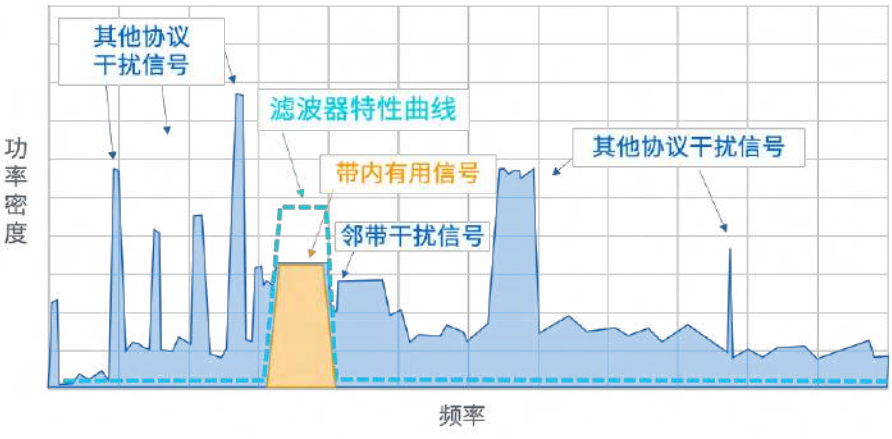


图 **3.36:** 接收通路的滤波特性

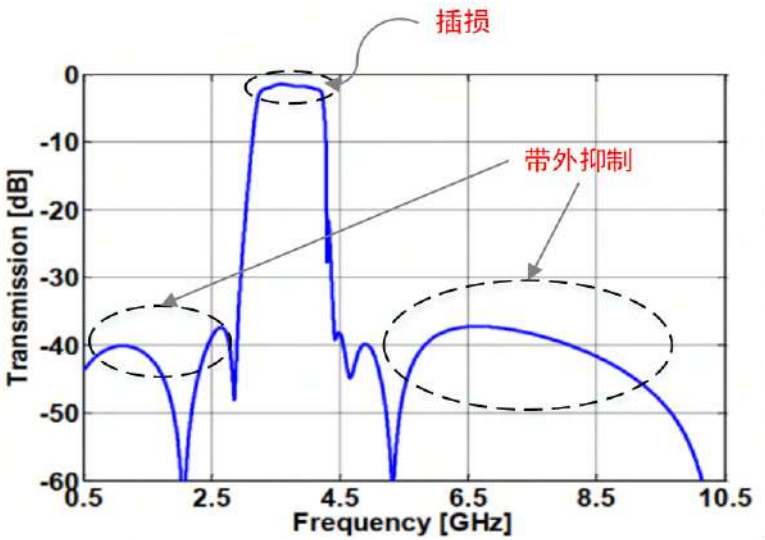


图 3.37: 典型的滤波器频率响应示意图

同时，放置在发射通路时，滤波器还需要考虑功率耐受特性以及温度特性，于是以下指标也是滤波器的主要评价指标：

1. 温度特性
2. 功率耐受

3.3.1.1 中心频率及带宽

带通 (带阻) 滤波器的中心频率是指通带 (阻带) 的中心频率，带宽是指通带 (阻带) 最高频率与最低频率之差。

滤波器作为选频器件，中心频率及带宽是其基本指标。滤波器的分类通常也是根据工作频率而分，如 B40、B3、B8、n41、n78 滤波器等，就是指工作在相应频段的滤波器。

滤波器的带宽一般由根据 3GPP 的需求而确定。比如，按照 3GPP 的规定，B3 的发射频率在 1.71-1.785GHz，接收频率在 1.805-1.880GHz，于是 B3 的发射滤波器和接收滤波器就需要分别覆盖 75MHz 的带宽。

3.3.1.2 插损

滤波器的插损是指滤波器对有用信号的衰减程度。

发射滤波器位于 PA 输出端，在 PA 输出功率相同的条件下，滤波器的插损越小，整个模块的输出功率就越大；同理，在模块发射功率相同的条件下，滤波器的插损越小，也就意味着需要 PA 输出的功率越小，PA 的功耗就越低，模块就更省电。

接收滤波器位于 LNA 之前，其插损直接决定了对接收链路噪声系数的贡献，可以通过以下噪声系数计算公式看出。插损越大，系统 NF 越大。

$$F = F_1 + \frac{F_2 - 1}{g_1} + \frac{F_3 - 1}{g_1 \cdot g_2} + \cdots + \frac{F_n - 1}{g_1 \cdot g_2 \cdot \cdots * g_{n-1}}$$

因此，其他条件不变的情况下，减小滤波器插损，可以有效提高系统接收灵敏度。

3.3.1.3 阻抗及收敛性

阻抗，即滤波器的输入输出阻抗，收敛性是指阻抗随频率的变化程度。

阻抗及收敛性直接影响滤波器的输入输出反射系数。滤波器的阻抗需要与外部电路匹配，越接近目标阻抗，对有用信号的反射越小，带内插损就越小；阻抗越收敛，带内波动就越小。图3.38为滤波器的阻抗收敛性。

此外，滤波器阻抗还会影响系统其它部分。

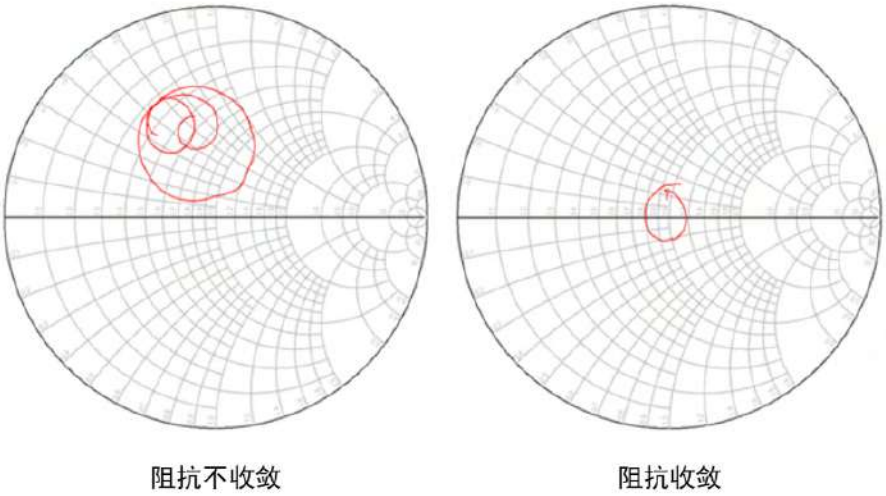


图 3.38: 反射与阻抗收敛性

如发射链路中，滤波器位于 PA 输出端，PA 设计时为了达到最大输出功率，会采用 LoadPull 方法得到最佳输出阻抗。

当滤波器阻抗偏离 50Ω 时，会造成下列问题：

1. 滤波器与匹配电路之间，以及滤波器与开关之间反射增加，链路插损增加；
2. PA 输出端的阻抗偏离最佳输出阻抗，PA 最大输出功率减小。

以上问题会导致发射链路输出功率减小，使系统性能下降。

3.3.1.4 带外抑制

带外抑制是指滤波器对通带以外的信号的衰减能力。滤波器进行带外抑制的原因如下：

发射通路

为了保证手机等终端能够正常工作，减小各频段间的干扰，3GPP 对终端发射通道的带外抑制指标提出了一系列要求，包括对谐波、互调等非线性分量的抑制，以及 CA/ENDC 场景下多频段共存的带外抑制指标等 (详见 3GPP 协议)。

由于 PA 本身对带外的抑制能力不足，所以需要在输出端加滤波器使最终的带外抑制指标符合 3GPP 规范。

接收通路

射频系统在接收时，来自外界或自身产生的干扰信号会导致接收端出现 desense，即灵敏度恶化。

为了避免或减小 desense，就需要滤波器对那些能够引入带内噪声的干扰进行抑制，具体要求就反映在带外抑制的指标上。

图3.39给出了滤波器需要考虑的一系列带外抑制点，具体抑制度需根据系统指标要求分析。

温漂特性

滤波器的工作频率会随温度的变化产生漂移，该特性称为温漂。

温漂特性会使滤波器通带产生偏移，会影响通带的插损，尤其是通带边缘位置的插损，以及带外抑制度，如图3.40所示。

图3.40中，在同样的 spec 要求下，图 (a) 温漂较大，通带和阻带的一部分在高低温下已经不满足指标的要求。经过温度补偿处理后，图 (b) 温漂较小，高低温下均满足指标的要求。在滤波器设计时，需要尽量减小温漂，适当拓宽通带的范围，减小过渡带范围，保证偏移后的通带仍能覆盖目标频段，带外抑制仍能满足指标要求。

	Tx滤波器	Rx滤波器	说明
带外抑制频段	Rx	Tx	Tx/Rx隔离需求
	Other Rx	Other Tx	其它共存频段Tx/Rx隔离
	-	Rx-Tx	可能产生互调干扰的频点
	-	(Tx+Rx)/2	
	-	Tx+Rx	
	-	2*Tx-Rx	
	-	2*Tx+Rx	
	2*Tx	-	Tx谐波
	3*Tx	-	
	4*Tx	-	
	2.4G~2.4835G	2.4G~2.4835G	2.4G ISM
	5.15G~5.85G	5.15G~5.85G	5G ISM
	1.164G~1.189G	-	GPS L5/BD B2a频点
	1.559G~1.608G	-	GPS L2/BD B1c频点

图 3.39: 滤波器带外抑制指标需求

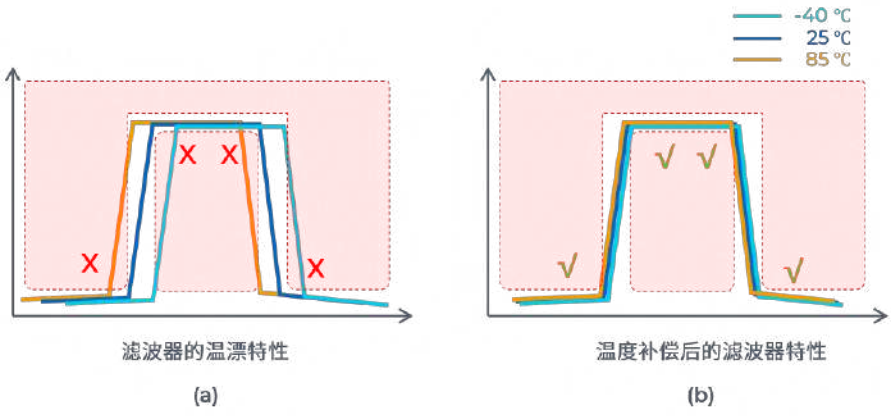


图 3.40: 滤波器温漂特性示意图

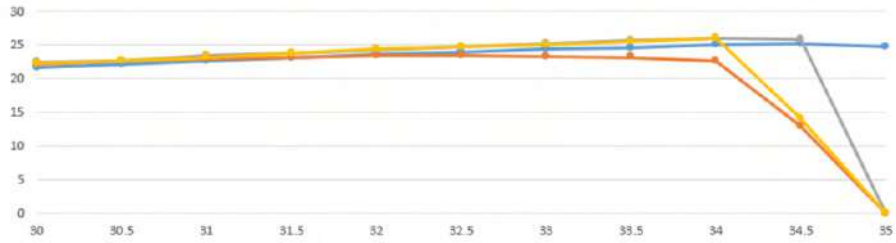


图 3.41: 不同滤波器滤波器功率耐受曲线

衡量滤波器温漂特性的指标是频率温度系数，即 TCF，其计算方法如下：

$$\mathrm{TCF}(ppm/^{\circ}C) = \frac{1}{f(T_0)} \cdot \frac{f(T_1) - f(T_0)}{T_1 - T_0} \times 1,000,000$$

其中，T_0(℃) 和 T_1(℃) 分别为两个温度点，$f(T_0)(MHz)$ 表示温度 T0 时对应的频率，，$f(T_1)(MHz)$ 表示温度为 T_1 时对应的频率，二者之差为频率偏移量。

计算 TCF 时，一般取通带左右两边固定插损 (如-20dB 或-30dB) 的频率点，计算其在不同温度下的偏移量。

对于 SAW 滤波器等声学滤波器，在不经过温补处理时，其温度系数一般为负值，即高温时整体频率向低频偏移，在使用时需要加以注意。

功率耐受

功率耐受指标表明滤波器对大功率的承受能力。

Tx 滤波器位于 PA 输出端，需要承受 PA 输出的大功率信号，因此需要对 Tx 滤波器进行功率耐受指标评估。

小功率条件下，滤波器输出功率随着输入功率的增加而线性增加，当输入功率到达功率耐受临界点时，滤波器输出功率达到最大，进一步增加输入功率，滤波器将发生暂时或永久损坏，此时插损会大幅增加，输出功率大幅下降。图3.41为不同型号滤波器的功率耐受测试曲线。其中三颗滤波器耐受功率在 34-34.5dBm。

在功率耐受测试中，需要重点关注的是通带高端的功率耐受性能。

因为大功率条件下，滤波器温度升高，通带向低频偏移，造成通带高端插损变大。插损越大，温度越高，会导致通带进一步向低频偏移，从而使通带高端插损更大，温度进一步上升。形成正反馈，进而导致器件损坏。

3.3.2 5G 射频前端模组中的滤波器

目前 5G 大规模商用的主要是 6GHz 以下频段。在这部分频段中，根据频率的不同，分为 Sub-3GHz 与 Sub-6GHz 两部分。

Sub-3GHz 频率位于 3GHz 频率以下，是原来 4G LTE 的频段的升级重新使用，所以又叫重耕（Re-farming）频段。这部分频段的特点是频谱众多、带宽较窄、较多 FDD 频段，需要对信号进行精准过滤才能够满足正常通信需求。

5G Sub-6GHz 一般指 6GHz 以下、3GHz 以上的新频段的部分，目前最主要的频段有 n77（包含 n78）、n79 两个频段，这部分频段带宽宽、旁边基本无干扰频段，并且是 TDD 频段，不需要考虑发射及接收之间的干扰，可以减轻对滤波器带外抑制的需求。不同频段对滤波器的需求如图3.42所示。

5G 射频前端模组中用到的滤波器主要分类如图3.43所示。

LC 滤波器是基于电感/电容的频率响应特性来进行滤波器设计，压电滤波器则利用材料的压电特性进行设计。在特性上，二者最大的区别就是带外抑制的区别。压电滤波器可以做到陡峭的带外抑制，适合于频谱拥挤、对 T/Rx 抑制有需要的 FDD 频段。

典型的 LC 滤波器与压电滤波器的频率响应对比如图3.44所示。以 TF-SAW 滤波器为例，其在 ±5% 的带宽处即可达到 30dB 以上的带外抑制，而 LC 滤波器中的 LTCC 滤波器的带外滚降要平缓很多。不过，LC 的滤波器可以做到大的通带带宽，这一点是压电滤波器较难做到的。

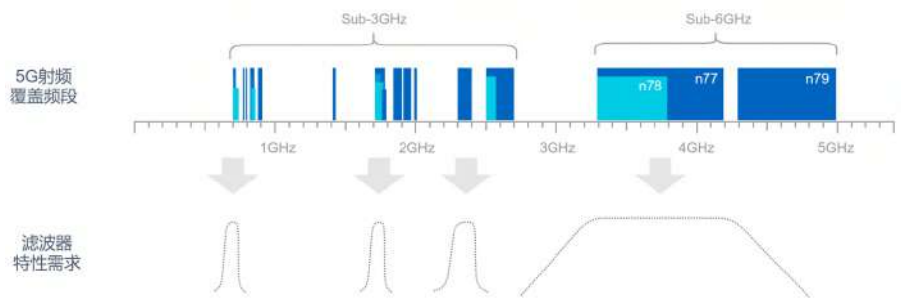


图 3.42: 5G 射频前端的频率覆盖，及对滤波器的特性需求

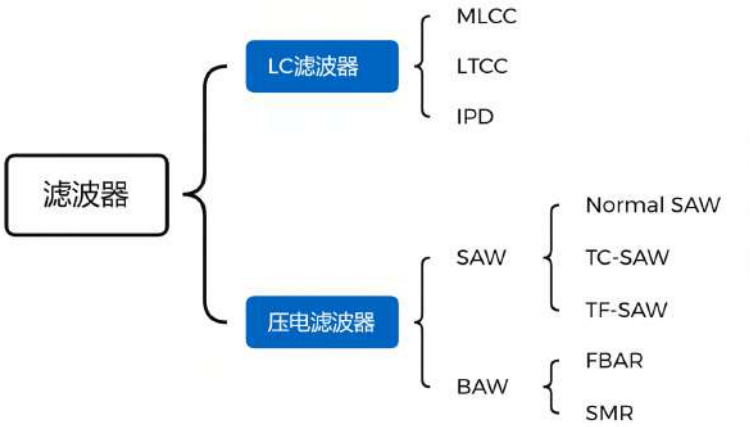


图 3.43: 5G 射频前端模组的滤波器分类

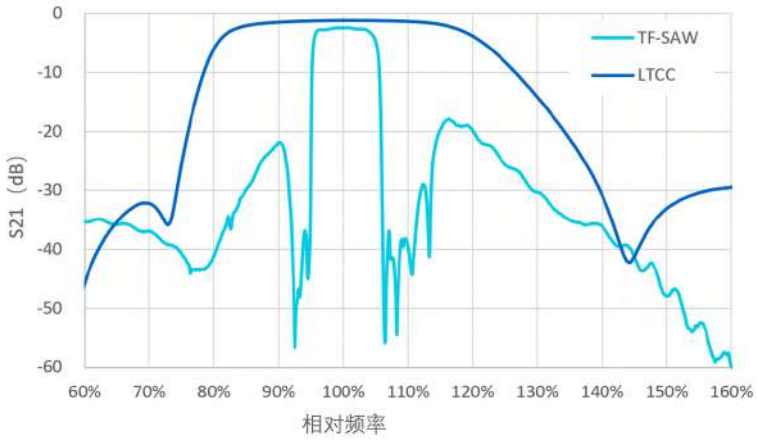


图 3.44: LTCC 与压电滤波器特性对比

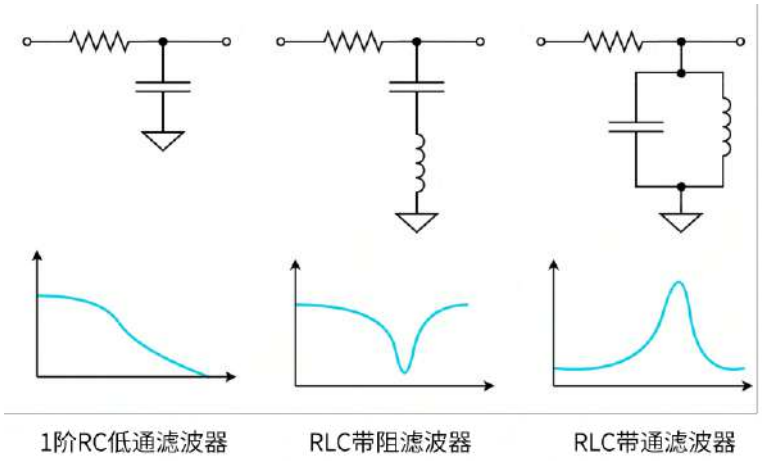


图 3.45: 典型 LC 滤波器频率响应

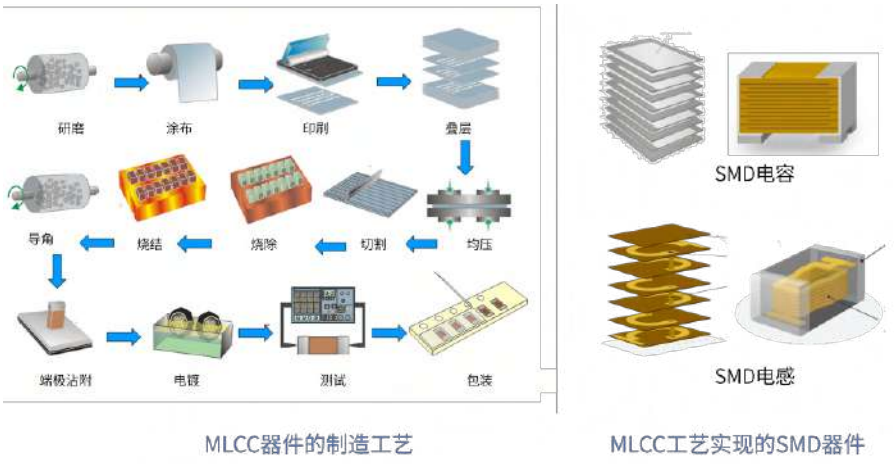


图 3.46: MLCC 工艺流程图示

3.3.2.1 LC 滤波器

LC 滤波器是基于电感/电容的频响来设计的滤波器。常见的电感/电容电路频率响应特性如图3.45所示。

利用多级滤波器的设计方法，就可以设计出给定带宽、中心频率以及带内波动的滤波器。在物理实现上，可以采用分立 SMD 器件、LTCC 以及 IPD 等工艺实现

分立 SMD 器件滤波器

采用分立 SMD 器件法是指将目标电容值与电感值以分立 SMD 的方式进行实现。这种实现方式的优点是设计灵活，易于调试；缺点是所占面积大，寄生效应影响明显。

分立 SMD 器件的制造工艺一般采用 MLCC 工艺，即多层陶瓷工艺。是将陶瓷浆粉通过研磨、涂布、叠层等工艺，在陶瓷衬底上，实现分立的 SMD 器件。图3.46为 MLCC 器件制造工艺图示。

LTCC 滤波器

LTCC 工艺的中文名称是低温共烧陶瓷工艺（Low Temperature Co-fired Ceramic ），也是利用陶瓷烧结技术进行多层电路基板设计。

与 LTCC 对应的工艺还有 HTCC（高温共烧陶瓷工艺，High Temperature Co-fired Ceramic），但因为 LTCC 工艺烧结温度低，只有 1000 度以下，这样就可以采用导电率高，但熔点较低的金属如金、银、铜等金属作为导体材料，所以 LTCC 工艺在高频、高速电路中有广泛的应用。

LTCC 工艺相比于 MLCC 也有很大的相似之处，都是利用陶瓷和导体材料共同烧结形成的物理器件。但

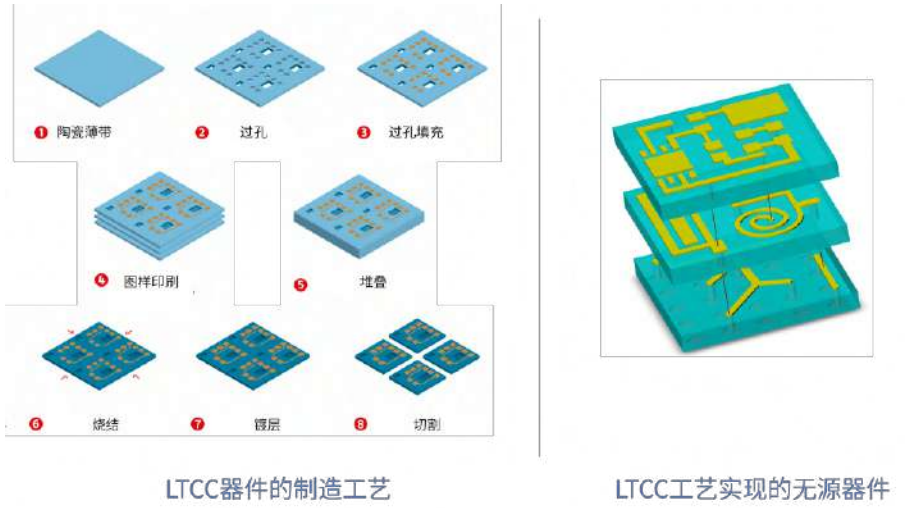


图 3.47: LTCC 工艺流程图示

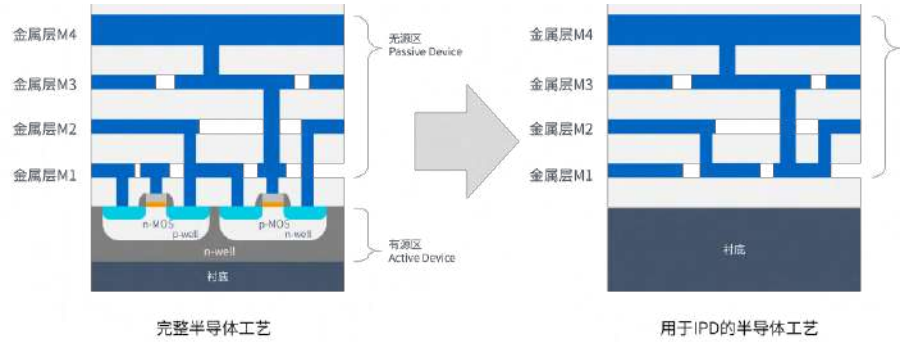


图 3.48: 完整半导体工艺 IPD 工艺关系示意

LTCC 工艺可以一次性的集成更多的电感、电容器件，所以适合设计制造如滤波器、集成基板等较为复杂的无源器件。图3.47为 LTCC 工艺流程。

正是由于以上特性，LTCC 工艺在 LC 滤波器中也得到了广泛应用。

IPD 滤波器

IPD 的全称是集成无源器件（Integrated Passive Devices）。由于名称中有“集成”（Integrated）二字，所以必定和“集成电路”（Integrated Circuits）有很强的联系。

完整的集成电路器件既包含有源器件（Active Devices），也包含无源器件（Passive Devices），同样利用半导体工艺，将其做最大化的简化，把有源器件舍弃，只做无源器件，就形成了 IPD（集成无源器件）。IPD 工艺与完整的半导体工艺对比如图3.48所示。

IPD 的优势是可以承接半导体工艺中的优势，即数量越大，平均单颗成本越低。但也因为需要采用半导体的工艺，在小批量时，不易获得成本优势。

另外，半导体工艺不容易实现较厚的金属，一般 IPD 器件的金属厚度在 um 级别，较薄的金属层也会对其 Q 值造成限制。

3.3.2.2 压电滤波器

压电滤波器（Piezoelectric Filter）是利用材料的压电效应所设计的滤波器。

压电材料的特性是可以将电信号转化为机械信号。射频中常用的压电效应是将电信号转化为机械信号中的弹性波信号，在机械信号中进行处理后，再转化为电信号输出。由于所使用的弹性波位于声波频率范围内，所以这种器件又叫声波器件（Acoustic Wave Device）。其工作原理如图3.49所示。

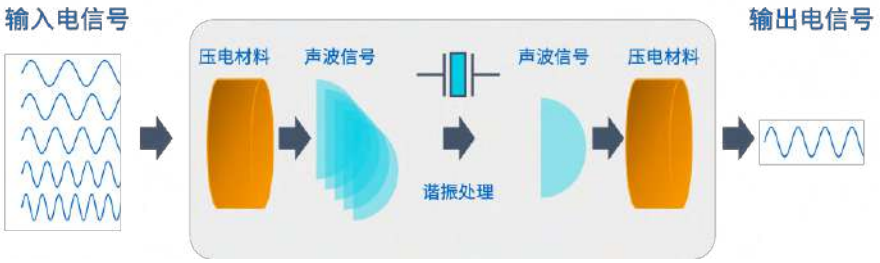


图 3.49: 声波器件工作原理

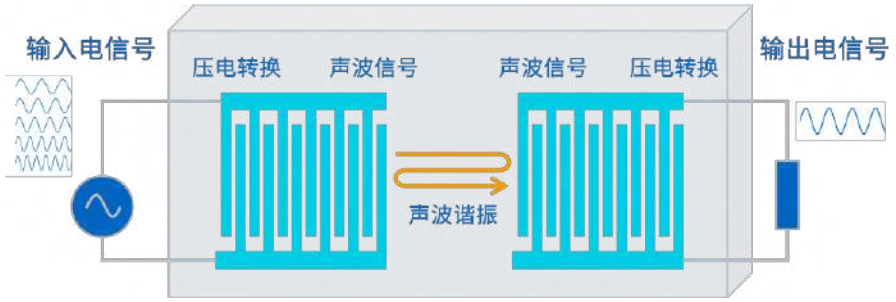


图 3.50: SAW 滤波器工作原理示意图

在声波器件中，最为常见的是 SAW 和 BAW 两种器件，两者全称分别为声表面波器件，和体声波器件（Surface Acoustic Wave，Bulk Acoustic Wave）。

从 SAW 和 BAW 的名字可以看出，二者都是利用了声学特性设计的器件，不过一种是使声波在表面传输，一种是使声波在“体”内传输。

图3.50及图3.51分别为 SAW 和 BAW 器件的工作原理示意图。

压电滤波器的优势是可以利用声学器件极高的 Q 值，设计出窄带高抑制、低插损的滤波器。缺点是必须要用到压电材料，与集成电路中的半导体工艺不兼容。并且工艺敏感，对设计和制造工艺提出了高的要求。

5G 射频前端模组中不同滤波器的对比如3.52所示。

3.3.2.3 5G 射频模组中的滤波器集成

不管是对于 IPD、LTCC 设计的 LC 滤波器，还是对于 SAW、BAW 技术设计的压电滤波器，都与射频前端中的 PA、LNA 以及开关设计中的 GaAs、SOI、CMOS 等半导体工艺不兼容，在模组设计中，采用系统级封装（SiP，System in Package）方式来集成实现模组集成。图3.53为 5G 射频前端模组的工艺实现。

在 5G 集成模组中，除了对上述滤波器的主要衡量指标有要求外，还对滤波器尺寸、二次封装能力等提出了更高要求。由于封装进模组之后无法进行二次调试和替换，5G 集成模组中滤波器还要求有高的一致性和可靠性。



图 3.51: BAW 滤波器工作原理示意图

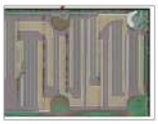
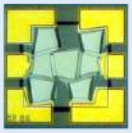
滤波器类型	特点	类型细分	实现方式	实物照片
LC滤波器	1.宽带 2.易实现 3.无法形成窄带高抑制	SMD分立	分立器件搭建	
		LTCC	LTCC一体实现	
		IPD	半导体工艺实现	
压电滤波器	1.窄带 2.声学器件实现 3.可形成窄带高抑制	SAW (包含TC-SAW、TF-FAW等)	声表面波器件实现	
		BAW (包含FBAR、SMR等)	体声波器件实现	

图 3.52: 5G 模组中不同滤波器的对比

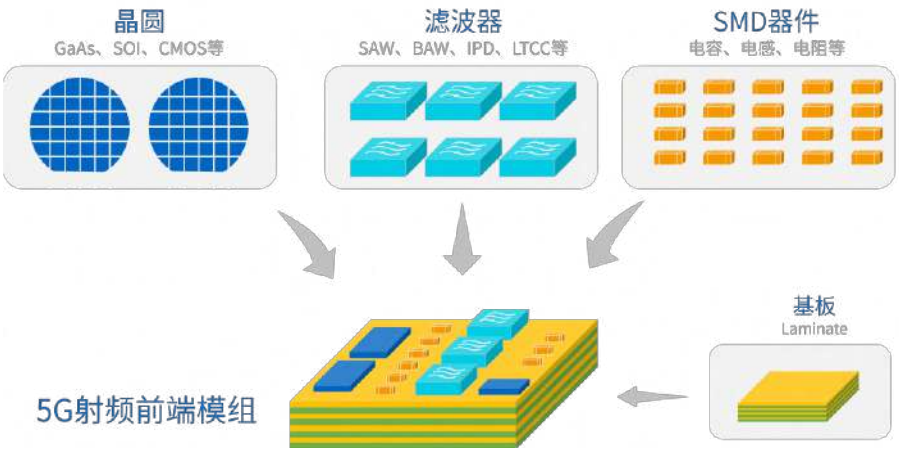


图 3.53: 5G 射频前端模组的工艺实现

3.3.3 总结

滤波器是射频前端中的重要器件，负责完成信号的“滤波”功能。在滤波器评估中，频率和带宽、插损、抑制、温度特性和功率特性是重要的评价指标。

5G 移动终端中常用的滤波器有 LC 滤波器和压电滤波器，分别在宽带、对抑制要求不高的场景，以及窄带、对抑制有高要求的场景中使用。

在 5G 高集成模组化演进中，滤波器也被不断集成进射频前端模组中来。在集成过程中，除了考虑常规滤波器指标外，还需要考虑滤波器的尺寸、二次封装能力、可靠性等指标。

在滤波器使用中，您觉得还有哪些需要注意的技术要点？欢迎扫描文章末尾二维码，在线留言讨论。

参考文献

1. https://www.murata.com/
2. https://commons.wikimedia.org/wiki/File:MLCC-Manufacturing-Process.png
3. https://www.minicaps.com/ltcc3d.html
4. Razavi, B. . ”RF Microelectronics, 2nd Edition.” (2012).
5. 朱晓明等著. 硬件十万个为什么-无源器件. 西安: 北京大学出版社
6. https://www.jcetglobal.com/uploads/pdf/20200110154805_87333.pdf
7. https://ewh.ieee.org/r5/denver/sscs/2005_03_Ruby.html
8. https://www.reverse-costing.com/teardown-notes/qualcomm-ultrasaw-technology-soitec-poi-substrate-/

在线阅读

本文摘自慧智微微信公众号，扫描以下二维码，在线阅读

3.4 5G L-PAMiD 的前世与今生

作者：彭洋洋 博士

最近十几年中，射频前端方案快速演进。“模组化”是射频前端演进的重要方向。射频前端的“模组化”究竟是什么，它是怎么来的，又有什么挑战？

带着以上问题，本文对射频前端模组的发展过程做一个梳理，对射频前端产品模组化进程中的挑战和未来可能的演进做一个讨论。

3.4.1 射频前端的模组化是什么？

射频前端是指天线后，收发机之前的部分。射频前端主要有 PA（功率放大器）、Switch（开关）、LNA（低噪声放大器）及 Filter（滤波器）构成。

射频前端的模组化方案（Integrated Solution）与分立方案（Discrete Solution）相对应。发射通路中的模组化是指将 PA 与 Switch 及滤波器（或双工器）做集成，构成 PAMiD 等方案；接收通路的模组化是指将接收 LNA 和开关，与接收滤波器集成，构成 L-FEM 等方案。模组化方案与分立方案的区别如图3.55所示。

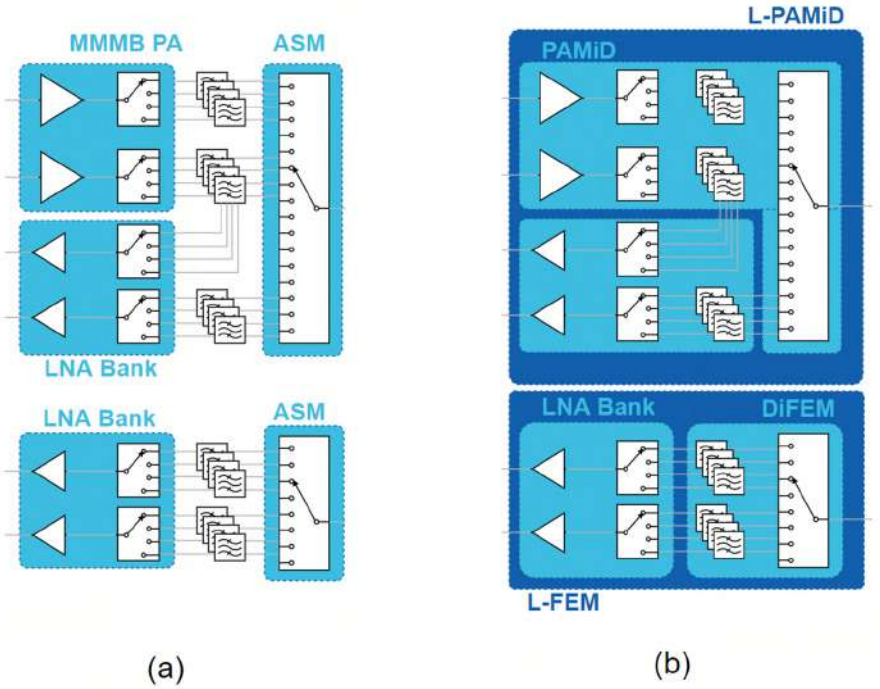
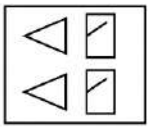
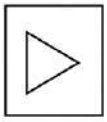
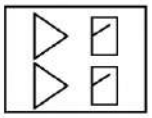
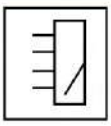
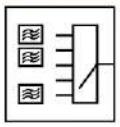
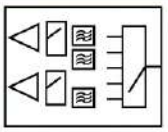


图 3.55: 分立方案（a）与模组方案（b）实现的射频前端系统

根据模组内集成器件的不同，射频前端模组也有不同的名称。常见的模组名称及集成的器件如图3.56所示。

在 3G 及 4G 的早期时代，手机需要覆盖的频段不多，射频前端一般采用分立方案。到了 4G 多频多模时代，手机需要众多器件才能满足全球频段的支持需求，射频前端也变的越来越复杂；同时，分立方案在一定程度上无法满足高集成度、高性能的需求，集成模组方案得到了规模化采用。目前，iPhone 中已经全面采用模组化方案，根据拆机分析网站 eWisetech 的拆机分析，在 2020 年至 2021 年华为、小米、OPPO、vivo、荣耀等多个厂商发布的手机中，处于 1500 至 2000 人民币价位带的多款手机已采用模组化方案 [1]。

缩写	LNA	LNA bank	PA	PAM
全称	Low noise amplifier		Power amplifier	PA module
集成子模块	LNA	多个LNA	PA	PA，开关
示意框图				

缩写	SW	ASM	DiFEM	L-FEM
全称	Switch	Antenna switch module	Diversity Front-end Module	LNA Front-end Module
集成子模块	开关	天线开关	开关，接收滤波器	开关，接收滤波器，LNA
示意框图				

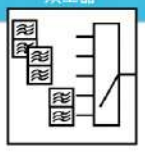
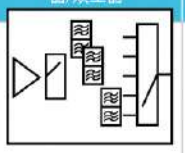
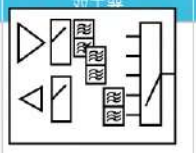
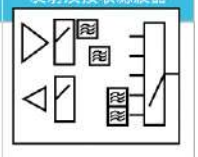
缩写	FEMiD	PAMiD	L-PAMiD	L-PAMiF
全称	Front-end Module integrated Duplexer	PA Module integrated Duplexer	LNA PA Module integrated Duplexer	LNA PA Module integrated Filter
集成子模块	开关，发射滤波器/双工器	PA，开关，发射滤波器/双工器	PA，LNA，开关，发射及接收滤波器/双工器	PA，LNA，开关，发射及接收滤波器
示意框图				

图 3.56: 不同射频前端简写及集成子模块

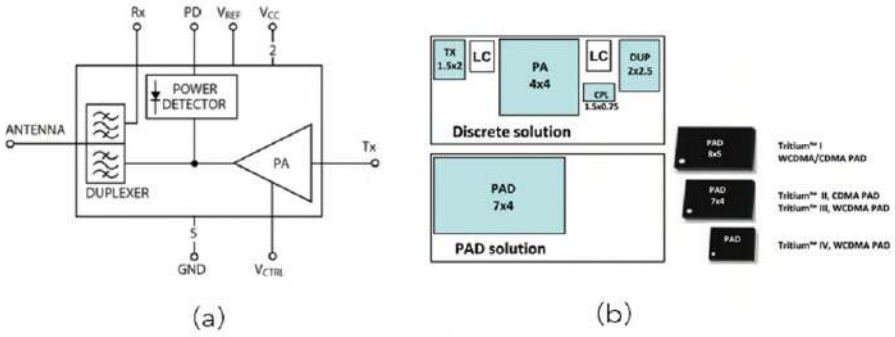


图 3.57: （a）Triquint 于 2003 推出的模组产品 TQM71312 （b）Triquint 对其模组产品的说明



图 3.58: iPhone 3G 所采用的 Triquint PAMiD 模组

3.4.2 5G 射频前端模组的前世

3.4.2.1 2000-2009 年：先驱者的尝试，PAMiD 萌芽的 10 年

射频前模组方案中，最具代表性的就是发射通路的 PAMiD 模组。PAMiD 是 PA Module integrated with Duplexer 的缩写，早期也被称为 PAD，是集成了 PA、开关与滤波器的模组。

最早的 PAMiD 可追溯到 2000 年初，两家先驱型射频前端公司 Triquint 及 Agilent 看到集成模组化带来高集成、高性能及低成本优势，开始做集成模组化的尝试，两家公司均实现了开创性的工作。

Triquint 是当时领先的 CDMA 射频前端供应商，在并购了滤波器厂商 Sawtek 后，Li, P., Souchuns, C., 和 Henderson, G. 于 2001 年左右开始模组化产品 TQM71312 的研发。2003 年，Microwave Journal 报道了该产品的工作，指出模组化设计将带来高性能、高集成度、小尺寸及高易用性，取得了 40% 的平均电流降低 [2]。这是行业内第一个公开发布和报道的集成模组产品，在后续行业综述中，这项工作被引用为集成模组产品的开端。如图3.57所示。

在报道中，Triquint 的集成模组产品系列命名是 Tritium^{TM}。功不唐捐，先驱者的付出并没有白费。苹果公司在 2008 年推出的首款支持 3G 的 iPhone 手机 iPhone 3G 中，首次采用了模组方案。而 iPhone 3G 中用于支持 3G 信号的射频前端就是 Triquint Tritium^{TM} III 系列模组芯片 [4]，如图3.58所示。Triquint2014 年与 RFMD 公司合并成立 Qorvo 公司，Triquint 在集成模组的优势，在 Qorvo 时代依然延续。

值得一提的是，当年 Triquint 参与业界首款开创性集成模组的 3 名设计人员中，有 2 位今天依然活跃在业界一线，引领和推动着行业发展，对工程师来讲射频行业实在是一个事业常青的领域。

关注到 PAMiD 的另外一家公司是 Agilent。Agilent 是有悠久历史和传承的射频前端厂商，源于 HP。Agilent 于 2001 年开始实现 FBAR 滤波器的量产，到了 2002 年，实现了千万级出货 [5]，将自己的射频 PA 产品与滤波器产品做整合变成了顺理成章的选择。AFEM-7731 是 Agilent 于 2005 年推出的 CDMA PAD 产品。与 Triquint 公司的 TQM71312 类似，AFEM-7731 内部集成一路 CDMA PA 及一个双工器，如图3.59所示。得益于 FBAR 的低插损，Agilent 表示 AFEM-7731 可以取得优秀的线性和效率性能 [6]。

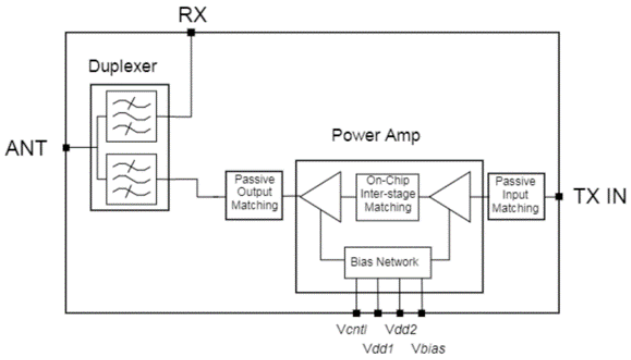


图 3.59: Agilent 于 2005 年推出的 CDMA 集成模组产品 AFEM-7731



图 3.60: iPhone5 (A1429 型号) 射频前端方案，采用模组化方案进行设计

或许是看到射频前端巨大的发展前景，2005 年 12 月 12 日，Agilent 的射频前端部分从 Agilent 独立出来，成立新公司 Avago，成为当时最大的非上市独立半导体公司，并于 2009 年上市。2016 年，Avago 与 Broadcom 合并，新公司更名为 Broadcom。

尽管 Avago 具有 FBAR 技术带来的滤波器性能优势，但在 2000 年初，它的射频功率放大器处于弱势，集成模组产品的进展并不尽如人意。直到 2010 年左右，基于新工艺和新功率合成架构的射频功率放大器获得性能优势，进而带动了集成模组产品的成功。2012 年起，Avago 在 PAMiD 的产品及之后的 Broadcom 公司的射频前端模组产品，被大量应用于 iPhone 系列手机中。

3.4.2.2 2010-2019：国际厂商推动，模组方案主流化的 10 年

苹果的引领

2010 年，苹果推出 iPhone4 手机，单款机型销量超过 5,000 万部，是当时最成功的 iPhone 手机。从 2010 年开始，苹果公司开始对智能手机的全面引领。在 iPhone4 手机中，依然采用 Triquint Tritium™ 系列 PAMiD 方案实现 3G 射频前端。

在 2012 年发布的首款支持 4G 的 iPhone 手机 iPhone5 中，iPhone 采用了 Triquint、Avago 及 Skyworks 的模组化产品 [7]。苹果继续坚定的采用模组化方案，如图3.60所示。

在这一时期射频前端供应商在模组化也进行了坚决的投入。为了实现模组化中模块的优势整合，一系列射频前端公司也进行了合并。如图3.61所示。

- 2014 年，RFMD 宣布与 Triquint 合并，成为 Qorvo 公司。
- 2014 年，Skyworks 与松下成立合资公司，2016 年 Skyworks 将合资公司全资收入旗下。
- 2017 年，高通宣布与 TDK 成立合资公司 RF360，2019 年高通将合资公司合资收入旗下。



图 3.61: 射频前端公司的整合



图 3.62: Skyworks 与 Qorvo 向公开市场推出 PAMiD 方案

除了在苹果手机中使用的定制化射频前端模组方案，各个射频前端供应商开始将模组化产品推向公开市场。Skyworks 在 2014 年推出 SkyOne® 方案，Qorvo 也在 2014 年推出 RF Fusion™ 方案。Skyworks 在对 SkyOne® 方案的介绍中指出：“SkyOne® 是首款将多频功率放大器及多掷开关同所有相关滤波、双工通信及控制功能整合在一个单一、超集成封装当中的半导体设备，所用空间还不到行业最先进技术的一半”[8]。如图3.62所示。

FEMiD：模组化的另外一种选择

虽然 PAMiD 模组化方案有诸多的性能优势，但其供应劣势也相对明显：厂商必须要同时掌握有源（PA 及 LNA，Switch）及无源（SAW、BAW 或 FBAR）等能力，才有办法设计出 PAMiD 模组。而同时掌握这些资源的厂商只有 Skyworks、Qorvo、Broadcom 及 Qualcomm 等少数具有完整资源的厂商。

于是，华为、三星等终端公司着手推动 FEMiD（Front-end Module integrated with Duplexer）方案。FEMiD 是将天线开关及滤波器整合为一个模组，交由滤波器公司提供；PA 依然采用分立方案，由 PA 公司提供。这种方案有效的发挥了无源公司与有源公司的特长。华为、三星等终端也因此摆脱了对 PAMiD 厂商的绝对依赖。两种方案的对比如图3.63所示。

2016 年，PAMiD 与 FEMiD 的主要供应商如图3.64所示。Broadcom、Skyworks 及 Qorvo 是主要的 PAMiD 供应商，村田和 RF360 是主要的 FEMiD 供应商 [7]。

3.4.3 5G 射频前端模组的今生

3.4.3.1 Phase6/7 系列 PAMiD 方案的归一

不过，与 iPhone 中模组化方案的绝对主流相比，早期公开市场的模组化方案推广并不顺利。原因是 Skyworks 与 Qorvo 各自定义，所推广的方案并不兼容，在技术上和供应上都给平台适配和客户使用造成困扰。

为了解决方案统一的问题，MTK 平台、国内头部手机厂商及 Skyworks/Qorvo 射频前端厂商联合发起 Phase6 系列射频前端集成方案定义。在 Phase6 方案中，Low Band (包括 2G) 与 Mid/HighBand 两颗 PAMiD 构成完整发射方案。如图3.65所示。

由于方案归一，并且终端厂商、平台厂商及芯片厂商联合参与定义，Phase6 系列方案自 2016 年推出后，得

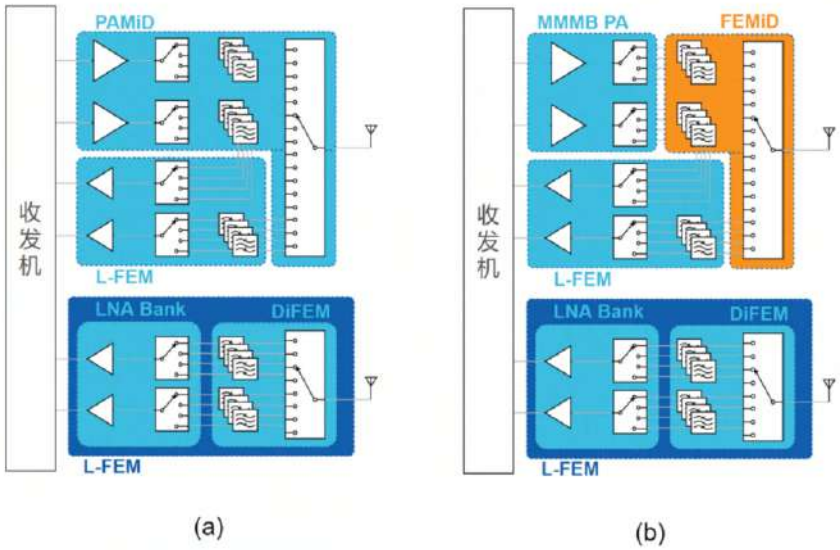


图 3.63: PAMiD 与 FEMiD 方案对比（a）PAMiD 方案（b）FEMiD 方案

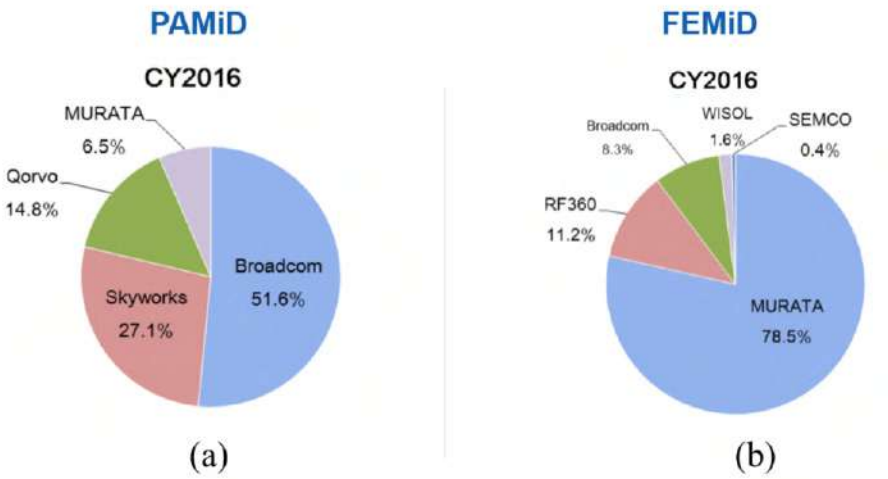


图 3.64: PAMiD 与 FEMiD 主要供应商

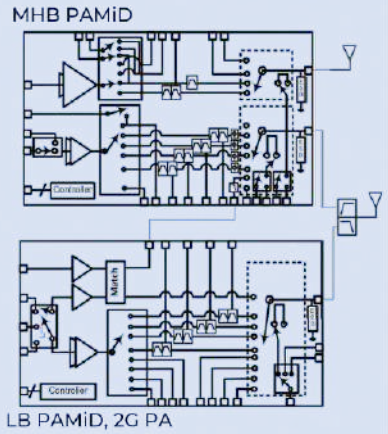
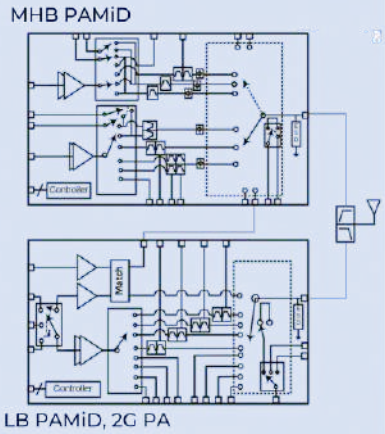
	Phase6	Phase6L
框图		
尺寸	• 8.6x6.5 mm² (MHB) • 7.6x6.0 mm² (LB)	• 8.6x6.5 mm² (MHB) • 7.6x6.0 mm² (LB)
集成器件	• PA、滤波器、双工器、ASM开关等	• PA、滤波器、双工器、ASM开关等
支持功能	• 2G GSM • 4G LTE Cat 16 • DL 4CC, UL 2CC • 3 Ant • ET	• 2G GSM • 4G LTE Cat13 • DL 3CC, UL 2CC • 2 Ant • APT
代表产品	• QM77031 (MHB) • Sky78185-11 (LB)	• QM77031 (MHB) • Sky78185-21 (LB)

图 3.65: Phase6 与 Phase6L 方案的定义

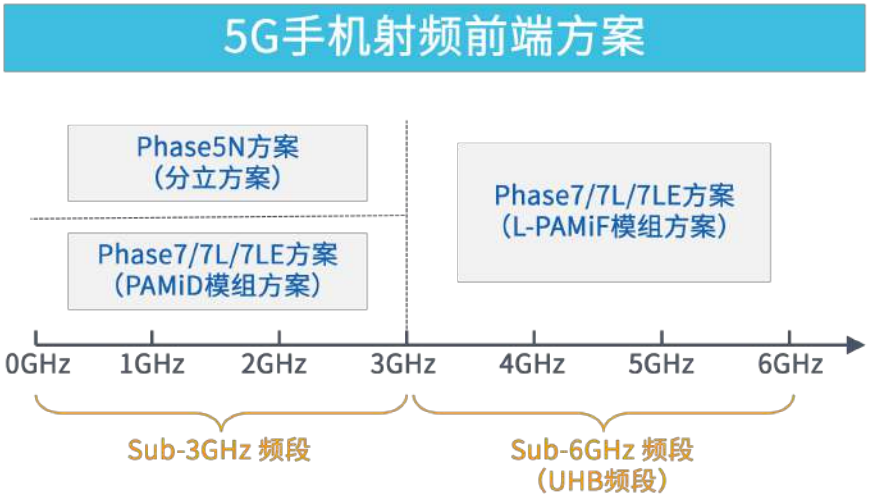


图 3.66: 5G 手机射频前端方案

到华为、小米、OPPO 及 vivo 等手机厂商认可。在对于性能及集成度有高要求的高端手机中得到使用，模组化方案得到了普及。5G 到来之后，Phase6 系列方案演进至 Phase7/7L，依然维持 PAMiD 模组化定义。

3.4.3.2 2020 至之后：国产开始形成突破

随着 2019 年底运营商 5G 陆续商用，2020 年 5G 元年正式开启。5G 到来之后，手机终端需要支持更多的频段。并且 5G 定义了 3GHz 以上，6GHz 以下的超高频（UHB，Ultra-High band）频段，对射频前端性能提出了更高要求。

经过两年的方案迭代，5G 方案已基本收敛。主要分为 Phase7 系列方案及 Phase5N 两种方案。两种方案在 Sub-6GHz UHB 新频段部分方案相同，均为 L-PAMiF 集成模组方案；在 Sub-3GHz 频段分别为 PAMiD 模组方案和 Phase5N 分立方案。如图3.66所示。

Sub-6GHz UHB 频段 L-PAMiF：国产已成熟商用

Sub-6GHz UHB 频段为 5G 新增频段，频率高、功率大，且增加 SRS 切换等复杂功能，集成 LNA、PA、滤波器、收发开关及 SRS 开关的 L-PAMiF 成为主流选择。

在 Sub-6GHz UHB L-PAMiF 产品中，国产厂商逐渐形成突破。2019 年 12 月，在中国 5G 正式商用的 2 个月之后，慧智微 n77/78/79 双频 L-PAMiF S55255-11 量产，国产射频前端厂商首次与国际厂商同时同质推出产品。S55255 系列产品在 2020 及 2021 年陆续被国际头部手机终端所采用。2020 年，该产品获得“中国芯”重大创新突破产品奖，2021 年慧智微获得 GTI 年度荣誉奖（2021 GTI Honorary Award）。

2021 年，国产其他射频前端厂商陆续推出 UHB L-PAMiF 产品，在未来演进中，国产 UHB L-PAMiF 产品会越来越有竞争力。5G UHB 主要产品如图3.67所示。

Sub-3GHz 频段：国产亟待突破

相比于 Sub-6GHz，虽然 Sub-3GHz 模组频率更低、功率更低，不需要复杂的 SRS 开关等，但由于 Sub-3GHz 频段较多，需要集成的滤波器及双工器更多，并且是 SAW、BAW 及 FBAR 等声学滤波器，对滤波器资源的获取、多频段的系统设计能力提出了高的要求。不同产品集成滤波器技术及数目如图3.68所示。

对于 Sub-3GHz PAMiD/L-PAMiD 模组产品设计，主要的挑战有：

- 全模块子电路的设计和量产能力
  需要射频前端厂商有模块内每个主要电路的成熟设计及产品化能力，如各频段的 PA、LNA 及开关等，并且各子模块无性能短板。
- 强大的系统设计能力

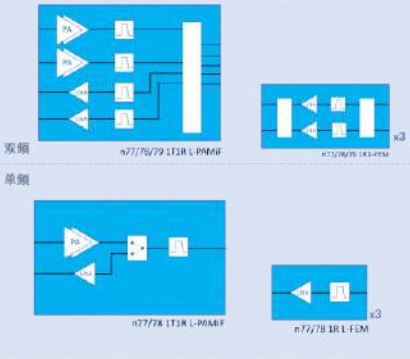
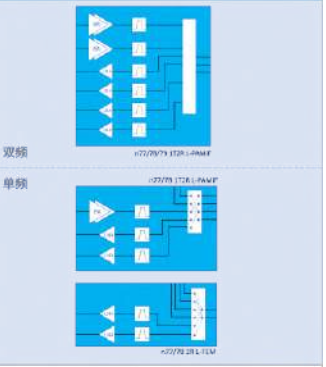
	Phase7/7L	Phase7LE
Sub-6GHz UHB n77/78/79 L-PAMiF及 L-FEM		
功能	1T1R	1T2R
代表产品	• 双频: • S55255+S15728 （慧智微） • Sky58255+Sky53728 • 单频 • S55207+S15720 （慧智微） • QM78207+Murata-507	• 双频: • S55281 （慧智微） • Sky58281 • 单频 • S55217+S15722 （慧智微） • Sky58277+Sky53759

图 3.67: 5G UHB 主要产品

	Sub-6GHz 频段		Sub-3GHz频段	
	单频	双频	Low-band	Mid/High-band
集成发射滤波器数目	1	2	4	8
集成滤波器频段	n77/78	n77/78 n79	Band8 Band12 Band20 Band26	Band1 Band25(2) Band3 Band34 Band39 Band40 Band7 Band41
集成滤波器技术	LTCC或IPD	LTCC或IPD	SAW、BAW或FBAR	SAW、BAW或FBAR

图 3.68: 不同产品集成滤波器技术及数目

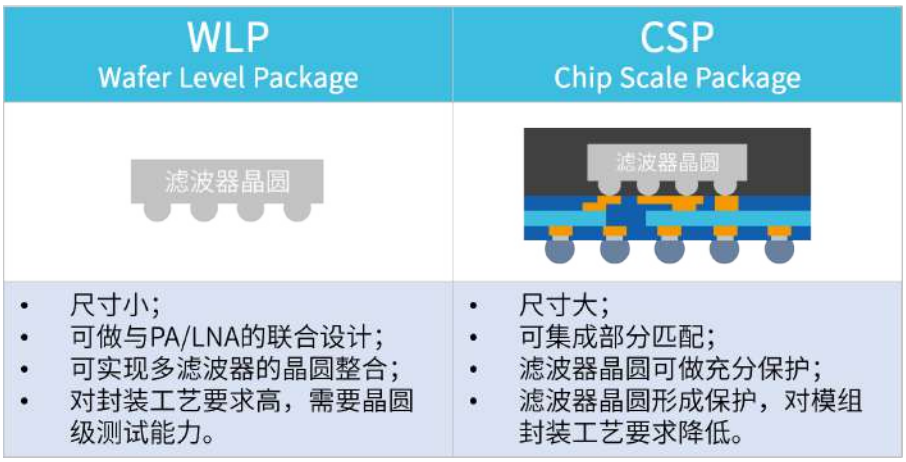


图 3.69: WLP 与 CSP 两种封装结构下的滤波器比较

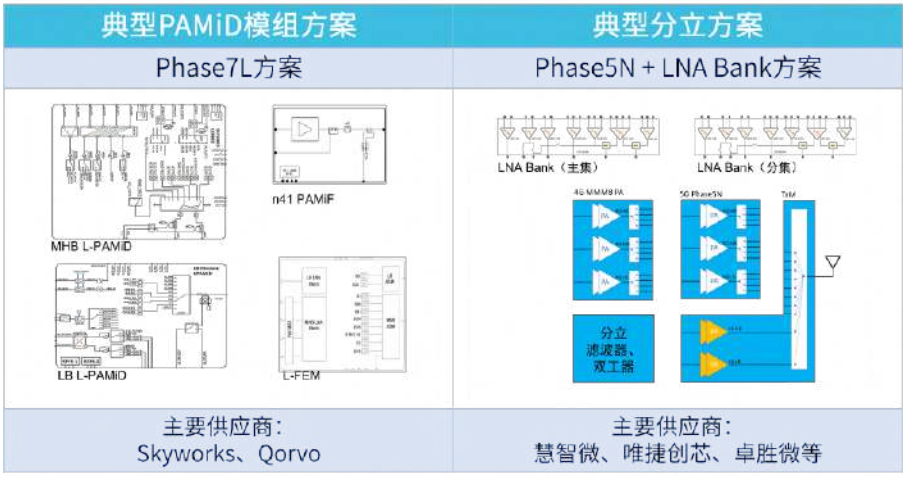


图 3.70: Sub-3GHz 典型的模组方案与分立方案比较

全集成模组本身构成一个复杂的系统，涉及到发射与接收之间隔离、各频段之间的抑制及载波聚合的通路设计等等问题。射频前端不再是一个单独的功能模块，需要厂商有强大的系统分析与设计能力。

- 小型化滤波器资源

  小型化可集成的滤波器资源是模组设计的稀缺资源，目前在 Sub-3GHz 用到的主要是 WLP（Wafer Level Package，晶圆级封装）或 CSP（Chip Scale Package，芯片级封装）两种封装结构的滤波器。两种滤波器的比较如图3.69所示。WLP 滤波器尺寸小、与模组内其他模块的设计中有优势，是未来模组内滤波器的发展方向。

以上能力的同时具备是设计 Sub-3GHz 模组产品的必要条件，也是国内射频前端厂商面临的挑战。在国内厂商对以上挑战未完全实现突破的情况下，国内厂商在 Sub-3GHz 只能提供分立方案。目前 Sub-3GHz 集成与分立方案的比较如图3.70。

虽然有诸多问题和挑战，射频前端模组仍是国内射频前端厂商必须攻克的产品类别。慧智微在这个产品方向上较早进行了积累，并形成了系列成果。慧智微射频前端的 PA/LNA 模块采用可重构架构，在模组产品设计中有以下优势：

- 集成度更高，晶圆更少，有助于兼容较大尺寸的滤波器。
- 具有软件调谐特性，有助于集成之后的二次适配。
- 多代产品持续积累，经过验证的封装管控能力。

3.4.4 总结

“模组化”是射频前端演进的重要方向，在这个过程中，滤波器厂商与模组厂商都面临巨大的挑战和机遇。慧智微期待用技术创新解决模组化产品设计中的问题，为客户带来更优的射频前端解决方案。

本文整理过程中，得到众多同行、同事的帮助，在此一并表示感谢。

参考文献

1. https://www.ewisetech.com
2. module. Microwave Journal, pp.10-15.
3. Wang, D., Wang, X.,Li, P., Tang, A., Souchuns, C., Chen, W., McNamara, B., Wang, X., Liu, L.,Apel, T. and Pavio, R., 2004, October. Recent advances in CDMA power amplifier module developments. In Proceedings. 7th International Conference on Solid-State and Integrated Circuits Technology, 2004. (Vol. 2, pp.1236-1241). IEEE.
4. https://www.ifixit.com/Teardown/iPhone+3G+ Teardown/600
5. Ruby, R., 2004, March. FBAR-from technology development to production. In Second International Symposiumon Acoustic Wave Devices for Future Mobile Communication Systems (p.5).
6. AFEM-7731Product Information, Operating Instructions.AFEM-7731 Datasheet.
7. RF Devices / Modules For Cellular, Navian, 2017.
8. Skyworks Product Introduction, http://sige.com/downloads/literature/CorporateBrochure_Simplified_Chinese.pdf
9. Qorvo Product Introduction, http://qorvo.com
10. Qualcomm Product Introduction, http://www.qualcomm.com
11. Mobile Experts. RF front ends (RFFE) for mobile devices 2021. Mobile Experts, 2021.
12. SmarterMicro 推出可重构全频段多频多模功率放大器, 简化 4G 手机设计. 中国集成电路, 23(11):2, 2014.
13. Li,P.,Peng, Y., and Li,Y.. Overview of development in reconfigurable multi-mode multi-band power amplifier design. In 2019 IEEE MTT-S International Wireless Symposium (IWS), pages 1–3. IEEE, 2019

在线阅读

本文摘自慧智微微信公众号，扫描以下二维码，在线阅读

3.5 尺寸缩小 40%：国产工规级 5G n77/n79 L-PAMiF 领先量产

作者：小慧，发布时间：2023 年 4 月 20 日

近年来，5G 发展如火如荼。5G 的全球大规模应用也对射频前端芯片提出了更多挑战，比如需要支持更多频段，需要更小尺寸，需要兼容支持多国家和地区的方案。

针对于 5G 全球大规模应用中的挑战。**慧智微 5G 产品持续领先演进，慧智微推出全新小尺寸、高集成 n77(78)/n79 双频工规级 L-PAMiF 产品** S55235。目前，S55235 产品在终端客户项目完成完整 5G 系统验证，产品成功实现量产。

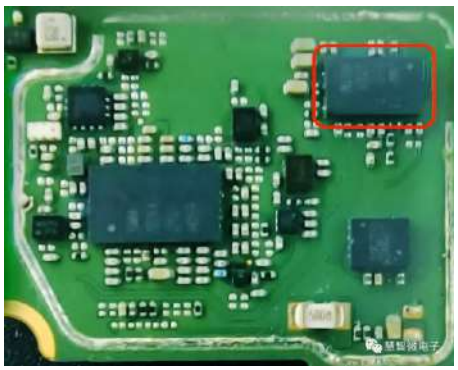


图 3.72: S55235 产品系统验证照片

S55235 产品实现的功能有：

- 支持 5G n77(78)/n79 双频 1T2R（1 路发射、2 路接收）
- PA、LNA、滤波器、SRS 开关、耦合器全集成
- 支持高通、MTK、展锐等各大 5G 平台

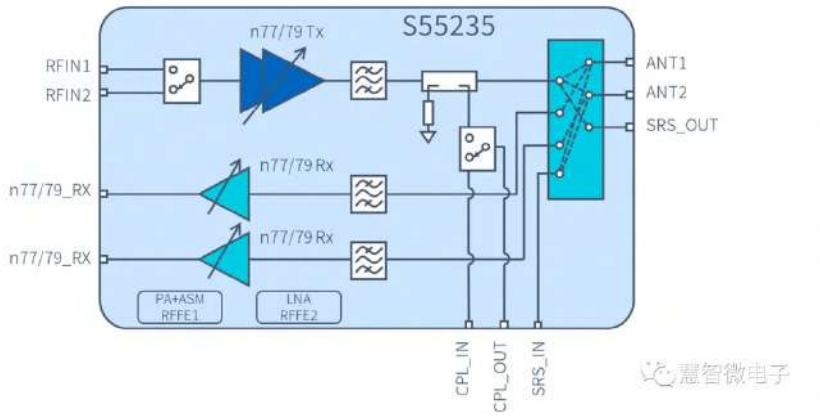


图 3.73: S55235 功能框图

在此基础上，S55235 实现以下重要突破：

- 尺寸由原来的 4.5mm x 5.5 mm，缩小至 3mm x 5mm，面积缩小 40%
- Pin 脚与 n77(78) 单频 1T2R 产品兼容，方便兼容设计
- 支持 3.4V 低压 PC2 功率等级，不再需要额外升压电源
- -40 到 85 度工规可靠性支持，满足工业需求

基于以上特性，S55235 可以在以下场景中实现优势应用：

- 需要兼容 n77 及 n77(78)/n79 设计的场景
- 需要实现小型化、全频段的 5G 应用

- 低压 3.4V 实现 5G PC2 高功率的应用场景



图 3.74: S55235 产品优势应用场景

S55235 产品还是首个国内产业合作厂商成功自定义的 5G 高集成模组芯片。国内终端厂商、平台厂商以及芯片设计厂商，在定义和设计、量产过程中紧密合作讨论，设计实现最符合当前 5G 方案应用的高集成模组芯片产品。以下就对此芯片的详细规格进行介绍。

3.5.1 双频、低压、小尺寸：懂您所需，予您所求

S55235 是慧智微基于第四代 AgiFEM5GTM 可重构射频前端平台开发的 n77(78)/n79 射频前端模组。S55235 具有以下特性：

- 支持 n77(78)/n79 5G UHB 双频频段；
- 集成功率放大器 PA、低噪声放大器 LNA 和射频收发开关；
- 集成滤波器，满足带外抑制需求，以应对复杂的 EN-DC 场景；
- 集成 SRS 输出开关，支持 SRS 快速切换；
- 集成天线切换开关，提高扩展能力；
- 全倒装设计，提升器件散热能力和射频性能；
- 满足不同调制波形在协议规定回退条件下的线性和 EVM 指标；
- 支持 LNA 七级增益挡位，匹配平台接收需求；
- 支持 1T4R/2T4R 方案；
- 在低压 3.4V 供电下，n78/n79 工作时，模组输出功率达 29dBm，支持整机天线口 26dBm PC2 发射功率需求；
- 3.0mmx5.0mm 小型化尺寸，pin 脚与 n77(78) 单频 1T2R 产品兼容设计，方便单、双频方案兼容使用；
- -40 至 85 度可靠性测试，满足工业需求；

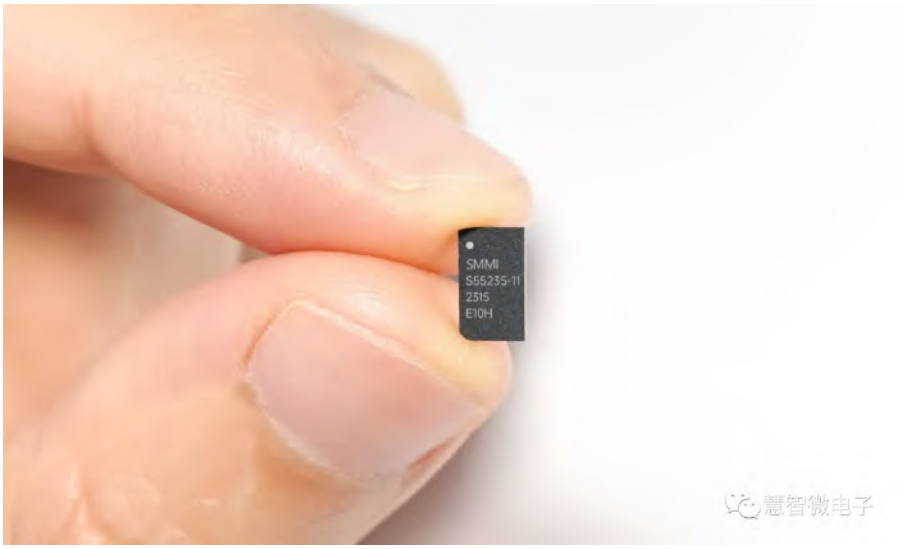


图 3.75: S55235 产品照片

3.5.1.1 尺寸缩小 40%：实现双频频段，且与单频兼容

根据 GSMA（全球移动通信协会）统计，2022 年，中国、美国等国家 5G 手机连接数在 40% 左右，全球 5G 领先国家在 2030 年连接数将在 80% 以上。除了中国、美国、日本等领先商用地区外，在欧洲、美洲等全球地区快速增长 [1]。

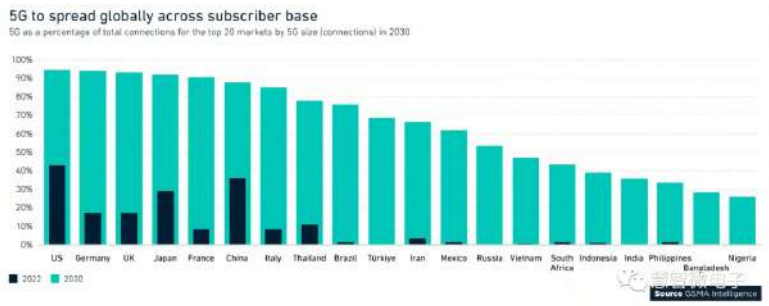


图 3.76: 全球 5G 未来迎来快速发展

相比于 4G 等以往协议，5G 需要满足高速率、大容量、低时延等应用。为了对这些应用有良好的支持，5G 开辟了 Sub-6GHz 以及毫米波新频段。

Sub-6GHz 指 6GHz 以下频段，一般专指 3GHz ~6GHz 频率范围。相比于原有 4G 频段，5G Sub-6GHz 频段的优势是带宽大，容量大，同时相比于毫米波频段来说，5G Sub-6GHz 频段覆盖范围广，穿透力强，部署成本低。以上特性特别适合 5G 创新的应用场景和业务需求，所以 Sub-6GHz 一经确定，就成为 5G 的核心拓展频段。

	频段范围	频段说明	代表频段	应用特点
Sub-3GHz	0.6-2.7 GHz	• 处于3GHz以下，称为Sub-3GHz • 原来的4G频段，也称4G重耕频段 • 以1/2GHz为界，划为低/中/高频	• 低频：n28（703-748 MHz） • 中频：n3（1710-1785 MHz） • 高频：n41（2496-2690 MHz）	• 覆盖范围广，部署成本低 • 带宽小，大多在在几十MHz
Sub-6GHz	3.3-6.0 GHz	• 5G新开辟频段 • 一般专指3.0-6GHz的频率范围 • 又被称为超高频（UHB）	• n77（3.3-4.2 GHz） • n78（3.3-3.8 GHz） • n79（4.4-5.0 GHz）	• 覆盖范围较广，部署成本较低 • 带宽大，数百MHz
毫米波	24-40 GHz	• 波长在毫米量级的频段 • 距离其他原有频段较远，方案与基站部署有较大调整	• n257（26.5-29.5 GHz） • n260（37-40 GHz）	• 覆盖范围小，部署成本高 • 带宽极大，数GHz

慧智微电子

图 3.77: 5G 不同频段的特点



图 3.78: 5G 不同频段的带宽及部署成本比较

由于 n77(78)/n79 频段带宽相比于原有 4G 频段可以有明显的提升，同时保持相对较低的部署成本，所以也成为 5G 应用中的主力频段。除美国、韩国重点发力毫米波外，世界多数国家对此频段进行 5G 大力部署。中国也是 5G Sub-6GHz 的重要推进地区，中国移动、电信、联通、广电也获得了此频段的授权使用。在 5G n77(78)/n79 频段，中国四大运营商频谱分配如下：

- 中国移动：4800-4900MHz（100MHz）的 5G n79 频段
- 中国电信：3400-3500MHz（100MHz）的 5G n77(78) 频段
- 中国联通：3500-3600MHz（100MHz）的 5G n77(78) 频段
- 中国广电：4900-5000MHz（100MHz）的 5G n79 频段

n77(78) 频段是全球重点部署的 Sub-6GHz 频段。除 n77(78) 外，全球越来越多的地区对 n79 进行部署。中国是全球 n79 频段应用的领先国家。除中国外，日本等国家也对 n79 进行商用支持，俄罗斯等一些欧洲国家有计划在接下来将 n79 进行商用。

全球应用的手机一般对 n77(78) 及 n79 频段进行了较好的支持。iPhone 全系列手机支持 n79 频段，vivo、OPPO、荣耀、小米等高端旗舰手机中也对 n79 进行支持。在三星销量最高的 A 系列部分手机中，也对 n79 进行支持。在三星 2021 年全球实现数千万部出货的 A22 手机中，就搭载了慧智微高集成 n77(78)/n79 模组芯片，实现了对 n77(78)/n79 的全频段支持 [4]。

n79 还是物联网应用的黄金频段，n79 大带宽、高传输特性非常适用于大数据传输、高可靠工业应用场景。中国也是对 n79 频段在物联网中的应用做了强制规定。根据《2023 中国移动行业终端产品白皮书》显示，物联网行业应用终端必须支持 n79 频段 [3]。

不过，**对 n79 频段的支持也给系统设计带来了挑战**。主要的挑战有：

- 一般 n79 与 n77(78) 是两个射频通路，这就意味着支持 n79 需要有另外的成本投入；
- 一般支持同时支持 n79 及 n77(78) 的芯片模组尺寸都要大于单 n77(78) 的芯片模组，造成无法进行共板设计，支持 n79 时必须要另起一个项目进行设计；
- n79 可选择的方案少于 n77(78)，造成目前方案不够优化；

在之前设计中，因为没有优化的射频方案，使得终端支持 n79 需要额外的设计投入，以及芯片成本提升。使得 n79 终端并未大规模普及开来。

根据中国信通院统计数据，在入网检测中，2020 年约有 80% 的手机包含 n79 频段。随着 2020 年运营商白皮书对 n79 频段进行了价位划分，4,000 元以上手机才会强制要求支持 n79，于是即使大带宽的 n79 会带来网络体验的明显提升，在设计挑战和成本挑战压力下，大部分手机还是选择放弃了对 n79 的支持。根据 2022 年第一季度的统计，入网检测的手机中，支持 n79 的比例为 22.2% [2]。

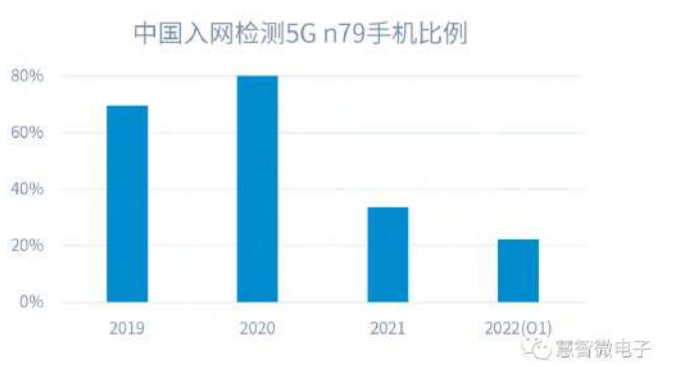


图 3.79: 中国信通院入网检测 5G 手机比例

未来随着 5G 的全球演进，物联网应用的深入，随着国内网络部署的丰富和完善，将会有越来越多支持 n79 的项目需求。于是，**慧智微联合终端厂商、平台厂商，共同定义开发全新 5G n77(78)/n79 L-PAMiF 芯片**，提供支持 n79 及 n77(78) 的优化解决方案。

在 S55235 量产之前，市面上类似功能的产品标准封装尺寸为 4.5mmx5.5mm。慧智微自 2019 年开始量产此标准尺寸的 5G UHB L-PAMiF 产品。

S55235 是慧智微在原有产品基础上，继续提升集成度、减小模组尺寸。S55235 尺寸只有 3mmx5mm，相比于原有产品，尺寸缩小 40%。

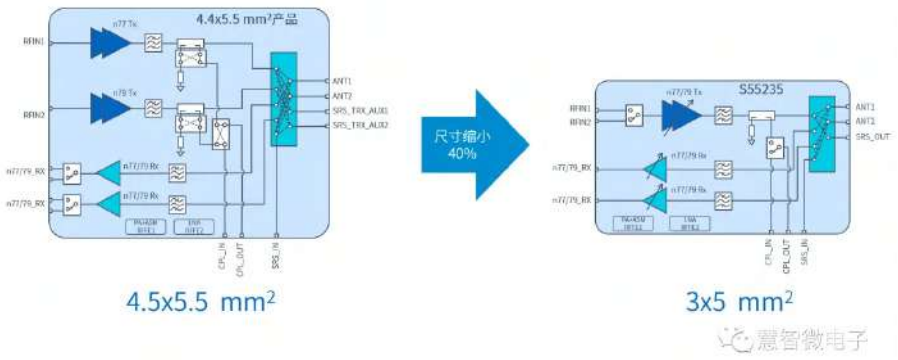


图 3.80: S55235 在原有高集成 5G UHB 双频 1T2R L-PAMiF 基础上，尺寸再缩小 40%

采用 3mmx5mm 封装尺寸后，支持 n77(78)/n79 的 5G UHB 双频 1T2R L-PAMiF 尺寸可以做到与 n77(78) 单频 1T2R L-PAMiF 相同。在管脚定义中，产品也做了兼容考虑，做到与 n77(78) 单频 1T2R L-PAMiF pin 脚兼容，方便用户在设计中实现 n77(78)/n79 项目与单 n77(78) 项目的兼容设计，极大的提升 5G 项目中 UHB 单、双频共板设计的便利性。

3.5.1.2 3.4V PC2：最新需求，必须满足

根据慧智微公众号文章：《手机“射频增强”的技术方案》讨论，移动终端的通信能力和信号的发射功率成正比。为了实现 5G 终端的良好通信能力，5G 协议对 n41、n77(78)、n79 等频段定义了更高的功率等级：Power Class 2 功率等级，简称 PC2。

在 PC2 的定义下，终端的发射功率由原来的 PC3 基础上提升 3dB，天线口功率提升至 26dBm。

	终端功率	代表频段
PC3	23dBm（±2dB）	• n28 • n3 • n7
PC2	26dBm（+2/-3dB）	• n41 • n77 • n79

图 3.81: 3GPP 所定义的功率等级

功率提升 3dB 换算成倍数就是增加为原来的 2 倍。根据《5G PA 的 Load-line 与 Load-pull》讨论，射频 PA 的输出功率严格依赖于供电电压，在只有 3.4V 电池供电的移动终端来说，功率提升一倍并非易事。

于是，在 5G 应用的早期，一般是采用电源升压芯片（Boost DC-DC），将 3.4V 电压转变到 5V 附近，再供给 PA 使用，来实现 5G PA 的大功率输出。

但这种应用也增加了成本，一般一颗 Boost DC-DC 芯片价格接近 1 美金，这在很多消费类终端以及成本敏感的工业应用终端中，是不可接受的。为了节省 Boost DC-DC 芯片的使用，终端需要 5G 射频 PA 可以在 3.4V 下即支持 PC2 的功率输出。



图 3.82: 5G 终端的 3.4V PC2 功率输出需求

在 PA 设计中，为了达到 3.4V PC2 功率输出，可以采用功率合成的方式（见《5G 射频 PA 架构》）或线性功率校正优化的方式，来实现低压下功率的提升。慧智微第四代 AgiFEM5GTM 可重构射频前端平台对低压高功率特性开展了深入研究，利用可重构技术，实现 3.4V 下的高功率输出。

S55235 产品采用慧智微最新的低压 5G PA 内核，支持 3.4V 电压下工作时输出 PC2 高功率。相比于之前产品必须借助外部升压电路，实现 4V、5V 以上电压后，再实现 PC2 功率输出，S55235 产品直接在 3.4V 下进行降压操作，就可以对全功率范围内支持，甚至可以支持直接连接电池的供电支持。

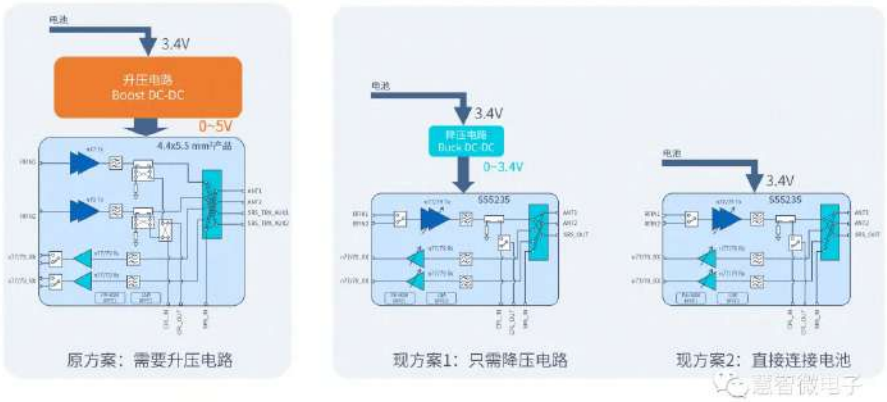


图 3.83: 支持 3.4V 下输出 PC2 功率等级，不再需要外部升压电路

S55235 的 3.4V PC2 特性，简化了电源方案的设计，为整个方案的小型化和集成化提供了便利。

3.5.1.3 不止商规：支持-40 到 85 度工规需求

虽然过去几年 5G 手机应用快速发展，但 5G 应用的前景绝不止是手机，5G 物联网应用中将得到更长远发展。

因 5G 技术可以为物联网提供高速率、低时延、大连接等特性，使得物联网的应用场景更加丰富和广泛，包括智能制造、智慧城市、智能交通、远程医疗、智能家居等。5G 技术可以实现物与物之间的高效通信和协作，为物联网带来更多的创新和价值。

根据 GSMA 的预测，自 2021 年后，蜂窝物联网将迎来快速发展。自 2022 年至 2023 年，蜂窝物联网连接数目将翻一番，从 26 亿，增长至 53 亿 [1]。

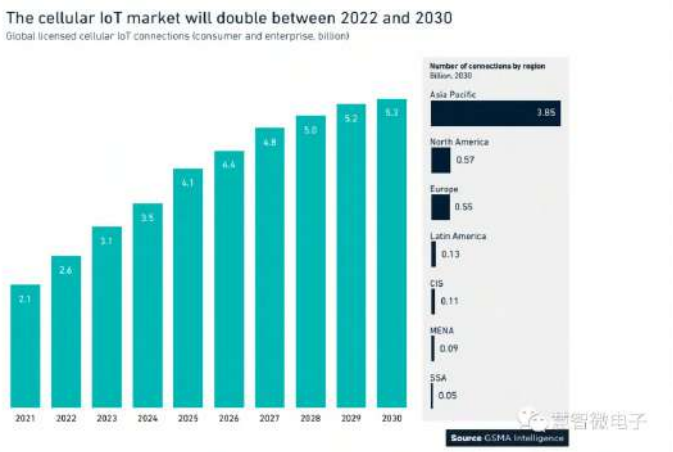


图 3.84: 2022 至 2030，蜂窝物联网市场将翻一番

和手机一般是在“有人“的环境中不同，物联网应用环境复杂，所以对芯片的可靠性也提出了更多的要求。

对于 5G 射频芯片来说也是一样，物联网应用中的 5G 射频芯片负责信号的无线传输，是物联网产品中的重要功能模块，需要严格确保产品可靠。与商规芯片相比，工业应用中的工规产品差异化主要有：

- 工作温度范围：
  工规芯片一般要求能在-40°C ~85°C 的环境下正常工作，而商规芯片一般只要求能在-20°C ~60°C 的环境下正常工作
- 可靠性标准：
  工规芯片一般要求符合 JESD47 等国际标准，进行各种高温、高湿、高压等可靠性测试，保证芯片的寿命和稳定性，而商规芯片的可靠性测试要求相对较低

- 应用领域：
  工规芯片一般应用于工业控制、通信、医疗、航空等对可靠性和安全性要求较高的领域，而商规芯片一般应用于消费电子、娱乐、家居等对可靠性和安全性要求较低的领域
- 价格差异：
  由于工规芯片的生产成本、设计难度、质量管控等都比商规芯片高，因此工规芯片的价格也比商规芯片高

5G 的万物互联应用，也对 5G 射频前端芯片的可靠性提出了更高的要求。为了支持 5G 产品在物联网市场应用，S55235 产品在设计时就按照工规标准进行设计，产品出厂前经过严苛、完整的可靠性测试，产品支持-40 到 85 度工规需求。

3.5.2 总结

S55235 是慧智微利用慧智微技术优势，开发的满足 5G 演进最新需求的射频前端产品。如果您在项目中有 5G n77(78)/n79 **双频、小尺寸单频兼容、3.4V PC2 的高集成、工规标准的** 5G L-PAMiF **模组需求**，欢迎联系慧智微在地销售，获取产品更多技术资料。

慧智微期待与您一起，化繁为简，使一切智慧互联。

参考文献

1. GSMA，Global Mobile Trends 2023
2. 中国信通院，https://digi.china .com /digi /2022 06 09 / 2022 06091 081880.html
3. 中国移动，2023 中国移动行业终端产品白皮书
4. 慧智微，三星新一代 Galaxy A 系列 5G 手机全球发布，慧智微 5G PA inside
5. 3GPP，TS38.101，User Equipment (UE) radio trans mission and reception

在线阅读

本文摘自慧智微微信公众号，扫描以下二维码，在线阅读

第四篇 Wi-Fi 技术与市场

本篇内容

- Wi-Fi：连接未来，缔造美好生活
- Wi-Fi 系统：技术方案与实现路径
- Wi-Fi 市场探析：今日应用与未来趋势
- Wi-Fi PA 和手机 PA，有什么不同？

在无线通信的广阔天地里，除了蜂窝网络这一重要支柱外，Wi-Fi 技术同样扮演着举足轻重的角色。随着互联网的普及和移动设备的激增，Wi-Fi 已成为我们日常生活中不可或缺的一部分，它为我们提供了便捷、高速的无线接入方式。而 Wi-Fi 协议也在快速演进，推动着 Wi-Fi 射频系统经历着翻天覆地的变化。

本章，我们将对 Wi-Fi 技术做一次系统的分析。从 Wi-Fi 的起源、发展历程，到当前的最新标准和技术特点，我们将一一剖析，帮助大家全面了解 Wi-Fi 技术的全貌。同时，我们还将深入解析 Wi-Fi 市场的发展现状和趋势，探讨 Wi-Fi 技术在各个领域的应用情况和市场前景。

在 Wi-Fi 射频系统方面，随着 Wi-Fi 标准的不断升级，对射频性能的要求也在不断提高。从最初的 802.11a/b/g，到如今的 802.11ax（Wi-Fi 6）及 802.11be（Wi-Fi 7），Wi-Fi 射频系统一直在不断追求更高的传输速率、更低的功耗和更好的覆盖范围。我们将详细介绍这些标准的技术特点和对射频系统的要求，以及射频技术在实现这些要求中所扮演的关键角色。

此外，我们还将探讨 Wi-Fi PA 与手机 PA 的异同。虽然 Wi-Fi PA 和手机 PA 在功能上都是用于放大射频信号，但由于应用场景和性能要求的不同，它们在设计和实现上存在着显著的差异。我们将从工作原理、性能指标、设计挑战等方面对两者进行比较，帮助大家更好地理解它们的异同点。

通过本章的材料，我们希望大家能够全面了解 Wi-Fi 技术的发展历程和市场现状，深入理解 Wi-Fi 射频系统的关键技术和挑战，以及掌握 Wi-Fi PA 与手机 PA 的异同和未来发展方向。这将有助于我们在实际开发中更好地选择和使用 Wi-Fi 技术。

4.1 Wi-Fi：连接未来，缔造美好生活

作者：彭洋洋 博士

2007 年，iPhone 诞生，它不仅改变了手机行业的格局，更彻底改变了我们的上网方式。从此，一个颠覆性的时刻到来，移动互联网时代的序幕被拉开。但苹果对世界的影响，远不止于此。

回溯到 1999 年，苹果推出了 iBook，这是第一次将无线网络技术融入到笔记本电脑中。那时的创新，为后来设备的移动联网奠定了基石。在 iBook 发布会上，苹果的传奇领导者乔布斯，为了证明电脑真的没有连接网络线，在台上用呼啦圈绕着电脑旋转，魔术一般的操作引起场下长时间惊呼，而这一幕也成为了科技史上的经典瞬间。



图 4.1: 1999 年苹果 iBook 发布中的经典一幕，乔布斯用呼啦圈证明笔记本的无线连接

当时，那台 iBook 所使用的无线网络技术，正是 802.11 系列协议，也就是我们今天所熟知的 Wi-Fi 协议的前身。当时谁也没能想到，这样一个无线通信协议，会在未来近 30 年里不断演进，成为连接数十亿设备的核心纽带。

如今，Wi-Fi 已如空气和水一般，成为我们生活中不可或缺的元素。甚至有这样的段子流传：“Wi-Fi 一断，全家人都开始重新认识彼此。”这足以见证 Wi-Fi 在现代生活中的重要性。



图 4.2: Wi-Fi 在现代生活中已经非常重要

那么，Wi-Fi 究竟如何从一个简单的无线通信协议，发展到现在这样无处不在的存在呢？它与蜂窝网络、蓝牙等无线网络又有何区别？更重要的是，Wi-Fi 的未来将向何处发展？

带着这些疑问，让我们一起探寻 Wi-Fi 技术的发展轨迹，看看这个让我们的生活变得更加美好的技术，是如何一步步走来，又将走向何方。

4.1.1 Wi-Fi 协议的诞生及演进

4.1.1.1 Wi-Fi 协议的起源：IEEE 802.11 协议

谈及 Wi-Fi 协议，就不得不提及 IEEE（电气电子工程师学会）的 802.11 协议。IEEE 推出的 802.11 协议旨在构建一个统一的无线通信标准。然而，作为一个专注于技术标准制定的组织，IEEE 并不涉及产品的测试与认证。因此，在初期，市场上的 802.11 设备普遍存在兼容性问题，制约了 Wi-Fi 技术的广泛推广与应用。

在这种背景下，1999 年 Wi-Fi 联盟应运而生，有效解决了这一问题。Wi-Fi 联盟是由多家领先公司共同发起的行业组织，其核心使命是负责 802.11 相关产品的测试与认证。联盟的名称“Wi-Fi”源自“Wireless Fidelity”的缩写，寓意着无线通信的可靠与稳定。同时，联盟还设计了独特的 Wi-Fi 标志，进一步提升了品牌辨识度。自此，Wi-Fi 作为无线通信的重要技术标准，开始广为人知。



图 4.3: （a）Wi-Fi 联盟的 logo 及（b）Wi-Fi 的认证标识

随着技术的不断发展，Wi-Fi 逐渐被越来越多的应用所接受，甚至在一些场景中直接成为无线局域网技术（WLAN）的代名词，人们直接用“Wi-Fi”来指代整个 WLAN 技术，足以见其影响力之广泛。

4.1.1.2 Wi-Fi 协议的今生：从 Wi-Fi4 到 Wi-Fi7 飞跃

随着科技的不断进步，IEEE 802.11 协议也历经了多次更新，从 802.11a、802.11b 逐步发展至 802.11ac，再到如今的 802.11ax（即 Wi-Fi 6）和最新的 802.11be（即 Wi-Fi 7）。这些更新不仅标志着技术的显著进步，更带来了数据传输速率的惊人增长。从最初的几 Mb/s 到如今高达 46Gb/s 的峰值速率，Wi-Fi 正迅速适应并满足现代社会对高速无线通信的迫切需求。

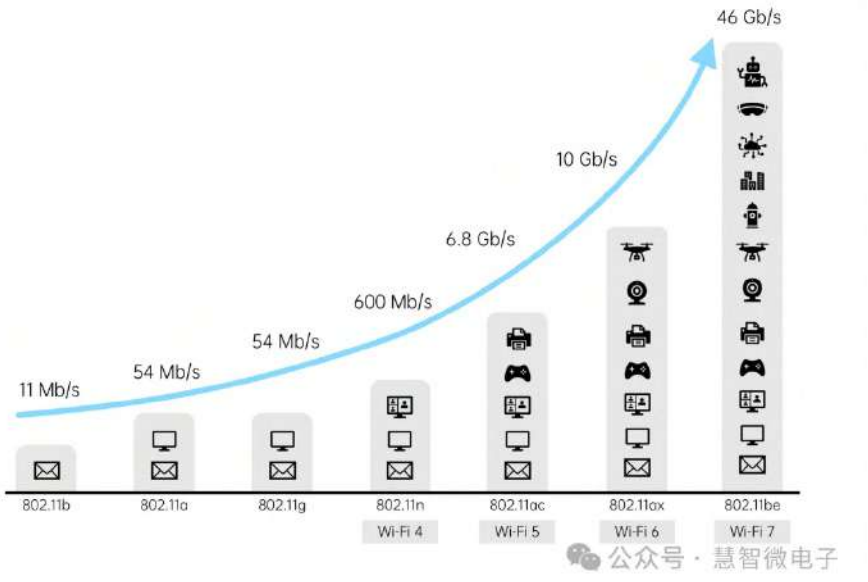


图 4.4: Wi-Fi 速率的飞跃发展

然而，随着 802.11 技术的持续升级，诸如 802.11b、802.11ax、802.11be 等复杂的版本号让普通用户感到困惑。这些版本号中的数字和字母组合不仅难以理解，而且并不按照协议发布的顺序排列，进一步增加了用户的认知负担。为了简化这一复杂性并推动 Wi-Fi 的更广泛应用，Wi-Fi 联盟决定对这些版本进行重新命名。

2018 年，802.11ax 推出，但由于名字和上一代的 802.11ac 过于相似，消费者根本不知道哪个才是最新的通信制式。于是，在 2018 年 10 月 4 日，Wi-Fi 联盟实施了一项重大的命名改革，将 802.11ax 协议更名为 Wi-Fi6，以便消费者更好地辨别。

由于 802.11ax 更名为 Wi-Fi6，前几代协议也跟着一起更名了，于是当时市面上大多数设备采用的 802.11ac 被称之为 Wi-Fi5, 而更早的 802.11n 协议被命名为 Wi-Fi4。

将复杂的版本号简化为更直观、易懂的“Wi-Fi 4、Wi-Fi 5、Wi-Fi 6、Wi-Fi 7”等名称，这一改革不仅极大地提升了消费者对产品技术水平的直观理解，还有力地促进了 Wi-Fi 技术的普及和应用。尽管现代设备有时仍会标注其支持的 802.11 系列协议版本，但人们更倾向于使用简洁明了的“Wi-Fi + 数字”命名方式。这次命名变革无疑是一次巨大的成功，并为其他技术领域提供了宝贵的借鉴经验。

	Wi-Fi4	Wi-Fi5		Wi-Fi6	Wi-Fi7
协议	802.11n	802.11ac		802.11ax	802.11be
		wave1	wave2		
年份	2009	2013	2016	2018	2021
频段	2.4/5GHz	2.4/5GHz		2.4/5GHz	2.4/5/6GHz
最大频宽	40MHz	80MHz	160MHz	160MHz	320MHz
最高调制	64QAM	256QAM		1024QAM	4096QAM
单流带宽	150Mb/s	433Mb/s	867Mb/s	1200Mb/s	2880Mb/s
最大带宽	600Mb/s	3.5Gb/s	6.8Gb/s	9.6Gb/s	46Gb/s
最大空间流	4x4	8x8		8x8	16x16
MU-MIMO	不支持	不支持	下行	上/下行	上/下行
OFDMA	不支持	不支持		上/下行	上/下行
MLO	不支持	不支持		不支持	支持

公众号 · 慧智微电子

图 4.5: 不同版本的 Wi-Fi 特性，及与 802.11 标准的对应

随着 Wi-Fi 在技术和市场上的广泛认可，其应用也日趋多样化。据 TechInsights 统计，截至 2023 年，全球 Wi-Fi 设备的连接数量已达到惊人的 72 亿，并将在未来 5 年内增长至 89 亿。Wi-Fi 技术在应用领域也实现了巨大的突破，深刻影响着我们的日常生活和工作方式。

4.1.1.3 Wi-Fi 中的关键技术

Wi-Fi 作为一种持续进化的无线通信技术，依托多项关键技术实现其功能并不断优化用户体验。以下将深入探讨 Wi-Fi 中的这些关键技术及其发展。

1. 使用频率和信道

Wi-Fi 主要使用的频率属于工业、科学和医疗（ISM）频段以及非授权频段，这包括 2.4GHz、5.8GHz 以及新兴的 6GHz 频段。选用这些频段的原因在于其全球通用性和无需授权的特性，这为 Wi-Fi 设备的全球普及和应用提供了便利，真正实现了“全球通”的愿景。

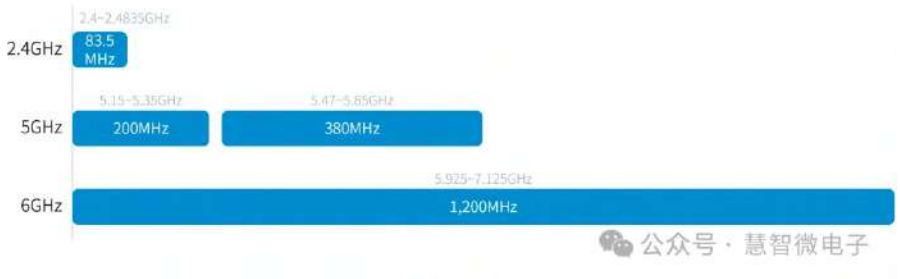


图 4.6: Wi-Fi 使用的全球主要国家和地区 ISM 及非授权频段

这些频段并非单一频点，而是一定范围的频率集合。例如，2.4GHz 频段被划分为多个宽度一定的信道，Wi-Fi 设备可以在这些信道间灵活切换，以优化信号质量。但需注意，信号间的频率重叠可能引发干扰，因此实际操作中常选择不重叠的信道进行使用，如常用的 1、6、11 信道。

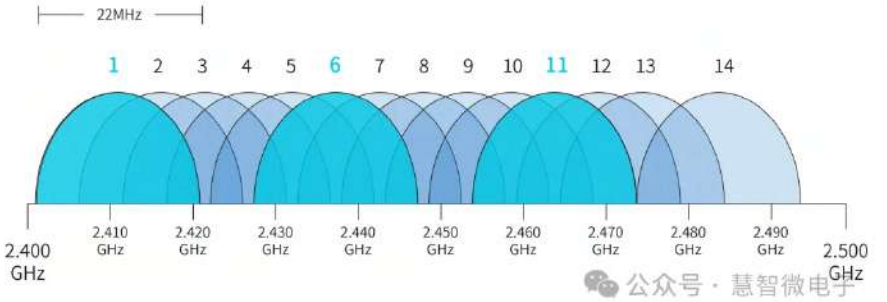


图 4.7: 2.4GHz Wi-Fi 的信道划分

值得一提的是，虽然非授权频段在全球大体统一，各国在非授权频段具体使用中还是存在部分差异。例如，CH14 的中心频率为 2.484GHz，这个信道在大多数国家都属于禁用频段，但日本除外。所以，如果将日本的 Wi-Fi 设备拿到其他国家使用，就需要在设备设置中禁用掉 CH14 信道，来确保设备之间的顺畅连接。

2. 高阶调制技术

Wi-Fi 广泛应用于室内高速数据传输场景，因此其调制和编码技术的演进至关重要。从 QPSK 到 4096QAM，Wi-Fi 的调制方式不断升级，显著提升了数据传输效率。例如，4096QAM 技术使得单个信号能够携带的信息量比 QPSK 增加了六倍。

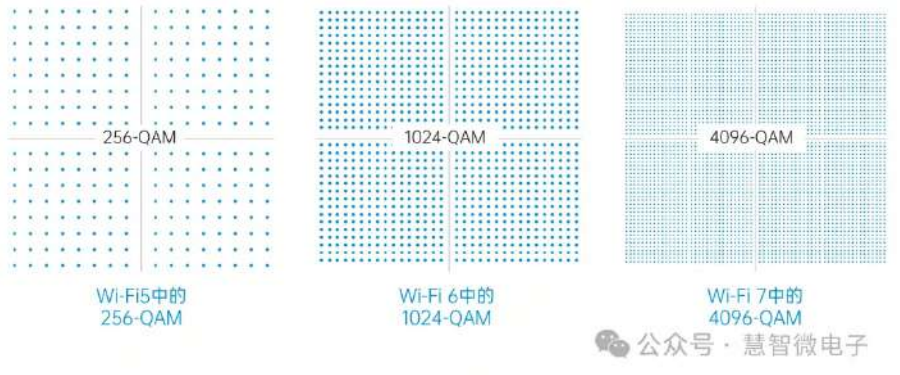


图 4.8: 不同制式 Wi-Fi 中的调制技术

更高的调制方式就像是在同一辆汽车上拉了更多的货物，在 Wi-Fi5 的 256-QAM 中，每一个符号携带 8bit 信息，到了 Wi-Fi6 的 1024-QAM 中，每个符号携带的信息增加到 10bit，而在 Wi-Fi7 中，每个符号位的信息量增加到 12bit。每一辆汽车上拉的货物增多了，信息的传输速率自然增加。

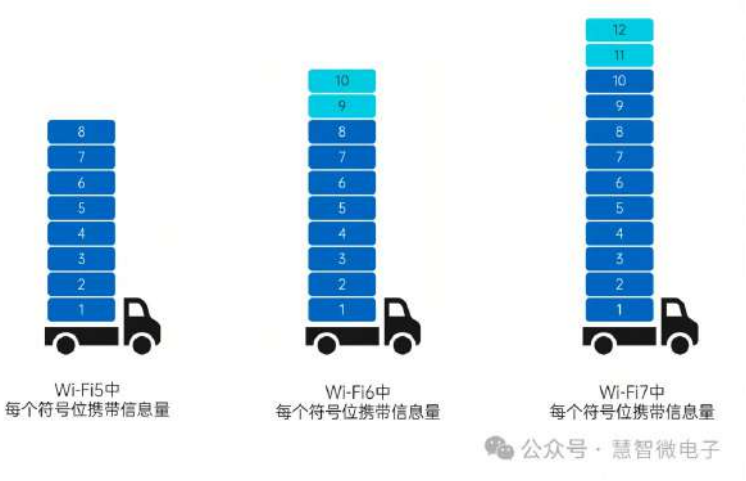


图 4.9: 不同制式 Wi-Fi 每个符号位携带的信息量

3. MIMO 技术

多输入多输出（MIMO）技术是无线通信领域的一项重要创新，也在 Wi-Fi 中得到了广泛应用。MIMO 技术通过多个天线同时发送和接收数据，显著提高了无线通信的吞吐量和可靠性。MIMO 技术就像是在起始点和目的地之间修了更多的通路，这样就可以完成多路的并行传输。

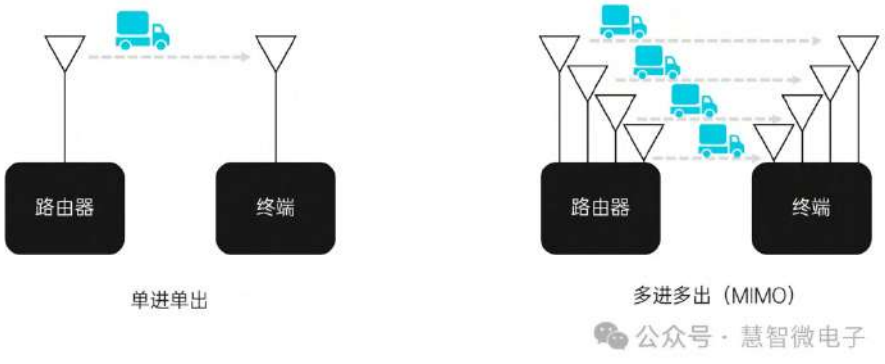


图 4.10: MIMO 技术示意图

在 Wi-Fi 中，MIMO 技术主要分为单用户 MIMO（SU-MIMO）和多用户 MIMO（MU-MIMO）两种模式。SU-MIMO 主要服务于单个用户，提升其与基站之间的通信效率；而 MU-MIMO 则允许多个用户同时与基站通信，大大提高了网络的整体效率。

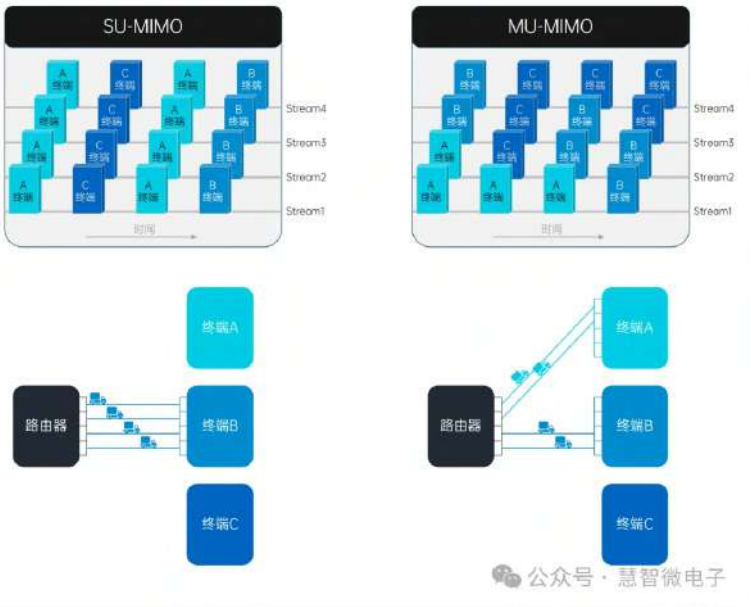


图 4.11: SU-MIMO 与 MU-MIMO 连接示意图

4. Wi-Fi 中的直连技术

尽管 Wi-Fi 网络通常包括接入点（AP）和站点（Station），但 Wi-Fi 技术并不完全依赖于中心化的网络结构。为了满足设备间直接通信的需求，Wi-Fi 直连技术应运而生。

Wi-Fi 中的直连技术指的是设备之间通过 Wi-Fi 网络直接进行连接而无需经过路由器或者基站。这种连接方式可以实现设备之间的点对点通信，提高了通信效率和速度。

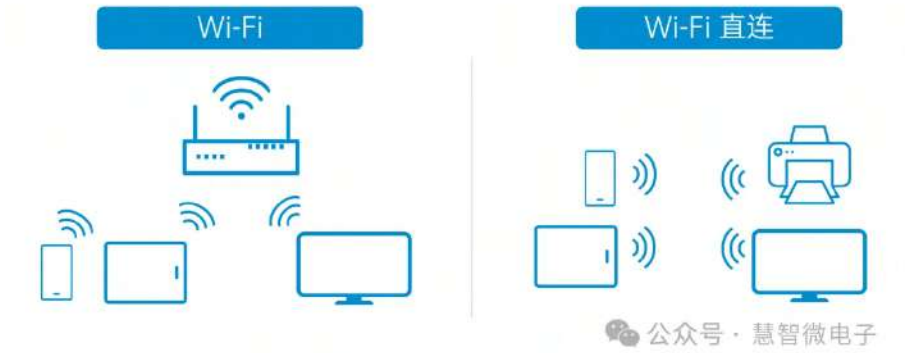


图 4.12: Wi-Fi 中的直连技术

在 Wi-Fi 的直连技术中，Wi-Fi Direct 和 Miracast 是两项代表性技术。Wi-Fi Direct 允许设备间直接建立连接，无需通过传统的 Wi-Fi 路由器。这一技术在打印机、媒体播放器等设备中得到了广泛应用。而 Miracast 则主要用于屏幕镜像和流媒体传输，如 Windows 设备的投屏功能就依赖于这一技术。



图 4.13: Wi-Fi 直连应用：打印机及 Windows 的直连

4.1.2 Wi-Fi 与其他技术的比较

在通信协议领域，各种技术之间的比较是不可避免的，这种比较有助于我们明确不同技术的定位和发展方向。

无线通信协议的评价指标有很多，如速率、成本、功耗、安全性和传输距离等，每个维度都有其独特的重要性。然而，在这些因素中，我认为传输距离是最为核心的考量点。传输距离直接决定了技术的应用场景，而不同的应用场景又会进一步决定功耗、安全性等其他需求。

以安全性为例，对于传输距离较短的蓝牙技术而言，由于其范围在人们可控之内，因此对安全性的要求相对较低，简单的配对步骤即可使用。然而，对于传输距离可达几十米的 Wi-Fi 技术来说，则需要设置密码以确保连接的安全性。当传输距离扩大至几千米的蜂窝网络层面时，由于无线信号可能覆盖大量设备，因此需要采用更为复杂的认证流程来建立设备间的联系，从而安全性的要求也大幅提升。另外不仅是安全性，功耗、成本、速率等其他技术指标也与传输距离紧密相关。

不同通信协议的覆盖范围及其应用场景可参考图4.14进行直观了解。

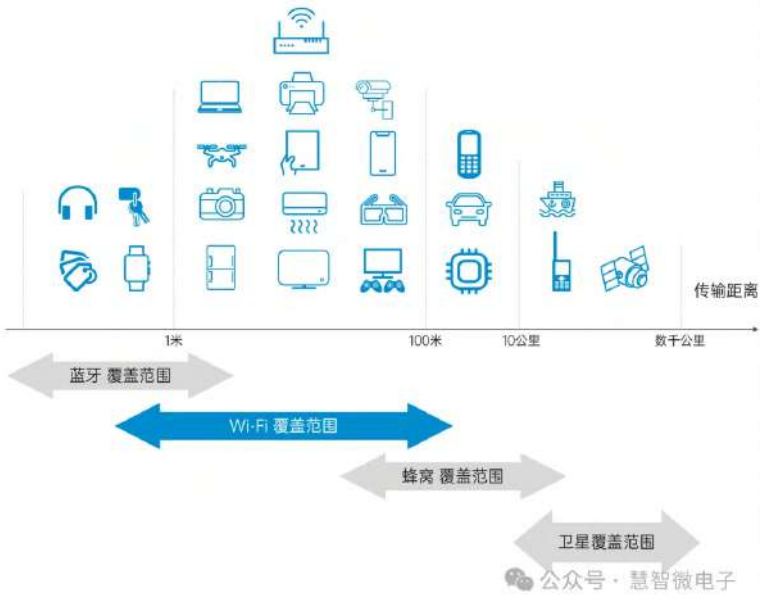


图 4.14: 不同通信协议覆盖范围及应用场景

在各类应用场景中，与 Wi-Fi 技术形成互补或存在重叠的主要是蜂窝技术和蓝牙技术。接下来，我们将通过对比这两种技术，来更深入地剖析 Wi-Fi 技术的特点与优势。

4.1.2.1 Wi-Fi 与蜂窝技术的比较

Wi-Fi 与蜂窝技术，两者有诸多相似之处，还曾一度被视为竞争对手，但随着技术的演进和市场需求的变化，人们逐渐认识到，这两种技术实则是相辅相成的。

Wi-Fi 与蜂窝技术共同点颇多：两者都是手机等设备的标配通信协议，都能满足日常高速通信的需求，且在许多使用场景中都有重叠。这些共通之处曾导致 Wi-Fi 和蜂窝技术阵营在发展过程中产生过竞争。在 4G 发展的初期，Wi-Fi 阵营试图将其发展的 WiMAX 协议，推广到 4G 的标准协议中。而蜂窝阵营，也一直尝试用 4G/5G 的部署，来取代只在室内使用的 Wi-Fi 协议。

然而，随着无线通信需求的激增，人们意识到，单一的通信协议无法满足所有场景的需求，Wi-Fi 与蜂窝技术逐渐从竞争关系转变为互补关系，共同为用户提供全面的无线通信服务。“地球足够大，容的下两个超级的无线通信协议”。

尽管有诸多相似，但 Wi-Fi 与蜂窝技术在设计理念、技术特性和发展路径上存在显著差异。

Wi-Fi 起源于互联网技术中的局域网技术，其核心目标是提供一种便捷的互联网接入方式，以替代传统的有线连接。因此，Wi-Fi 注重开放频段、全球标准接口和去中心化的连接方式，旨在丰富全球互联网网络。

而蜂窝技术则传承自广域网的设计理念，其核心目标是实现全球范围内的可靠连接。这一要求使得蜂窝网络必须是中心部署且统一调度的，以确保广泛性和可靠性。

这些不同的设计理念导致两者在发展方式上也有所不同。Wi-Fi 技术更侧重于接入速率，因此其通信速率一直领先于蜂窝网络。而蜂窝网络则更注重覆盖范围和覆盖人数，在网络覆盖面积和多用户处理方面具有显著优势。

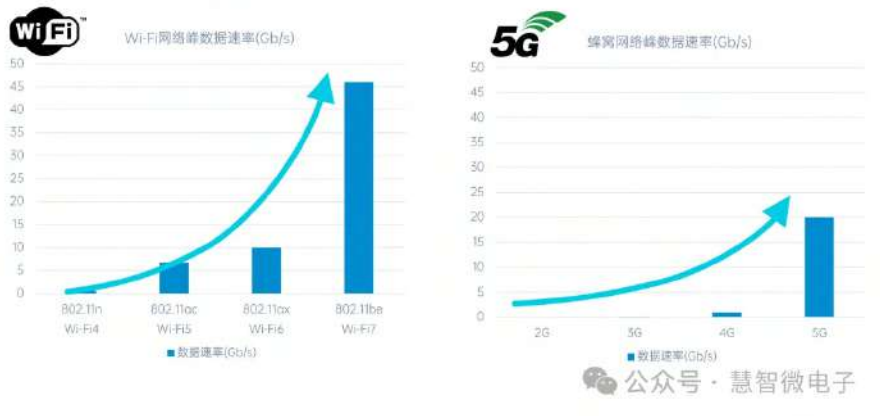


图 4.15: Wi-Fi 与蜂窝峰值速率比较

然而，这两种技术并非孤立发展，而是相互借鉴、相互促进。例如，蜂窝网络就一直在借鉴 Wi-Fi 中的复杂 QAM 调制、MIMO 等技术来提升速率，而 Wi-Fi 也借鉴了蜂窝网络中的多用户管理、载波聚合等技术以实现多用户覆盖和多频率使用。

总的来说，Wi-Fi 与蜂窝技术各有千秋，它们在无线通信领域中发挥着不同的作用。通过相互借鉴和融合，这两种技术将继续推动无线通信行业的发展，为用户提供更加便捷、高效的服务。

4.1.2.2 Wi-Fi 和蓝牙技术的比较

Wi-Fi 和蓝牙技术都是我们日常生活中经常使用的通信技术，它们之间存在许多共同点。两者都工作于 ISM 频段，为近距离的设备提供接入服务。无论是手机、平板电脑还是笔记本电脑，这些与我们生活息息相关的设备都广泛支持并应用了这两种技术。然而，Wi-Fi 和蓝牙在传输距离上存在显著差异，这也是它们主要的区别之一。Wi-Fi 主要设计用于覆盖数十米的固定空间，实现设备间的远程连接。相比之下，蓝牙则专注于几米范围内的短距离连接，更适用于人手边的设备。

蓝牙对于应用场景的精准的定位使得其能够进行更为细致的优化。由于蓝牙连接的设备通常随人移动，因此需要对其进行极致的功耗优化，从而免除频繁充电的烦恼。同时，这些设备通常不需要过高的通信速率，因此在通信能力方面可以进行适当的减弱。此外，当蓝牙设备作为人类感知器官的延伸时，例如在使用 TWS 耳机时，对延时等指标有较高要求，因此蓝牙技术在这方面也进行了特别的优化。

尽管 Wi-Fi 协议功能强大且拥有庞大的组织支持，但在处理人手边小设备的小连接方面却显得力不从心。Wi-Fi 在保障大带宽、中距离互联的同时，难以兼顾小带宽、短距离的连接需求。而这些需求正是由蓝牙等协议来补齐的，两者在无线通信领域中各自发挥着不可或缺的作用。

4.1.3 Wi-Fi 技术的未来演进展望

在众多无线通信协议中，Wi-Fi 与人们的日常生活联系最为紧密。从手机、平板电脑到电视，Wi-Fi 的应用范围广泛，几乎渗透到我们生活的方方面面。无论是办公上网还是视频娱乐，Wi-Fi 技术都发挥着不可或缺的作用。

展望未来，Wi-Fi 技术的持续演进将进一步提升我们的生活品质。



图 4.16: Wi-Fi 技术的未来演进展望

4.1.3.1 传输速率持续提升

作为高速大带宽传输的佼佼者，Wi-Fi 的传输速率将继续保持领先地位。

在不久的未来，借助超高速 Wi-Fi 连接，我们将能够无缝观看高清甚至 8K 视频，彻底告别缓冲和画质问题的困扰。家庭成员将能够同时畅享 AR/VR 游戏、超高清视频会议或大文件传输，而这一切在最新 Wi-Fi 技术的加持下将如丝般顺滑。

4.1.3.2 设备互联的广泛支持

随着科技的迅猛发展和智能家居、物联网的普及，无线通信技术正迎来前所未有的发展契机，Wi-Fi 技术亦是如此。未来，Wi-Fi 将不再局限于为手机、电脑等传统设备提供网络服务，而是将扩展到支持更多设备的互联互通。智能家居设备、移动设备、智能城市基础设施以及各种传感器等都将通过 Wi-Fi 实现连接和数据传输，从而构建一个庞大而复杂的网络系统。

为了适应更大规模的数据互联需求，Wi-Fi 技术也需进行相应的优化和改进。例如，Wi-Fi 需解决多用户同时连接时的干扰问题，以确保多用户连接的稳定性和可靠性。

4.1.3.3 连接方式的便捷性革新

随着越来越多的设备接入 Wi-Fi 网络，传统的连接和输入密码的方式将逐渐显得繁琐不便。用户期待更加简便的网络接入方式，以便轻松、快速且安全地连接到网络。未来的 Wi-Fi 连接将不仅限于人与互联网的连接，还将广泛连接物联网设备。显然，当前基于密码的连接方式已无法满足物联网设备的接入需求。因此，未来的 Wi-Fi 将不仅作为网络的提供者，更将成为便捷生活的推动者。通过提供更便捷的连接方式，Wi-Fi 将使网络连接无处不在，让生活更加简单高效。

4.1.3.4 AI 与无线网络的深度融合

相比其他通信协议，Wi-Fi 的应用场景更为多样化且部署灵活，这为无线网络与 AI 的融合提供了有利条件。在 AI 时代，Wi-Fi 技术的发展将更加充满想象空间。

随着 AI 技术的融入，智能家居的控制有望变得更加智能化和精细化，同时可能带来前所未有的创新体验。例如，通过众多 Wi-Fi 连接的智能家居设备可以实时收集和共享数据，并利用 AI 进行分析以实现更高级的自动化控制。这将实现比现在更复杂、更丰富的场景控制及设备联动功能。

同时，Wi-Fi 技术也将助力 AI 的进一步发展。借助 Wi-Fi 技术，AI 生成的内容或所需的数据可以在各种设备之间实现高速灵活地传输。这将使各类设备更加紧密地协同工作，共同打造一个智能互联的未来世界。

4.1.4 总结

Wi-Fi，这一曾被视为简单的无线通信协议，如今已成为我们生活中不可或缺的一部分。无线网络技术逐渐融入我们的日常生活，开启了移动互联网的新篇章。Wi-Fi 不仅改变了我们的上网方式，更推动了科技的发展和社会的进步。

经过数十年的演进，Wi-Fi 技术已经从最初的 802.11 标准发展到现在的 Wi-Fi 7，传输速率大幅提升，应用场景不断拓展。从家庭娱乐到办公网络，从智能家居到物联网，Wi-Fi 技术无处不在，连接着数十亿设备，构建起一个庞大而复杂的网络系统。

在与蜂窝网络、蓝牙等其他技术的比较中，Wi-Fi 也展现出了其独特的优势。尽管蜂窝网络在广域覆盖和多用户处理方面具有显著优势，但 Wi-Fi 在速率和灵活性方面更胜一筹。而蓝牙则更适用于短距离、低功耗的场景，与 Wi-Fi 形成了互补。

展望未来，Wi-Fi 技术的持续演进将为我们的生活带来更多美好。传输速率将进一步提升，支持更多设备的互联，连接方式将更加便捷，AI 与无线网络的深度融合将开启全新的智能互联时代。随着科技的不断发展，我们有理由相信，Wi-Fi 还将继续引领无线通信技术的潮流，为我们的生活带来更多便利和惊喜。

您认为 Wi-Fi 未来会给我们的生活带来什么样的惊喜，欢迎留言讨论。

在线阅读

本文摘自慧智微微信公众号，扫描以下二维码，在线阅读

4.2 Wi-Fi 市场探析：今日应用与未来趋势

作者：彭洋洋 博士

在当今这个数字化时代，Wi-Fi 技术已成为连接世界的桥梁，它无声地渗透到我们生活的每一个角落，从家庭、办公室到公共场所，无所不在。随着科技的飞速发展，Wi-Fi 不仅承载着数据传输的基本功能，更在推动着社会进步和产业升级中扮演着举足轻重的角色。从最初的 Wi-Fi 标准到如今即将普及的 Wi-Fi 6/6E，乃至未来展望中的 Wi-Fi 7，每一次技术的迭代都为我们带来了前所未有的连接体验和无限可能。

在上一篇材料《Wi-Fi：连接未来，缔造美好生活》中，我们从技术角度分析了 Wi-Fi 协议的发展历史和应用定位，材料也得到了很多业界专家的讨论。在和 Wi-Fi 行业专家的讨论中，我们逐渐理解了在这个日新月异的市场中，Wi-Fi 技术的应用现状。也了解到 Wi-Fi 技术面临着的挑战与机遇，同时，也看到 Wi-Fi 技术未来塑造我们数字生活的发展趋势。于是，我们将收集到的信息做一个整理，从市场应用角度，探讨 Wi-Fi 技术的今日应用与未来趋势。

4.2.1 今日 Wi-Fi 应用概览

根据 Techinsights、TSR 等统计机构的统计数据，2023 年全年 Wi-Fi 终端的出货约 36 亿部。

4.2.1.1 应用市场维度

36 亿的终端出货分别包含手机、路由器及 AP、平板电脑及 PC、智能设备等。随着智能手机、平板电脑、智能设备的普及，Wi-Fi 已成为这些设备连接互联网的主要方式。以下将从应用市场的维度，来看 36 亿 Wi-Fi 终端出货的市场构成。



图 4.18: Wi-Fi 终端市场分布

- **手机市场：**
  在智能手机中，Wi-Fi 已经成为不可或缺的一部分。2023 年全球智能手机出货约 12 亿部，其中，带有 Wi-Fi 功能的智能手机约为 11.3 亿部。Wi-Fi 已几乎已成为智能手机的标配。在功能机市场，也有少量的手机拥有 Wi-Fi 功能，根据 TSR 的统计，带有 Wi-Fi 功能手机约有 1,700 万部。数量上比智能手机少近两个数量级。在整个手机市场，2023 年约有近 11.5 亿部拥有 Wi-Fi 连接的手机出货。

- **路由器及 AP 市场：**
  在路由器及接入点（AP）市场中，Wi-Fi 应用同样占据核心地位。据统计，在 2023 年，企业级 AP、网关、路由器，以及家庭路由器和移动路由器市场，合计出货约为 5 亿部。

- **平板电脑及 PC 市场：**
  在平板电脑和 PC 市场，Wi-Fi 应用同样占据着至关重要的地位。平板电脑和移动笔记本电脑几乎已全部配备 Wi-Fi 功能，部分台式电脑也可对 Wi-Fi 进行支持。2023 年的出货中，约 1.6 亿部的平板电脑，1.8 亿部笔记本电脑，以及 6 千万的台式电脑具备 Wi-Fi 连接功能。

- **智能家居及智能办公：**
  在智能家居及智能办公领域，Wi-Fi 正发挥着日益重要的作用，应用也越来越广泛。智能家居主要包含家中智能设备，智能办公主要包含企业智能设备、打印机等。这些应用的设备在 2023 年达到约 3.8 亿部，并且未来可能会继续快速增长。

- **电视及播放器市场：**
  电视机虽然是一个传统市场，但现代电视中的智能应用功能普遍需要 Wi-Fi 模块进行联网，所以电视也成为 Wi-Fi 应用的重要市场。在 2023 年出货中，约 1.9 亿台电视带有 Wi-Fi 联网功能。除此之外，还有 9 千万带 Wi-Fi 功能的电视盒子、播放设备等出货。

- **网络音箱及摄像头：**
  在网络音箱及摄像头市场中，Wi-Fi 应用同样展现出强大的影响力和广阔的应用前景。2023 年出货的 Wi-Fi 设备中，约有近 1 亿部的智能音箱，及约 1.7 亿部的网络摄像头实现出货。

- **车载市场：**
  在车载应用中，Wi-Fi 的主要应用为娱乐及通信。其中娱乐应用的 Wi-Fi 出货量更多，约为 7,600 万部，用于通信的出货量近 2,000 万部。合计出货数量为近 1 亿部。

- **智能穿戴市场：**
  在智能穿戴市场，Wi-Fi 是提升用户体验和扩展设备功能的关键技术之一。可穿戴设备主要构成是智能手表。2023 年，约有 9,000 万部带有 Wi-Fi 功能的智能手表出货。

- **游戏及 VR 市场：**
  在游戏及 VR 中，Wi-Fi 技术的重要程度不言而喻，它直接关乎到用户体验的流畅度和沉浸感。2023 年，在此市场出货的 Wi-Fi 设备约 5,000 万部。其中游戏机仍然占主要部分，约 4,300 万部。VR 设备约 600 万部，仍在增长中。

- **其他市场：**
  作为一项通用无线连接技术，Wi-Fi 应用广泛，覆盖教育、医疗、零售、交通等多个领域。所以在除以上几个重要市场外，Wi-Fi 在其他市场应用中还有较大规模的出货。年出货数量约 3.5 亿部。

4.2.1.2 通信制式维度

在《Wi-Fi：连接未来，缔造美好生活》中，我们回顾了 Wi-Fi 技术的发展历史。其中按 Wi-Fi4/5/6/7 划分的通信制式便于推广和用户记忆，成为 Wi-Fi 推广中重要的命名方式。

在目前的 Wi-Fi 出货中，最主要的出货制式是 Wi-Fi4、Wi-Fi5、Wi-Fi6 及 Wi-Fi7。

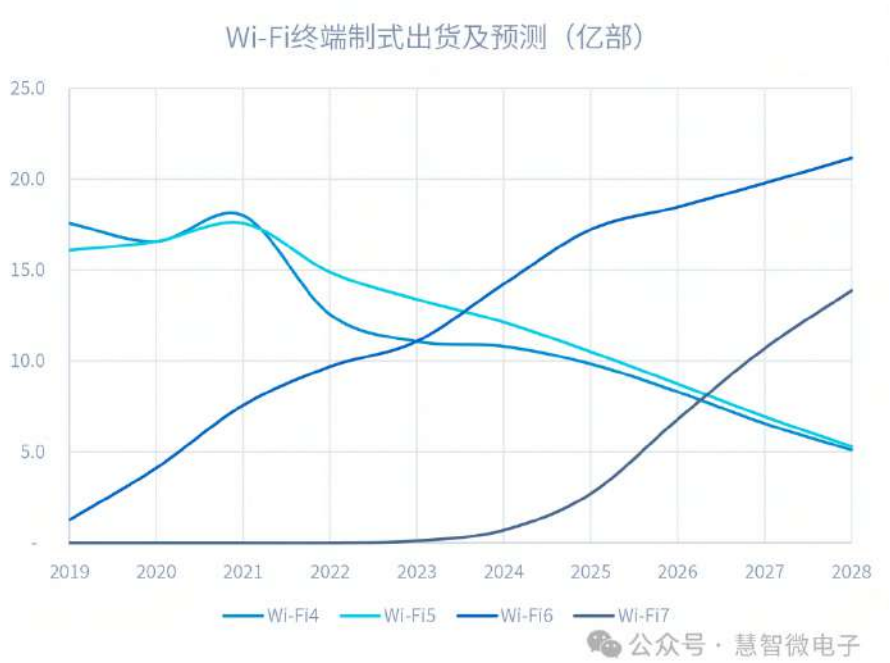


图 4.19: 不同制式的 Wi-Fi 终端出货数量及预测

- **Wi-Fi4/5：**
  Wi-Fi4 和 Wi-F5 做为早期、成熟的通信制式，在家庭和小型办公场所有着广泛的应用。随着智能手机的普及，高清视频、在线会议等应用的兴趣，Wi-Fi4/5 市场规模逐步扩大，占据在广泛的应用领域。虽然 Wi-Fi 技术不断发展，但在目前市场中，Wi-Fi4/5 仍有可观的出货数量。在 2023 年，Wi-Fi4 与 Wi-F5 终端出货数量分别达到 11 亿部及 13 亿部。

- **Wi-Fi6：**
  Wi-Fi6 作为新的通信制式，一经推出便迎来快速发展。随着 Wi-Fi 应用的日益广泛，以及物联网、智慧城市的发展，高密度、高并发、大速率的网络需求日益增长。Wi-Fi6 凭借其高效的频谱利用率、大带宽等特性，在这些领域展现出强大的市场竞争力，并将逐渐成为无线网络的主流制式。在 2023 年，Wi-Fi6 设备的出货数量将达到 11 亿部，并且预计在 2028 年超过 20 亿部。

- **Wi-Fi7：**
  作为 2024 年最新推出的通信制式，Wi-Fi7 也被视为未来无线网络市场的潜力股。未来网络需要更高速率、

更低延迟和更稳定的网络连接，Wi-Fi7 凭借其前所未有的高速率和低延迟特性，有望在这些领域开辟出全新的市场空间，推动无线网络市场进一步拓展。根据预测，Wi-Fi7 设备将以每年数亿部的增长速度快速增加，尽管 2024 年 Wi-Fi7 设备刚刚推出，但在 3 年之后的 2027 年，出货数量就有望超过 10 亿部。

4.2.2 Wi-Fi 市场未来机遇

过去近 30 年的发展过程中，Wi-Fi 技术快速发展，成为连接上百亿终端设备的通信制式。未来，Wi-Fi 市场仍充满无限发展机遇。

4.2.2.1 从市场维度看未来发展



图 4.20: Wi-Fi 终端年出货量预测

根据 TSR 等机构预测，随着新兴技术的不断涌现，Wi-Fi 作为无线网络的核心技术，出货数量也将随之增长。尤其在手机、路由器及 AP、智能家居及智能办公、网络音箱及摄像头，以及其他市场，Wi-Fi 都将迎来广泛的发展空间。

例如在增长迅速的智能家居及智能办公领域，年出货量将由 2023 年的 3.8 亿部，增长到 2028 年的 5.6 亿部。在其他领域，也将由 3.6 亿部，增长至 6.0 亿部。

4.2.2.2 6GHz 新频段 Wi-Fi 的增长

在 Wi-Fi 频段中，自 Wi-Fi6E 起就引入了 6GHz 新频段。在 Wi-Fi7 中，6GHz 也成为重要频率基础。

6GHz 频段是指位于 5.925 ~7.125GHz 频率范围中的部分，由于带宽宽、干扰少，可以带来 Wi-Fi 设备的速率提升、容量增加以及延迟降低。

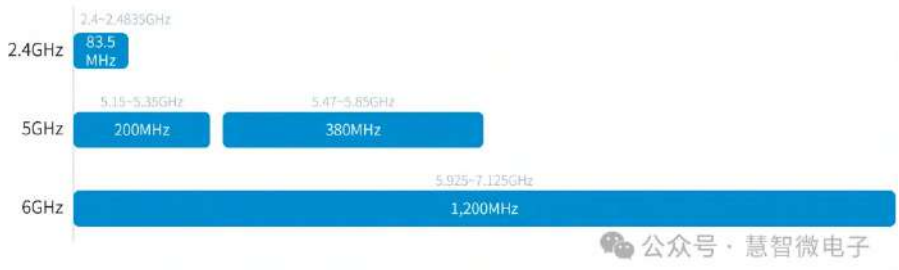


图 4.21: Wi-Fi 使用的全球主要国家和地区 ISM 及非授权频段

虽然 6GHz 频段在不同区域的法律法规中还存在差异。比如说北美在积极的推动 6GHz 在 Wi-Fi 等非授权协议中的应用，而中国希望这部分频段应用于蜂窝网络，欧洲则是将频段一分为二，在 5.925 ~6.425GHz 的 500MHz 给 Wi-Fi 使用，6.425 ~7.125GHz 的 600MHz 给预留给蜂窝使用。

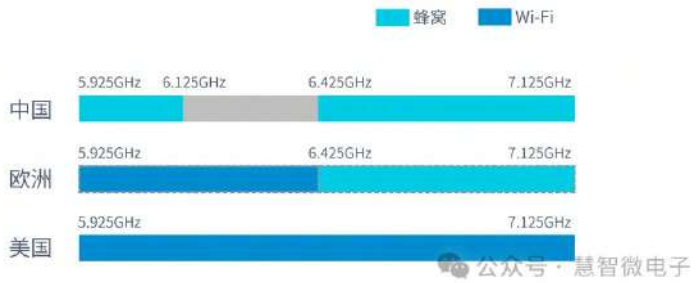


图 4.22: 6GHz 频段频率划分

法规的不收敛并不影响 6GHz Wi-Fi 的广泛使用。因为技术进步及市场需求的增加，6GHz 部分的 Wi-Fi 将持续增长。根据 TSR 的统计，6GHz Wi-Fi 终端在 2023 年出货约 2.2 亿部，随后将迎来快速增长，预计到 2028 年，这一数字将达到 7.7 亿部。

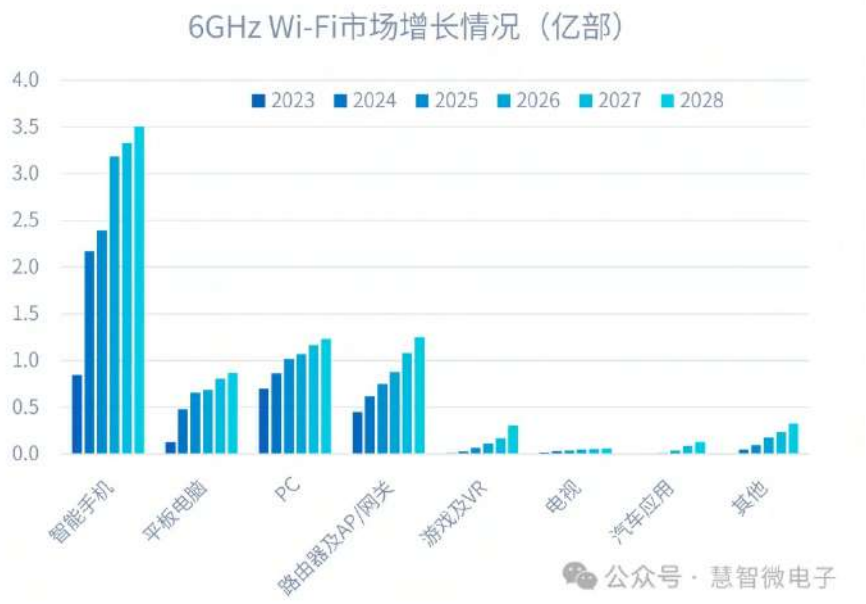


图 4.23: 6GHz Wi-Fi 市场增长情况

在 6GHz Wi-Fi 中，市场最大，增长最快的是智能手机市场。其次，平板电脑、PC、路由器及 AP/网关市场等，也都将迎来快速增长，到 2028 年时，达到年亿部出货量级

4.2.2.3 增长中的 Wi-Fi6/Wi-Fi7 市场

在关注 6GHz 新频段带来增长的同时，也不能忽视 Wi-Fi6/Wi-Fi7 本身也处于快速增长的阶段。如前文所述，作为最新的 Wi-Fi 协议，Wi-Fi6/Wi-Fi7 近几年发展快速，几乎每个应用中的 Wi-Fi6/Wi-Fi7 的终端出货数量都在强力增长。

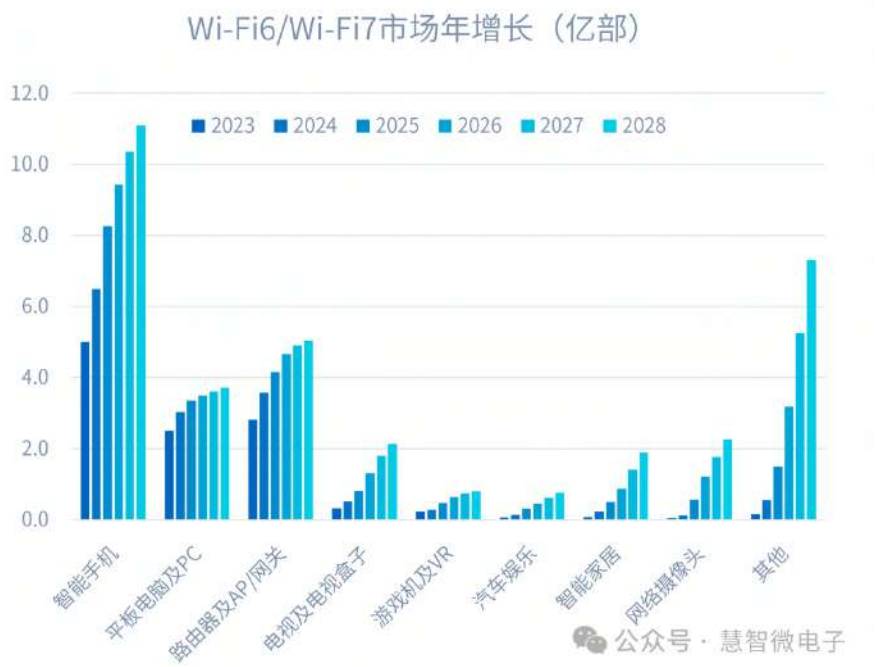


图 4.24: Wi-Fi6/Wi-Fi7 市场增长

具体而言，2023 年出货的智能手机中已有 4.5 亿部支持 Wi-Fi6 和 Wi-Fi7。这一趋势预计将持续，并且未来每年支持 Wi-Fi6 和 Wi-Fi7 的智能手机出货量将以数千万部的速度增长。按此增长速度推算，到 2028 年，支持 Wi-Fi6 和 Wi-Fi7 的智能手机出货量将突破 10 亿部。

在路由器和网关市场，Wi-Fi6 和 Wi-Fi7 设备的增长同样显著。2023 年，已有 2.4 亿部支持 Wi-Fi6 和 Wi-Fi7 的设备出货，而到 2028 年，这一数字预计将攀升至约 4.5 亿部，接近 2023 年的两倍。

此外，Wi-Fi6 和 Wi-Fi7 技术在电视、电视盒子、智能家居、网络摄像头等多个领域的应用也将迅速增长，进一步推动整个生态系统的蓬勃发展。

4.2.3 结论

在这个迅速发展的数字化时代，Wi-Fi 技术不仅仅是一种连接手段，更是推动社会进步和产业升级的重要引擎。通过当前应用和未来趋势，我们可以清晰地看到，Wi-Fi 正以其不可或缺的姿态渗透到各个领域，无论是智能家居、智能办公，还是车载、游戏、VR 市场，Wi-Fi 的广泛应用和未来发展前景都充满了无限可能。

从最初的 Wi-Fi 标准到如今的 Wi-Fi 6/6E，乃至未来的 Wi-Fi 7，每一次技术的迭代都为我们的生活带来了前所未有的改变。尤其是在 6GHz 频段的引入和 Wi-Fi 7 的推出下，未来的网络将变得更加高速、低延迟和稳定。这不仅满足了日益增长的数据需求，也为创新应用和新兴产业提供了坚实的基础。

总结来看，Wi-Fi 技术的发展不仅是技术进步的体现，更是我们追求更好生活品质的重要标志。随着技术的不断演进，Wi-Fi 将继续在连接世界、推动进步的道路上，发挥着不可替代的作用。

您认为 Wi-Fi 技术有哪些未来趋势？欢迎留言讨论。

在线阅读

本文摘自慧智微微信公众号，扫描以下二维码，在线阅读



4.3 Wi-Fi 系统：技术方案与实现路径

作者：彭洋洋 博士

在当今的数字化时代，无线网络已成为连接世界的桥梁，而 Wi-Fi 技术则是这座桥梁中最为关键的基石之一。随着移动互联网、物联网（IoT）以及智能家居的迅猛发展，Wi-Fi 不仅肩负着数据传输的重任，更成为了现代生活中不可或缺的一部分，广泛应用于企业办公、家庭娱乐以及城市智能化管理等多个领域。在前文《Wi-Fi 市场探析：今日应用与未来趋势》中可以看到，Wi-Fi 的年出货数量已达到 36 亿，并且仍在持续增长。

然而，随着应用场景的日益多样化，Wi-Fi 系统也在不断变化以适应新的需求。在 Wi-Fi 的应用领域，并没有一种一统天下的方案，而是根据具体的应用场景来定制。本文将从应用场景的角度出发，对 Wi-Fi 系统方案进行梳理，以期对 Wi-Fi 方案有更深入的理解。

本来从以下几个方面，对 Wi-Fi 系统的技术方案与实现路径进行探讨：

1. Wi-Fi 市场划分：AP 和 Station
2. AP 侧的通信能力划分
3. Station 侧的数传与 MCU Wi-Fi
4. Wi-Fi 中的外置 FEM

4.3.1 Wi-Fi 市场划分：AP 和 Station

在深入探讨 Wi-Fi 系统方案之前，首先需要了解 Wi-Fi 市场中的两个核心概念：AP 和 Station。AP（Access Point，接入点）的功能是提供无线连接的接入点，在 Wi-Fi 网络中起到终端和以太网的桥梁作用。而 Station（有时简称为 STA）通常指终端设备，如智能手机、智能家居设备等。

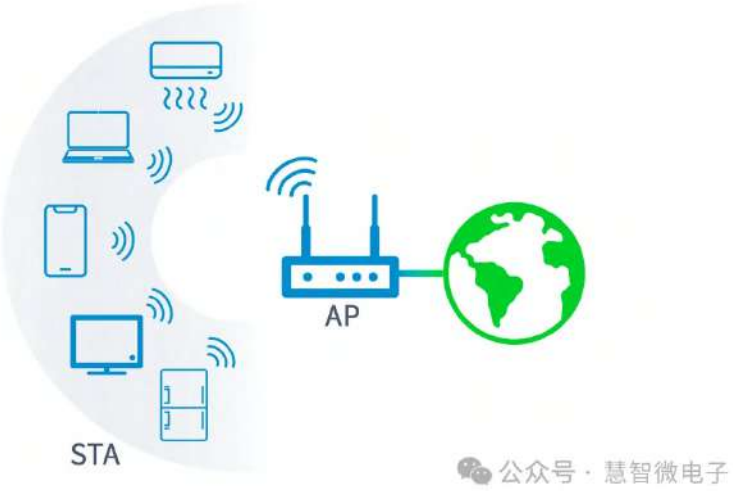


图 4.26: Wi-Fi 中的 AP 设备和 STA 设备

4.3.1.1 Access Point（AP）

Access Point，简称 AP，是提供无线连接服务的设备，通常指无线路由器或无线接入点。AP 在 Wi-Fi 网络中起到桥梁作用，它连接有线网络（如以太网）和 STA 设备通，从而实现无线设备与有线网络的互联。

AP 的核心功能包括信号覆盖、用户接入管理和数据包转发，通过无线信号覆盖一定范围内的区域。AP 一般允许多个 STA 设备同时连接，AP 还需要对这些设备进行管理，确保每个设备的连接稳定性。此外，AP 负责将来自 STA 的无线数据包转发到有线网络，并将来自有线网络的数据包转发给相应的 STA。

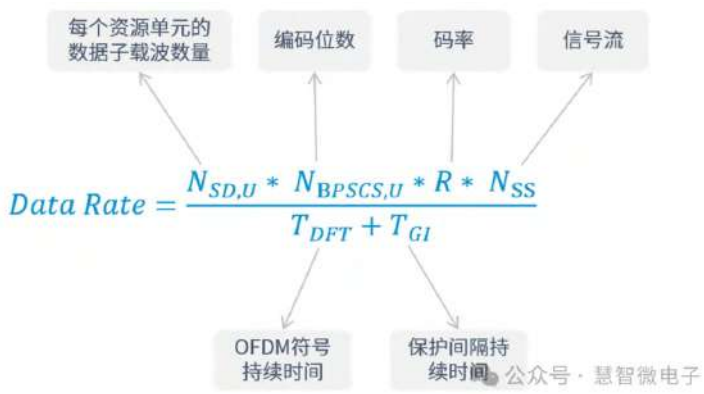


图 4.27: 不同通信协议覆盖范围及应用场景

在实际应用中，AP 设备不仅存在于家庭和办公室中，还被广泛应用于大型场馆、公共场所和智慧城市建设中。它们提供了广泛的无线覆盖，支持大量用户设备的同时接入，并通过智能化的管理功能提升网络的整体性能和用户体验。

4.3.1.2 Station（STA）

Station，简称 STA，通常指的是终端设备，如智能手机、笔记本电脑、平板电脑、智能家居设备等。STA 是连接 Wi-Fi 网络的用户设备，它通过无线信号与 AP 进行通信。STA 的主要职责是发送和接收数据包，从而实现互联网浏览、数据传输、流媒体播放等功能。

在现实生活中，STA 设备无处不在。无论是在家中使用的智能电视、智能音箱，还是在办公场所中使用的笔记本电脑和智能手机，都是 STA 设备。STA 设备通过 Wi-Fi 连接到 AP，实现了无缝的无线通信。这些设备的性能和体验直接影响用户对 Wi-Fi 网络的满意度。

与 AP 形态相对较为统一不同，STA 的形态展现出多种多样的特点。STA 作为网络中的终端设备，其形态可以根据具体的应用场景和需求进行灵活设计和调整。无论是智能手机、平板电脑、笔记本电脑等便携式设备，还是智能电视、智能家居设备等固定终端，都可以作为 STA 存在于网络中，并且随着未来万物互联的到来，STA 设备也将越来越丰富。这种多样性使得 Wi-Fi 网络能够支持广泛的应用场景，从家庭娱乐到工业控制，从办公自动化到智能城市。

4.3.2 AP 侧：按通信能力划分

在 Wi-Fi 网络中，AP 设备的能力差异显著，不同的 AP 设备能够提供不同的通信能力和性能。这些能力通常以速率（如 AX1800、AX3000、BE6400）来表示。这种分类方式主要基于 AP 设备支持的最高速率标准，以帮助用户快速了解设备性能并选择适合的 AP 设备。

在这种命名中，前面两个字母如 AX 表示 AP 设备支持的 Wi-Fi 标准。比如 AC 指 Wi-Fi5 设备，AX 指 Wi-Fi6 设备，BE 指 Wi-Fi7 设备。后面的数字表示设备的理论最大速率，单位为 Mb/s。

比如 AX1800，表示支持 Wi-Fi6 的 AP 设备，这个设备可支持的最大通信速率是 1,800Mb/s。BE6400 表示支持 Wi-Fi7 的 AP 设备，这个设备可支持的最大通信速率是 6,400Mb/s。

在 AP 的系统设计中，不同的通信能力对于不同的方案实现方式。越大的速率对应于越大的通信带宽，以及更多的发射、接收空间流。以 Wi-Fi6 为例，理论速率可用以下公式计算：

Wi-Fi6 的参数如下：

路由器	最高支持协议	2.4GHz 频段			5.8GHz 频段			合计最大速率 (Mb/s)
		信道带宽	天线流数	最大速率 (Mb/s)	信道带宽	天线流数	最大速率 (Mb/s)	
AC1200	Wi-Fi5 (802.11ac)	40	2	300	80	2	867	1,167
AC1900		40	4	600	80	3	1,300	1,900
AX1800	Wi-Fi6 (802.11ax)	40	2	574	80	2	1,201	1,775
AX3000		40	2	574	160	2	2,402	2,976
AX5400		40	2	574	160	4	4,804	5,378
BE3600	Wi-Fi7 (802.11be)	40	2	689	160	2	2,882	3,571
BE6400		40	2	689	160	4	5,765	6,454
BE7200		40	4	1,377	160	4	5,765	7,142

公众号 · 慧智微电子

图 4.30: 不同能力的 AP 路由器的设计方案

协议	Modulation	N_{BPSCS}	R	N_{SD}				T_{DFT}	T_{GI}		
				20MHz	40MHz	80MHz	160MHz		Short	Medium	Long
Wi-Fi 6 802.11ax	BPSK	1	1/2	234	468	980	1960	12.8 μs	0.8 μs	1.2 μs	3.2 μs
	QPSK	2	1/2 & 3/4								
	16 QAM	3	1/2 & 3/4								
	64 QAM	6	1/2 & 2/3 & 3/4								
	256 QAM	8	2/3 & 5/6								
	1024 QAM	10	3/4 & 5/6								

公众号 · 慧智微电子

图 4.28: Wi-Fi6 的参数

基于以上参数，可以计算出 Wi-Fi6 不同配置在下最大速率。

802.11ax中最大理论速率计算			
模式	最大理论速率 (Mb/s)	发射/接收天线	最大理论速率计算 (Mb/s)
1x1 20MHz	143	1Tx / 1Rx	234 x 10 x (5/6) x 1 / (12.8+0.8) ≈ 143
2x2 20MHz	287	2Tx / 2Rx	234 x 10 x (5/6) x 2 / (12.8+0.8) ≈ 287
1x1 40MHz	287	1Tx / 1Rx	468 x 10 x (5/6) x 1 / (12.8+0.8) ≈ 287
2x2 40MHz	574	2Tx / 2Rx	468 x 10 x (5/6) x 2 / (12.8+0.8) ≈ 574
1x1 80MHz	601	1Tx / 1Rx	980 x 10 x (5/6) x 1 / (12.8+0.8) ≈ 601
2x2 80MHz	1,201	2Tx / 2Rx	980 x 10 x (5/6) x 2 / (12.8+0.8) ≈ 1,201
1x1 160MHz	1,201	1Tx / 1Rx	1,960 x 10 x (5/6) x 1 / (12.8+0.8) ≈ 1,201
2x2 160MHz	2,402	2Tx / 2Rx	1,960 x 10 x (5/6) x 2 / (12.8+0.8) ≈ 2,402

公众号 · 慧智微电子

图 4.29: Wi-Fi6 不同配置在下最大速率

根据以上速率，可以对应算出不同能力的 AP 路由器的设计方案。

使用这种命名方式，可以直观地了解 AP 设备的具体方案，例如天线数量和带宽等。以下是几种常见通信能力的 AP 设备示例及其特点：

AX1800

- Wi-Fi6
- 频段：2.4GHz 和 5GHz
- 4 天线（每个频段 2x2 MIMO）
- 最大带宽：2.4GHz 40MHz，5.8GHz 80MHz

- 总速率：约 1800Mb/s（1201Mb/s 在 5GHz + 573.6Mb/s 在 2.4GHz）
- 适用场景：家庭或小型办公室，满足普通上网和流媒体需求。

AX3000

- Wi-Fi6
- 频段：2.4GHz 和 5GHz
- 4 天线（每个频段 2x2 MIMO）
- 最大带宽：2.4GHz 40MHz，5.8GHz 160MHz
- 总速率：约 3000Mb/s（2402Mb/s 在 5GHz + 573.6Mb/s 在 2.4GHz）
- 适用场景：中型办公室或家庭，支持更多设备和高带宽应用。

BE3600

- Wi-Fi7
- 频段：2.4GHz 和 5GHz
- 4 天线（每个频段 2x2 MIMO）
- 最大带宽：2.4GHz 40MHz，5.8GHz 160MHz
- 总速率：约 3600Mb/s（2882Mb/s 在 5GHz + 688Mb/s 在 2.4GHz）
- 适用场景：大型办公室、公共场所或高密度环境，支持大量设备和超高带宽需求。

由于上述命名方式能够清晰地展现出 Wi-Fi AP 路由器的性能特点，因此它已成为路由器行业广泛采用的命名惯例。以 TP-Link 的 Wi-Fi 7 路由器命名为例，我们可以明显看到，“字母”+“数字”的组合已成为标识不同路由器能力的主要命名方式。



图 4.31: TP-Link 的 Wi-Fi7 无线路由器命名

S



图 4.32: 数传 Wi-Fi 与 MCU Wi-Fi 应用场景

4.3.3 TA 侧：数传 Wi-Fi 与 MCU Wi-Fi

在 STA 侧（Station 侧），我们通常将其应用分为两大场景：数传 Wi-Fi 和 MCU Wi-Fi。数传 Wi-Fi 主要用来传数据，MCU Wi-Fi 主要用来传指令。

4.3.3.1 数传 Wi-Fi

数传 Wi-Fi，顾名思义，其核心任务是传输数据。作为 Wi-Fi 的基础应用和基本功能，数传 Wi-Fi 在无线通信领域扮演着至关重要的角色。

数传 Wi-Fi 是 Wi-Fi 的基础应用和基本功能，其主要功能是完成数据传输。衡量数传 Wi-Fi 性能的标准包括带宽、延迟、吞吐量和稳定性。数传 Wi-Fi 的典型应用场景包括视频流、文件传输、在线游戏、虚拟现实等。它对带宽和延迟有较高要求，因此在技术方案上注重以下几个方面：

- 高带宽
- 低延迟
- 安全性
- 覆盖范围

由于数传 Wi-Fi 主要的功能是完成数据传输，所以通常对传输速率有要求，比如 Wi-Fi 数传设备甚至可能会使用 AX3000 标准，采用 160MHz 的最大带宽，每个频段 2x2 的 MIMO，提供高达 3000Mb/s 的速率，从而满足大数据量的传输需求。

在方案设计上，数传 Wi-Fi 也通常会根据传输速率的不同而选择不同的方案设计。比如 AP 通信能力、信号带宽、天线架构等。另外，为了达到安全的数据传输，通常也会在加密设计、认证机制方面做更多的设计。

4.3.3.2 MCU Wi-Fi

随着万物互联的到来，Wi-Fi 不仅用于高速数据传输，还应用于还需要满足物联网（IoT）应用中一些控制指令的传输，这些应用不需要很高的传输速率，也不需要非常强大的 CPU 来处理复杂的数据传输问题。于是就出现了 MCU Wi-Fi 产品。

不同于数传 Wi-Fi 用于传输数据，MCU Wi-Fi 则更多地被用于传输指令。在物联网应用中，Wi-Fi 技术不仅需要完成基础的数据传输任务，还需要与更多的控制功能及应用场景相结合。MCU Wi-Fi 产品应运而生，它将微控制单元（MCU）与 Wi-Fi 技术相结合，不仅具备 Wi-Fi 通信功能，还能进行复杂的数据处理和控制操作。这种产品在设计上更加注重成本控制、低功耗以及易于集成和使用，特别适用于工业自动化、智能农业、智能安防等应用场景，能够实现设备的智能控制、远程监控以及数据采集与分析。

MCU Wi-Fi 芯片的设计思路主要是为了在保证低功耗和高集成度的同时，提供足够的计算能力和通信能力，以满足智能家居、可穿戴设备等物联网设备的需求。通过优化芯片内部的各个模块，并采用先进的电源管理技术，MCU Wi-Fi 芯片能够实现高效的无线通信和稳定的设备控制。

在 MCU WiFi 领域，中国厂商展现出了其独特的优势与擅长之处。凭借着在物联网技术方面的深厚积累与不断创新，中国厂商在 MCU WiFi 产品的研发与生产上取得了显著成就。他们不仅注重产品的性能与稳定性，还

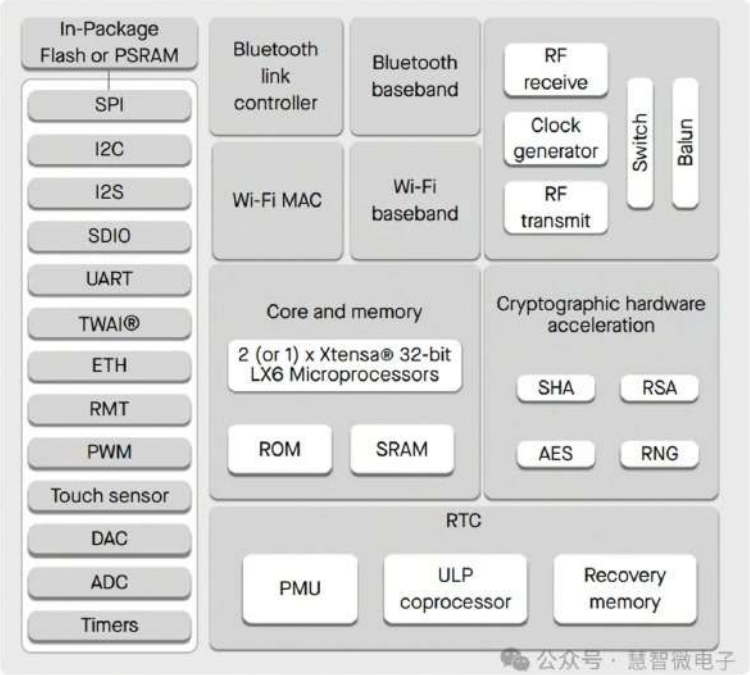


图 4.33: 乐鑫公司的 Wi-Fi SoC 产品

致力于优化成本控制，使得 MCU WiFi 产品更加符合市场需求。此外，中国厂商在供应链管理与生产效率方面也具备显著优势，能够快速响应市场变化，为客户提供定制化的解决方案。这些优势使得中国厂商在 MCU WiFi 领域脱颖而出，为全球物联网技术的发展贡献了中国智慧与力量。

4.3.4 Wi-Fi 中的外置 FEM

随着 Wi-Fi 标准的不断演进，从 Wi-Fi 5 到 Wi-Fi 6，再到新兴的 Wi-Fi 7，传输速度、覆盖范围以及信号稳定性都得到了显著提升。这一系列的进步背后，不仅离不开 Wi-Fi 处理器的提供，也离不开一个关键组件的默默支撑——外置射频前端模组（FEM）。

外置 FEM 作为 Wi-Fi 系统中的核心部件，对信号的发送与接收起着至关重要的作用。它不仅能够增强信号的传输功率，扩大覆盖范围，还能有效降低噪声，提高接收灵敏度。随着在 Wi-Fi6、Wi-Fi7 中，高频段、多通路、高功率 Wi-Fi 应用的普及，信号衰减和传输距离问题日益凸显，外置 FEM 的重要性也愈发显著。

4.3.4.1 外置 FEM 的介绍及功能

外置 FEM（前端模块，Front-End Module）在 Wi-Fi 技术中扮演着至关重要的角色。外置 Wi-Fi FEM 主要由功率放大器（PA）、低噪声放大器（LNA）、天线开关构成，部分 Wi-Fi FEM 还会集成滤波器等组件。其主要功能包括：

- **信号放大：**
  功率放大器增强发射信号的强度，确保信号在长距离传输时仍能保持足够的强度，从而扩大 Wi-Fi 的覆盖范围。
- **噪声抑制：**
  低噪声放大器在接收信号时有效降低噪声，提高信号接收的灵敏度，进而提升通信质量。
- **频段选择：**
  天线开关和滤波器用于在不同频段之间切换和过滤干扰信号，确保 Wi-Fi 设备在不同频段下的正常工作，并减少相互干扰。

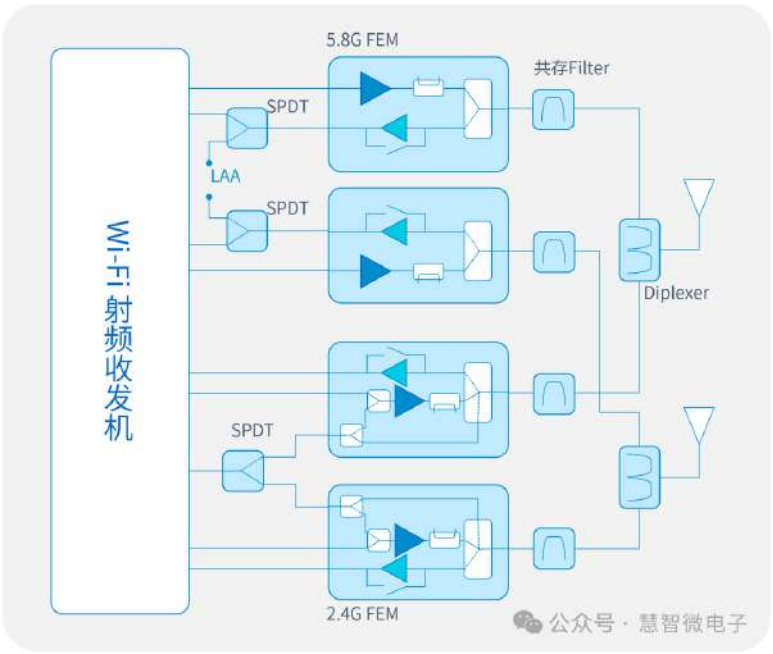


图 4.34: Wi-Fi 系统中的外置 FEM

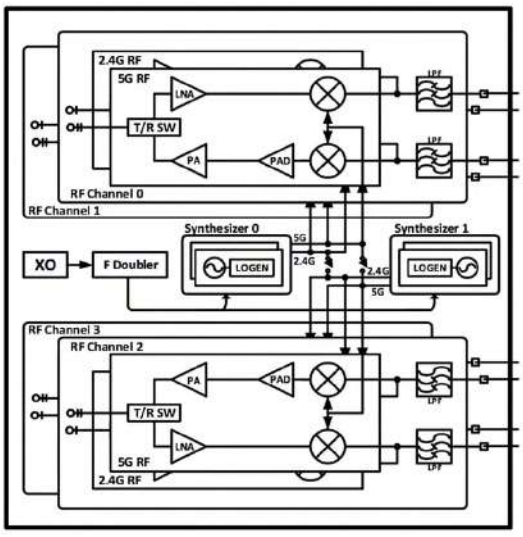
外置 FEM 的综合运用可以显著提升 Wi-Fi 设备的性能，包括更大的覆盖范围、更高的传输速率和更好的信号稳定性。

4.3.4.2 内置 Wi-Fi FEM

虽然外部 FEM 可以明显提高射频性能，但考虑到功耗、成本、设计复杂度，在一些应用场景中，外置 FEM 并不是必须。在方案应用中，一般有外置 FEM 与内置 FEM 两种方案。

和外置 FEM 不同，内置 FEM 通常集成在 Wi-Fi 芯片内部，相较于外置 FEM，它具有更高的集成度和紧凑的设计。内置 FEM 也由功率放大器（PA）、低噪声放大器（LNA）和部分滤波器等组件组成，只不过采用与 AP、Transceiver 共同的半导体工艺，在一个半导体晶圆上实现。

下图为 MTK 的 Wi-Fi6 Transceiver 产品框图及设计指标，以及版图设计。Transceiver 包含内部的 PA、LNA 以及开关。可以看到，MTK 的 Wi-Fi6 产品输出功率达到 20dBm，已经可以达到一些场景的应用需求。



		This work
WLAN Standards		4x4 11abgn/ac/ax
Process (nm)	CMOS	55
Integrated TRSW	2.4G	Yes
	5G	Yes
RF Bandwidth	MHz	160
TX Psat (dBm)	2.4G	28.5
	5G	29
TX Pout 2.4G (dBm)	54M @ -28 dB EVM	21.5
	MCS11 40M @ -35dB EVM	20.0
TX Pout 5G (dBm)	54M @ -28 dB EVM	21.5
	MCS9 80M @ -32dB EVM	20.8
	MCS11 160M @ -35dB EVM	20.0
TX EVM (dB)	2G MCS7 40MHz	-44.0 @ 0dBm
	2G MCS11 40MHz	-44.0 @ 0dBm
	5G MCS9 80MHz	-42.0 @ 0dBm
	5G MCS11 80MHz	-41.5 @ 0dBm
	5G MCS11 160MHz	-41.0 @ 0dBm
RX NF (dB)	2.4G	4.2
	5G	4.8
RX EVM (dB)	2.4G MCS11 40MHz	-48 @ -25dBm
	5G MCS11 80MHz	-45 @ -25dBm
RX Sensitivity (dBm)	2.4G, LGS4M	-79.5
	5G, LGS4M	-78.5
	5G, MCS9 80MHz	-66.0
	5G, MCS11 80MHz	-59.5
RF power consumption (mW)	RX 2.4G	322
	TX 2.4G	4324 (4SS+1LO, 20dBm, MCS11)
	RX 5G	420 (4SS+1LO, MCS11)
	TX 5G	4850 (4SS+1LO, 20dBm, MCS11)
PLL/LO IPN (degree)	2.4G (10k to 10MHz)	0.105 @ 2.4GHz
	5G	0.154 @ 5.925GHz

公众号 · 慧智微电子

图 4.35: MTK Wi-Fi6 Transceiver 设计框图



图 4.37: 慧智微产品开拓之路

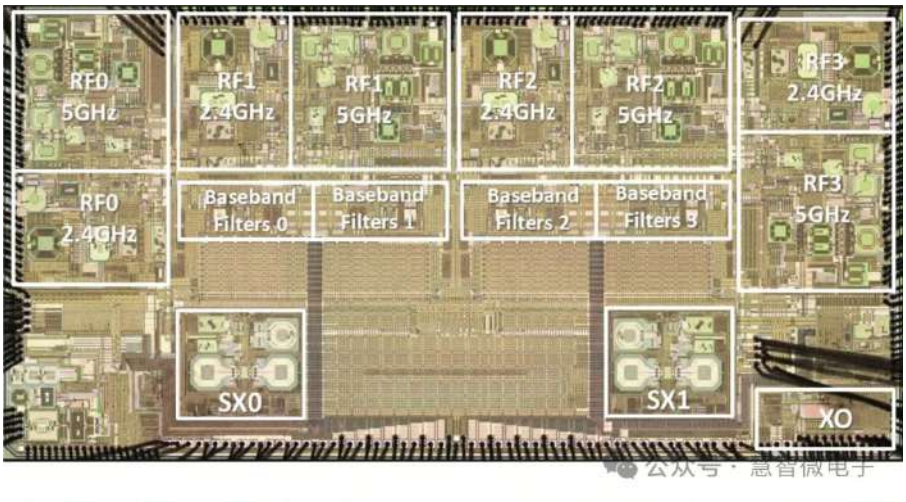


图 4.36: MTK Wi-Fi6 Transceiver 设计版图

4.3.4.3 外置 FEM 的选择和使用

在决定是否在 Wi-Fi 系统中采用外置 FEM 时，需要综合考虑多个因素进行考量。

首先是应用场景的需求，比如企业级 Wi-Fi 网络或智能家居系统通常需要更大覆盖范围和更高数据传输速率，这种情况下**外置 FEM 的引入显得尤为重要**。

其次是设备功耗的问题，外置 FEM 会增加设备的功耗，因此对于便携式设备或依赖电池供电的物联网设备，需要在性能提升和功耗之间找到平衡。这时就会考虑选用功耗较低的外置 FEM，甚至选择不加外置 FEM 袜。成本也是一个重要的考量因素，外置 FEM 会增加硬件成本，因此在一些成本敏感的市场或产品中，可能会选择不采用外置 FEM 以降低整体成本。

所以在外置 FEM 的选择中，需要根据具体应用需求、功耗预算、成本控制和技术能力等方面进行全面评估，决定是否在 Wi-Fi 系统中使用外置 FEM。目前看来，在高性能 Wi-Fi 系统中，外置 FEM 成为必须的选择。另外在更为先进的 Wi-Fi 系统，如 Wi-Fi7 中，由于 Transceiver 及 SoC 采用更为先进的工艺进行设计，此时集成内置高性能射频前端成为挑战，也需要采用外置 FEM 来设计实现射频前端方案。

作为专注于高性能射频前端产品设计的公司，慧智微深知外置 FEM 在 Wi-Fi 系统中的重要性。针对路由器、手机、物联网等市场，慧智微开发了系列高性能、低功耗 Wi-Fi FEM 产品，涵盖 Wi-Fi 5/6/7 通信标准，可应用于高通、MTK、Realtek 等系列 Wi-Fi 系统平台。

慧智微 Wi-Fi FEM 产品采用先进的射频技术和优化的电路设计，实现了产品的高性能。无论是路由器要求的大范围信号覆盖，还是物联网设备需要的稳定数据传输，或是手机等便携设备追求的无缝连接体验，慧智微

的 Wi-Fi FEM 产品都能提供出色的性能表现。

通过不断的技术创新和严格的质量控制，慧智微 Wi-Fi FEM 产品已经在市场上赢得了良好的声誉。慧智微期待不断提升 Wi-Fi FEM 的射频性能，为客户带来更了的用户体验。**欢迎与慧智微销售联系，了解慧智微 Wi-Fi FEM 的更多信息。**

4.3.5 总结

在数字化时代的大潮中，Wi-Fi 技术作为连接世界的桥梁，其重要性与日俱增。在 Wi-Fi 过去 30 年的发展中，我们不仅见证了其在企业办公、家庭娱乐、城市智能化管理等多个领域的广泛应用，还深刻理解了其随应用场景变化而不断演进的灵活性与创新性。

从 AP 与 Station 的基本概念出发，我们认识到 Wi-Fi 技术在构建稳定、高效的无线网络中的不可或缺作用。AP 作为接入点，其通信能力的多样化（如 AX1800、AX3000、BE6400 等标准）为不同场景下的网络需求提供了精准匹配。而 STA 侧，数传 Wi-Fi 与 MCU Wi-Fi 的细分，则进一步展现了 Wi-Fi 技术在数据传输与控制指令传输中的双重优势。

尤为关键的是，外置 FEM 作为提升 Wi-Fi 系统性能的关键组件，其在增强信号传输、降低噪声、提高接收灵敏度方面的显著作用，为我们在设计实现路径时提供了重要射频技术支撑。通过合理选择与应用外置 FEM，可以在确保性能的同时，有效控制设备功耗与成本。

在此背景下，作为高性能射频前端产品设计的佼佼者，慧智微始终致力于通过技术创新实现射频前端产品的优势迭代。在 Wi-Fi 产品线，慧智微产品涵盖 Wi-Fi 5/6/7 全系列标准，广泛应用于路由器、手机、物联网等市场。慧智微将持续关注市场需求变化，不断优化产品设计，为全球客户提供更加优质、高效的解决方案。期待通过慧智微的不懈努力与卓越追求，为 Wi-Fi 技术的未来发展贡献更多力量。

在线阅读

本文摘自慧智微微信公众号，扫描以下二维码，在线阅读

4.4 Wi-Fi PA 和手机 PA，有什么不同？

作者：彭洋洋 博士

近年来，在无线通信领域，最吸引人注意的两个通信协议就是手机里的 5G 通信协议，和 Wi-Fi 中的 Wi-Fi 7 协议了。

5G 通信协议是人类有史以来发展最快的通信协议，自 2019 年正式商用以来，4 年之间已经实现了超过 15 亿的终端连接。2024 年，5G 即将进入 5.5G 时代，3GPP 将 R-18 之后的 5G 版本定义为“5G-Advanced”，业界也称之为 5.5G。预计在 5.5G 协议中，峰值速率将由 5G 的 20Gbps，提升至 40Gbps，用户体验速率将由 1Gbps 提升到 10Gbps。

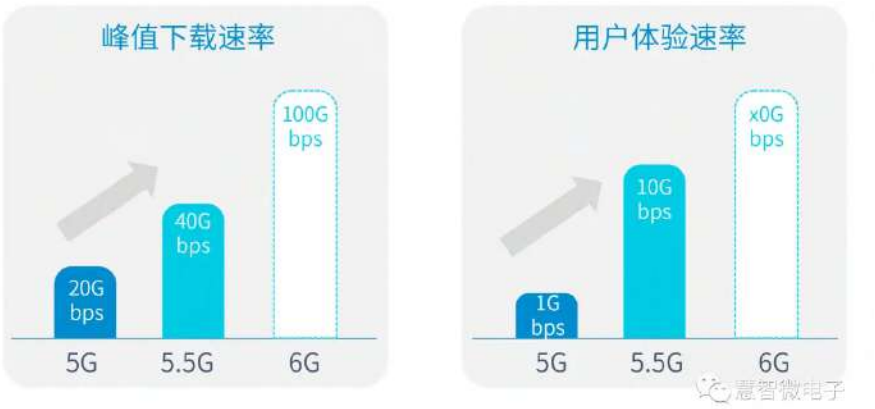


图 4.39: 预计 5.5G 的性能提升

2024 即将到来的不止有 5.5G，还有 Wi-Fi 7。虽然 Wi-Fi 7 协议的最终版本要到 2024 年才正式发布，但一些芯片厂商、路由器厂商已经摩拳擦掌，开始展示基于 Wi-Fi 7 技术的测试结果。2023 年 11 月 6 日，华为发布 Wi-Fi 7 AP 测试结果，称“面向全球市场，华为在美国 Tolly 实验室完成业界首个 Wi-Fi 7 AP 性能测试，整机性能超 13Gbps，单终端峰值性能超 4.33Gbps。此数据颠覆历史所有 Wi-Fi 性能测试结果，刷新最快速率记录。”[1]



图 4.40: 华为发布的 Wi-Fi 7 AP 测试结果

过去几年，手机蜂窝网络与 Wi-Fi 网络之间你追我赶，共同把手机变成具有强大通信能力的智能终端。二者仿佛有一个共同的目标，就是将手机的通信速率不断提升。两种通信协议也从最初的“瑜亮之争”，走到现在的互相补充、互相促进，甚至有些技术在互相趋同。手机蜂窝技术和 Wi-Fi 技术有哪些异同，二者所使用的 PA 又有什么一样或者不一样？

4.4.1 Wi-Fi 协议与蜂窝协议的异同

4.4.1.1 协议相同点 1：都是无线通信协议

在现代射频通信中，随着集成电路、信息论等技术的加持，现代射频系统的发展进入了快车道。现代射频系统的产业链已经非常复杂，依靠单一厂商已经很难独立完成某项射频通信技术。于是，射频通信的协议组织出现了。射频通信协议组织是一些专门负责制定和发布射频通信技术标准和规范的机构或者联盟，它们的存在是为了促进射频通信技术的发展和应用，保证不同的设备、系统、网络能够有效的交换信息，实现互联互通。

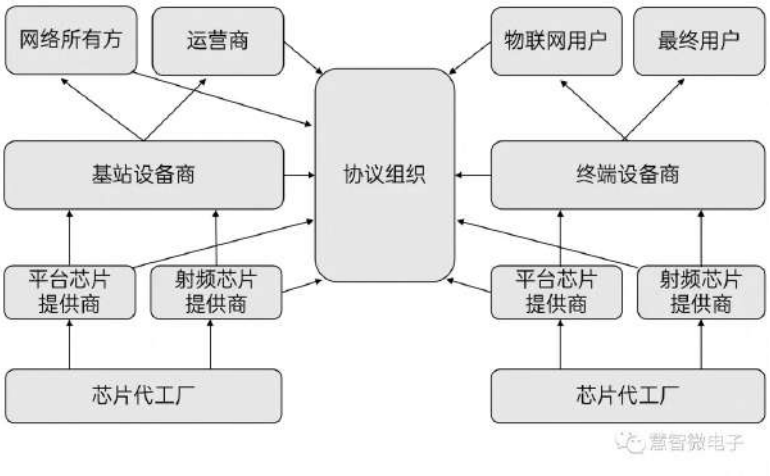


图 4.41: 现代射频通信的产业链协同

无线通信协议是一种规定了无线电波在空中传输数据的方式和格式的技术。无线通信协议有很多种类，其中最常见的就是 Wi-Fi 协议和手机蜂窝协议了。它们都可以让我们的手机、电脑、平板等设备无线地连接到互联网。

4.4.1.2 协议相同点 2：都实现了数十亿连接

无线通信协议不止有 Wi-Fi 与手机蜂窝，还包括蓝牙、NFC、UWB 等，在这些协议中，Wi-Fi 协议与手机蜂窝协议最为普及，均实现了数十亿的连接数。

连接数是指使用该协议的设备数量，它反映了该协议的普及程度和承载能力。根据 GSMA 的数据，2023 年全球蜂窝协议终端连接数目为 85 亿。根据 Strategy Analysis 的数据，在不包含手机及 PC 的情况下，2023 年 Wi-Fi 设备连接数已达到 35.32 亿台。二者都形成了充分完整的产业链。

产业链的完整支撑了行业的稳定发展，不论是 Wi-Fi 协议还是手机蜂窝协议，每年都有数十亿的设备出货。根据 TSR 的数据，智能手机年出货量接近 15 亿部，Wi-Fi 设备的年出货数量接近 40 亿部。



图 4.42: 全球终端设备出货

4.4.1.3 协议相同点 3：都实现了宽带连接

宽带通信是指单位时间内可以传输高的数据量，它反映了该协议的速度和效率。Wi-Fi 协议和手机蜂窝协议都具有很高的带宽，它们都可以提供高速的互联网接入。

例如，Wi-Fi 协议的最新标准 Wi-Fi 6 可以达到 9.6 Gbps 的理论速度，而手机蜂窝协议的最新标准 5G 可以达到 20 Gbps 的理论速度。这意味着我们可以用 Wi-Fi 协议和手机蜂窝协议来观看高清视频，玩在线游戏，进行视频通话，下载大文件等各种高流量的活动。均满足了人类日益提高的通信速率需求。

根据 ITU 的数据预测，全球数据用量在过去 10 年增长了 10 倍，在未来 10 年还将有 10 倍的增长。Wi-Fi 与蜂窝通信提供的大带宽特性，将有力支撑未来无线通信需求的演进。

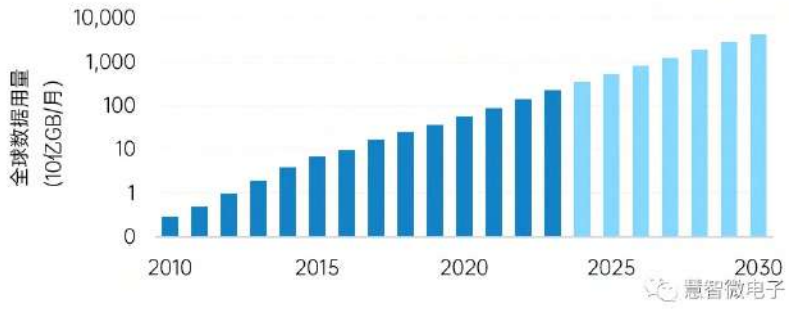


图 4.43: 全球数据用量

4.4.1.4 协议不相同点 1：协议组织不同

Wi-Fi 协议是由 IEEE（电气和电子工程师协会）制定的一系列标准，主要用于在局域网中实现无线接入。手机协议是由 ITU（国际电信联盟）和 3GPP（第三代合作伙伴计划）等组织制定的一系列标准，主要用于在广域网中实现无线通信。这两种协议的组织方有不同的目标和需求，因此它们的设计和实现也有所不同。

在产业链构成上，手机蜂窝协议产业链需要涵盖数公里范围覆盖的基站、终端设备，还需要涵盖能传输数千公里的核心网，其产业链更为庞大和复杂。也更加强调协议内设备的统一。由于整个产业链的复杂以及功能上的大一统，蜂窝协议的部署与演进经常与全球政策联系起来。比如全球蜂窝网络的部署节奏与频段划分会在联合国下属的 ITU（国际电信联盟）中讨论，各国的蜂窝网络部署也会在政府监管机构下统一进行。

而 Wi-Fi 协议是 IEEE 由有线以太网协议发展而来，最初是将有线以太网的物理层由有线改为无线实现，所以 Wi-Fi 协议关注在空中接口部分，有线传输网络仍然是基于原有的以太网。这也就使得 Wi-Fi 协议只需要关注在数百米范围之内的无线连接设备即可。协议组织的复杂度大大降低。但同时其对政府、政策等的影响力较蜂窝也明显下降。

4.4.1.5 协议不相同点 2：使用频段不同

不同的组织形式与在政策上的话语权，带来两种通信协议完全不同的两种运行方式，也决定了二者使用频段的不同。

Wi-Fi 协议使用的频段是 ISM 频段。ISM 是 Industrial Scientific Medical 的缩写，ISM 频段是由 ITU 定义的，用于工业、科学、医学研究使用的频段。这些频段是免费开放的，任何设备都可以使用。虽然免费开放意味着使用上的自由，更容易形成全球统一的标准，从而不需要像蜂窝网络一样切换运营商。但免费开放也意味着没有强制管理，也意味着会有其他设备同时使用这个频段，使用设备众多时也容易受到干扰和拥塞。比如在 2.4GHz 频段，微波炉、蓝牙等设备就会与 Wi-Fi 共享使用这个频段。

常见的 ISM 频段有：

- 2.400 ~2.4835GHz 频段（带宽 83.5MHz）
- 5.15 ~5.875GHz 频段（带宽 725MHz）
- 5.925 ~7.125GHz 频段（带宽 1.2GHz）

与 Wi-Fi 协议与其他设备“抢”频段不同，蜂窝协议在定义之初就会把频段确定下来。这些频段是由政府分配给运营商的，运营商需要付费使用，这些频段费用通常价格不菲，但是也可以保证服务的质量与覆盖范围。

频谱是非常宝贵的资源，国际运营商通常以拍卖的形式参与竞购。以 5G 频段为例，在 2019 年 5G 建设之初，德国电信用于拍卖 2GHz 附近的 40MHz 频谱就花费了 22 亿欧元。巨大的经济压力使得单一运营商无法购买连续的大带宽频谱，另外，由于无线频谱管理的限制，使得可用于蜂窝通信的频段十分分散，仅在 6GHz 以下频段，在 3GPP 的定义中，就有近百个用于 4G/5G 通信的频段，这些频段窄则几 MHz，宽则几百 MHz，散落在 6GHz 以下频段之中。

这些散落的频段也给手机射频设计提出了更复杂的要求。在中高端手机中，一般需要支持的频段达 30 个以上，于是“多频段多模式”的设计理念就开始在手机射频设计中出现，依靠单一射频器件对多个频段的支持，来减少射频电路设计的复杂程度。这给手机蜂窝射频器件的设计提出了更高的挑战。

频率复杂度的增加也催生了“载波聚合”技术。为了达到等同于宽带传输的效果，运营商要求将其所掌握的多个零散频段“聚合”起来，达到更高的传输速率。载波聚合技术使射频器件的复杂度明显提升。

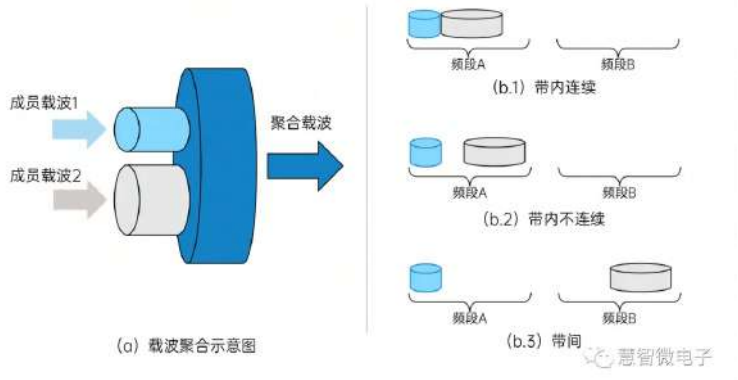


图 4.44: 载波聚合技术示意图

4.4.1.6 协议不相同点 3：部署理念不同

在两种协议组织运行的历史里，一直在讨论如何才能取代另外一个协议组织。不论是 2001 年 IEEE 就发起讨论的 WiMax 标准，准备将其作为 4G 的候选技术全球推广，还是 5G 到来之时，蜂窝协议雄心满满的认为手机有了 5G 就不需要连接 Wi-Fi。两个协议组织在过去 20 年里相爱相杀，互相竞争。

随着两种协议逐渐成熟，两种协议也均找到了自己的定位。二者有不同的部署和发展理念，人们越来越意识到，二者不是竞争关系，而是互补关系。



图 4.45: 互为补充的蜂窝协议与 Wi-Fi 协议

Wi-Fi 协议考虑的是灵活、低成本的局域网，它主要依靠无线路由器和接入点来提供服务，用户可以自由地连接和断开。据统计，在无线通信的使用场景中，85% 的时间发生在室内。Wi-Fi 协议很好地满足了这部分通信协议的需求，释放了蜂窝基站通信的压力。

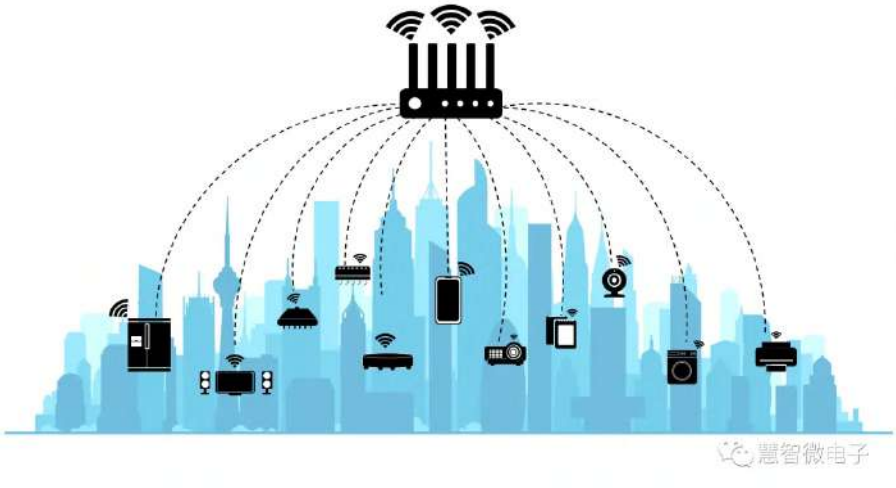


图 4.46: Wi-Fi 协议提供灵活的局域互联

手机蜂窝协议考虑的是统一、高可靠的广域网，它主要依靠基站和交换机来提供服务，并由运营商统一运营，用户付费之后，就可以在不同的区域和国家之间无缝地切换。手机蜂窝协议提供了可靠的、随时随地的互联，满足人类即时互联的需求。

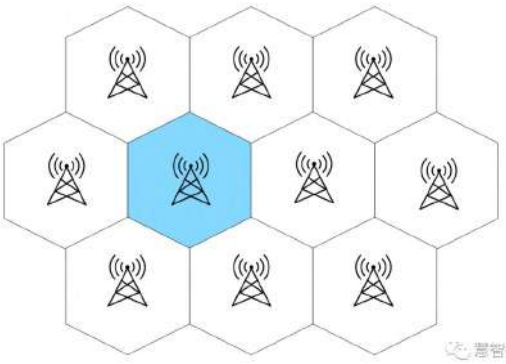


图 4.47: 蜂窝协议提供可靠的广域互联

4.4.2 Wi-Fi PA 与手机 PA 的异同

射频 PA（Power Amplifier，功放）是射频电路中的重要组成部分，对射频系统的性能有着重要的影响。其功率输出能力决定了设备的通信能力，其功耗的大小决定了设备的待机时长。

在 Wi-Fi 协议与手机蜂窝协议越来越趋同的发展中，Wi-Fi PA 和手机蜂窝 PA 设计也出现越来越多的相同点。

（注：尽管手机中也有可能包含 Wi-Fi PA，但作为蜂窝设备的代表，“手机 PA”一般指的是“手机中的蜂窝 PA”。）

4.4.2.1 PA 相同点 1：架构基本相同

Wi-Fi 系统是典型的 TDD 架构，收发通路在不同时间维度上切换，完成信号的收与发。

手机蜂窝早期是 FDD 架构，即收发通路在同一时刻都工作，但在频率维度上区分。在 4G 时代到来后，为了更有效地解决手机上下行能力需求不对称的问题，蜂窝网络引入 TDD-LTE 结构，与 FDD-LTE 共同构成 4G 手机蜂窝系统。到了 5G 之后，TDD 结构在手机蜂窝继续发扬光大，5G 的新频段，如 n77/78/79 等，均直接采用 TDD 结构。

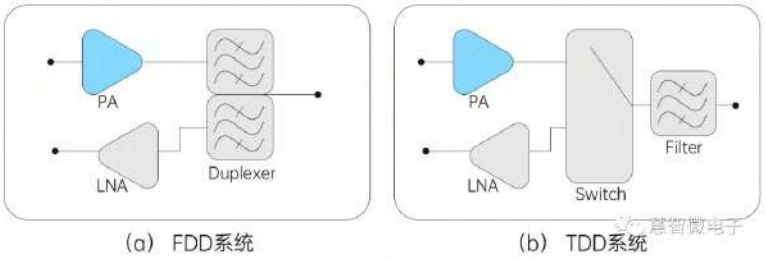


图 4.48: FDD 系统与 TDD 系统

由于均为 TDD 系统，Wi-Fi 与 5G 前端模组架构也逐渐趋同。下图为慧智微应用于 5G 蜂窝系统的模组产品 S55207-11 与慧智微应用于 Wi-Fi 系统的模组产品 S56540-11 产品架构比较。可以看到，二者均集成一路 PA 与一路 LNA，并且集成切换开关用于收发切换。不同的是，蜂窝系统中需要集成滤波器，并且需要根据系统需求集成 MIPI 控制器、定向耦合器等模块。

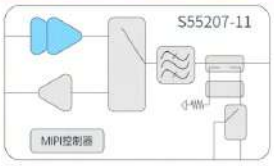
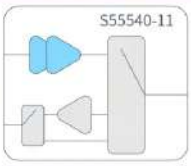
	5G n77前端模组	5GHz Wi-Fi模组
芯片型号	S55207-11	S56540-11
使用场景	5G蜂窝	Wi-Fi
架构框图		
集成模块	PA、LNA、开关、滤波器	PA、LNA、开关
工作频率	3.3-4.2GHz	5.15-5.85GHz
芯片尺寸	3x5 mm²	3x3 mm²

图 4.49: 慧智微 5G 蜂窝与 Wi-Fi 模组产品架构比较

4.4.2.2 PA 相同点 2：设计方法基本相同

随着协议的演进，Wi-Fi 系统与蜂窝系统均引入多 QAM、OFDM 等技术，以达到数 Gbps 的传输速率，在两种通信系统中，均需要用到对高阶调制友好的线性 PA 架构进行设计（有关 PA 的架构选择请参见《5G 射频 PA

架构》）。虽然在手机蜂窝 PA 与 Wi-Fi PA 中都在提 Doherty 架构等非线性架构，但是只是在线性度与效率之前做 trade-off 来进行效率提升，本质还是在以线性 PA 的设计思路在进行设计。

另外，Wi-Fi 通信能力的提升也要求 Wi-Fi PA 有更高的输出功率。虽然 Wi-Fi 并不需要向蜂窝网络一样传输数公里的距离，但高功率的 PA 有助于采用更高调制的信号，从而提升传输速率。在目前外置 Wi-Fi PA 模组中，PA 的功率能力与手机蜂窝 PA 能力基本相当。

对 PA 的性能、功率需求的相同，也使得二者的供应链逐渐趋同。目前在 Wi-Fi PA 设计中，绝大部分的 Wi-Fi PA 采用与手机蜂窝 PA 相同的 GaAs HBT 工艺进行设计（有关 PA 的工艺选择请参见《5G 射频芯片中的半导体》）。

	5G蜂窝PA	Wi-Fi PA
PA类型	线性PA	线性PA
PA饱和功率	30-38 dBm	28-35 dBm
PA主要架构	单端 Push-Pull Doherty ET	单端 Push-Pull Doherty
PA主要工艺	GaAs HBT	GaAs HBT

慧智微电子

图 4.50: 5G 蜂窝 PA 与 Wi-Fi PA 的对比

4.4.2.3 PA 相同点 3：供应商趋同

由于应用场景的接近，终端用户的统一，供应链的一致，手机蜂窝 PA 与 Wi-Fi PA 的供应商也开始趋同。手机蜂窝 PA 与 Wi-Fi PA 供应商趋同有诸多好处。比如：

- **技术复用：**
  厂商可以将蜂窝 PA 和 Wi-Fi PA 领域的技术经验和成果相互借鉴和应用，从而提高产品技术水平和创新能力。比如可以将蜂窝 PA 的高效率技术应用到 Wi-Fi PA 中，提高 Wi-Fi PA 的效率。也可以将 Wi-Fi PA 高线性度的技术应用到蜂窝 PA 中，提升蜂窝 PA 性能。

- **供应链复用：**
  虽然 Wi-Fi 终端一年数量有 30 亿部，但大部分 Wi-Fi 终端不含有外置 PA，并且由于 Wi-Fi 的频段数目远少于蜂窝，Wi-Fi PA 对半导体的用量仍明显小于蜂窝 PA。利用在蜂窝 PA 的规模优势，可以更容易建立起与供应商的合作，从而获得更优惠的价格和更快的交期，降低原材料和元器件的成本和风险。另外也可以通过共享和协调产测设备的方式，提高产品生产测试的效率和质量。

- **完整方案：**
  提供手机与 Wi-Fi PA 产品的厂商可以为客户提供更完整和更灵活的解决方案，从而满足客户在不同应用场景和需求下的无线通信需求。完整的解决方案可以使得客户在技术支持、商务支持中均得到收益。

在以上优势驱动下，射频厂商开始开发手机蜂窝与 Wi-Fi 的完整解决方案。国际厂商如 Skyworks、Qorvo、高通等厂商均同时提供手机蜂窝与 Wi-Fi 射频产品，国内厂商如慧智微等，也有手机蜂窝、Wi-Fi 产品线，提供全系列解决方案。

4.4.2.4 PA 不同点 1：方案不同

尽管 Wi-Fi 和 5G 蜂窝在架构上有很多类似之处，但还是有些细节造成二者方案有所不同。

Wi-Fi 与 5G 方案不同的最大来源就是使用频率的不同。5G 使用的是零散的频率，并且在 3G/4G 向 5G 演进过程中，还有不断的新频率加入进来。在手机需要“全球通”漫游的前提下，手机需要将这些频段均完美支持起来，并且还需要满足 4G/5G 双连接、载波聚合等额外的需求。这一切使得蜂窝系统极其复杂。在文章《除了调匹配，射频人还需要掌握系统知识》中，仅仅分析清楚一部 5G 手机中的射频系统，就已经不是易事，更不用说设计新的射频系统了。

相比之下，Wi-Fi 系统就简单许多。全球频率的归一使得方案中不需要对频率做太多考量；制式的向下兼容也简化了不同代制式的双连接问题。这都使得 Wi-Fi 系统比蜂窝系统要简化很多。

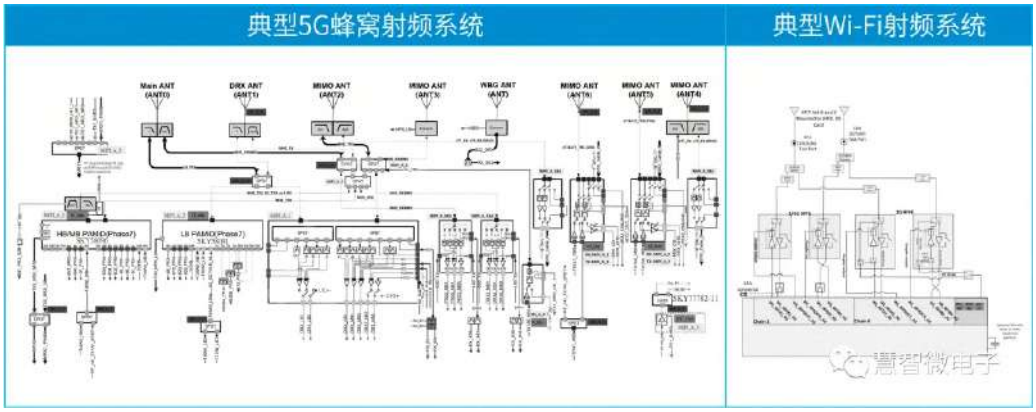


图 4.51: 典型蜂窝射频系统与 Wi-Fi 射频系统对比

4.4.2.5 PA 不同点 2：功率要求不同

与手机蜂窝网络中对手机功率有严格的限制不同，Wi-Fi 协议并没有对功率有严格的要求。这就给 Wi-Fi PA 设计带来了很大的自由度。尽管在高性能、高制式 Wi-Fi PA 设计中，功率甚至与手机蜂窝 PA 相近，但仍有大量的低端、近距离使用的 Wi-Fi 终端只使用收发机的内置低功率 PA 就可以完成传输。

Wi-Fi 对功率要求的不同也使得 PA 厂商开发出多个功率等级的 PA 供终端选择，以平衡功率与效率、成本之间的关系。在 Wi-Fi PA 中，一般根据功率的大小，分为高功率、中高功率、中功率、低功率 PA。另外，收发机内置的 PA 一般称为 IPA（Internal PA）。不同 PA 可覆盖的功率等级一般称法如下表（注：由于没有标准定义，此处只代表典型称法）。

Wi-Fi 6 FEM功率范围	级别分类
≥21dBm	高功率
18~20dBm	中高功率
16~18dBm	中功率
15~16dBm	低功率
≤15dBm	IPA（内置PA）

图 4.52: Wi-Fi PA 的功率级别（MCS11，-43dB EVM）

4.4.2.6 PA 不同点 3：线性度要求不同

尽管手机蜂窝 PA 与 Wi-Fi PA 都是线性 PA，但由于编码方式不同，造成对线性度的需求也不同。

比如在 Wi-Fi 6 中，最高采用 1024-QAM 调制，一个符号可承载 10bit 信息。在 Wi-Fi 7 中，最高将采用 4096-QAM 调制，一个符号可承载的信息提升到 12bit。而在 5G 蜂窝协议中，最高只是采用 256-QAM。

更高阶的调制对 EVM 提出了更高的要求，要求信号误差向量要足够小，这样才不会出现信号位的错误。这也就是在 Wi-Fi 系统会更关注 EVM，而在蜂窝系统更关注 ACLR（Adjacent Channel Leakage Ratio，相邻频道泄漏比）的原因。

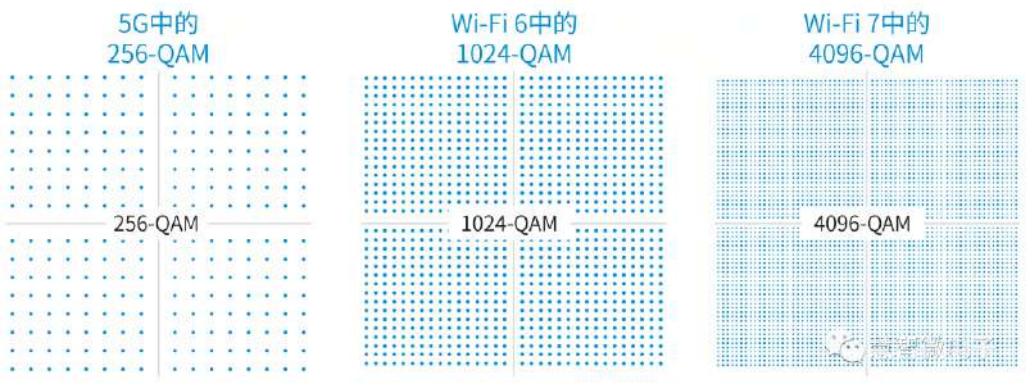


图 4.53: 不同协议中的编码要求

说 Wi-Fi 对于 EVM 的高要求还体现在 DEVM 指标上，DEVM 的全称是 Dynamic EVM（动态 EVM），它反映了 Wi-Fi 信号在发射过程中由于 PA 的瞬态热响应而引起的失真程度。由于 PA 的稳态建立需要一定的时间，在这个稳定时间到来之前，不希望系统传输有用信息。而 Wi-Fi 系统的帧结构复杂，在头部的 Preamble、header 帧中带有同步符号、样本信号、传输参数等信息，这就需要 PA 快速建立，用来响应帧头部的重要信息。DEVM 指标的特殊要求也是 Wi-Fi PA 与手机蜂窝 PA 的重要不同点。

以上就是手机蜂窝 PA 和 Wi-Fi PA 的主要异同点，总结如图4.54所示。

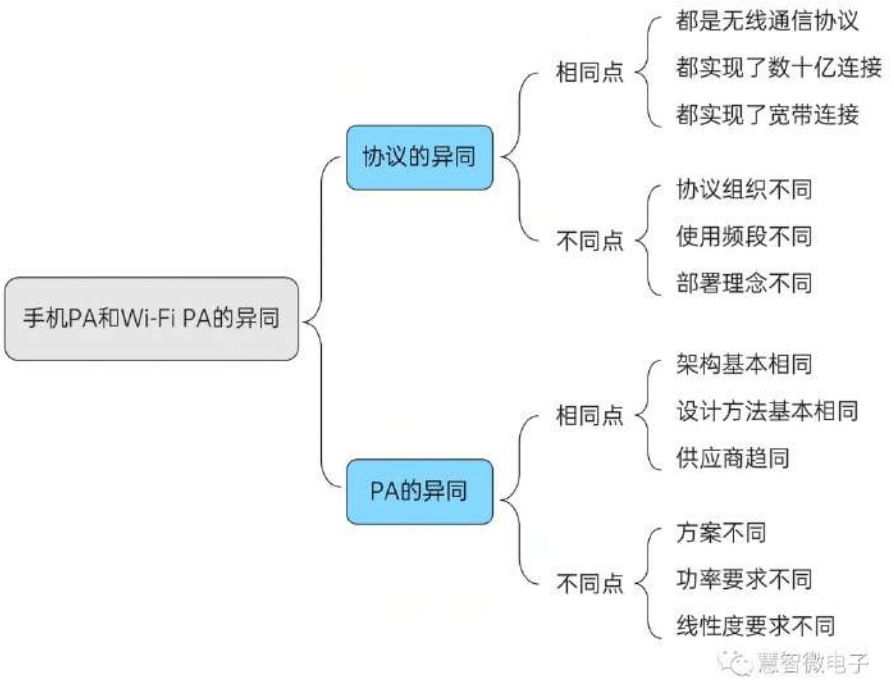


图 4.54: 手机 PA 与 Wi-Fi PA 的异同点

4.4.3 手机 PA 与 Wi-Fi PA 的未来趋势

由于产业历史和技术演进原因，手机蜂窝 PA 和 Wi-Fi PA 在过去的发展中较为独立。但随着技术的演进、协议的相互借鉴、系统的相互融合、产业链的逐步发展，手机蜂窝 PA 和 Wi-Fi PA 的设计也在逐步趋同。在未来发展中，可能有以下几个趋势：

- **共同向高性能发展：**
  手机蜂窝 PA 和 Wi-Fi PA 都将向高性能的方向发展，以满足用户对于无线通信的更高的要求，比如更高的数据速率、更低的延迟、更好的覆盖和更强的抗干扰能力。为了实现高性能，手机 PA 和 Wi-Fi PA 都需要

采用更先进的技术和材料，比如高线性度的电路和算法、高集成度的封装和模块、高稳定性的温度和电压控制等。

- **产业链逐步趋同：**
  手机蜂窝 PA 和 Wi-Fi PA 的产业链将逐渐趋同，以实现更好的协作和技术演进，为客户提供更为完整的解决方案。无论在芯片制造，还是设计公司提供的完整方案上，手机蜂窝 PA 和 Wi-Fi PA 的产业链都将逐渐趋同。

- **射频链路复用：**
  虽然目前手机蜂窝 PA 和 Wi-Fi PA 分属不同的制式，被不同的平台 SoC 及射频收发机所管辖，但随着未来技术的发展，平台 SoC 通信协议的整合，手机蜂窝 PA 和 Wi-Fi PA 在射频链路上有可能出现整合和复用，以简化射频系统设计。

射频链路的整合同样有助于应对日后复杂的频率规划。比如在最新的 6 ~7GHz 频段，不同国家和地区给出了不同的规划，中国目前将此频段范围均规划给蜂窝使用，而美国将此频段都规划给了 Wi-Fi，欧洲则是可能各分一半。频段划分的不统一给射频系统设计带来极大的复杂性。如果手机与 Wi-Fi 链路统一，就可以从容应对频谱划分中的不确定性。

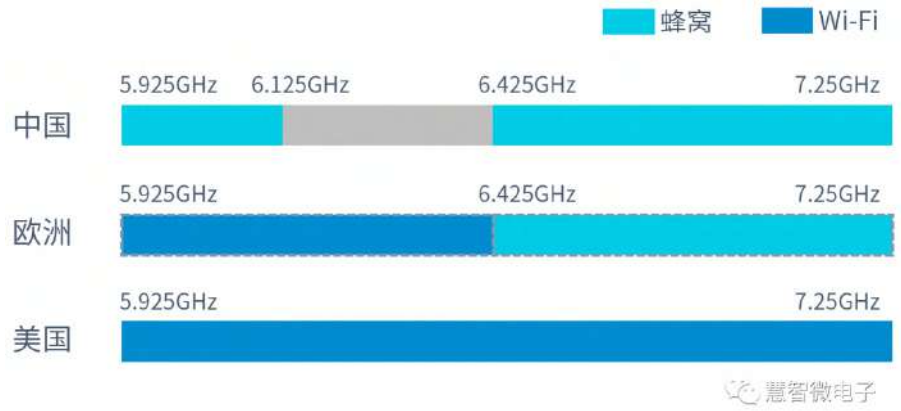


图 4.55: 6 ~7GHz 频段的初步划分

4.4.4 总结

手机蜂窝与 Wi-Fi，这两个发源于上世纪无线通信大爆发时代的协议在互相竞争中互相促进，共同快速成长。经过几十年的发展，两种通信协议汇集到我们手中的手机之中。成为我们离不开的技术。

在手机蜂窝协议和 Wi-Fi 协议发展过程中，二者既合作又竞争。它们都致力于提高无线通信的速度和质量，安全性和可靠性。它们之间的竞争促进了无线通信技术的创新和进步，比如蜂窝网络中引入的 MIMO、OFDM 等技术来自于 Wi-Fi 的最早使用，而 Wi-Fi 7 中也将借鉴蜂窝网络的“载波聚合”技术来实现多频连接。

在 PA 侧也是一样，协议的趋同使得两种场景中工作的 PA 也有大量的设计方法可以互相借鉴；产业链的趋同使得芯片供应商可以综合考虑，提供出完整的解决方案。

未来的世界将是一个万物智慧互联的世界，慧智微期待和你一起探索更适合未来智慧互联的射频技术。努力“化繁为简，使一切智慧互联”。

参考文献：

1. 华为企业级 Wi-Fi 7 全球最快，实测性能超 13Gbps，助力园区无线网络体验升级！华为数据通信公众号，2023/11/6
2. Pahlavan, K., Krishnamurthy, P. Evolution and Impact of Wi-Fi Technology and Applications: A Historical Perspective. Int J Wireless Inf Networks 28, 3–19 (2021).
3. Deepender, Manoj, U. Shrivastava and J. K. Verma, "A Study on 5G Technology and Its Applications in Telecommunications," 2021 International Conference on Computational Performance Evaluation (ComPE), Shillong, India, 2021, pp. 365-371
4. Charting an integrated future: IoT and 5G research papers, https://futurenetworks.ieee.org/
5. 5G 射频 PA 架构，慧智微
6. 5G 射频前端中的半导体，慧智微
7. 除了调匹配，射频人还需要掌握系统知识，慧智微

在线阅读

本文摘自慧智微微信公众号，扫描以下二维码，在线阅读

第五篇 射频系统设计

本篇内容

- EVM 详解
- 除了调匹配，射频人还要掌握系统知识
- 射频中的“Bypass”与“Decoupling”
- 协议再演进：从 5G-Advanced 到 6G
- 手机射频增强的方法
- 如何评价一部手机的通信能力？
- 今年，是时候了解“毫米波相控阵”了
- “iPhone 13 支持卫星通信”，支持的究竟是什么？
- 为什么有的手机能连卫星？
- 宽带卫星通信市场综述
- “有线”与“无线”通信，有什么异同？
- MIPI RFFE 协议解读
- 为了流畅上网，射频工程师们都做了什么
- 蜂窝移动终端的 Cat 指的是什么？
- 5G 物联网之 RedCap
- 车，怎么联网？
- 车联网，需要什么样的射频芯片

射频技术与其他技术领域的一个显著不同点在于，射频芯片的性能指标与系统整体指标之间存在着紧密的耦合关系。这种耦合性使得在射频芯片的设计和应用过程中，必须同时考虑芯片级别的性能以及它们在系统级别上的表现。然而，在芯片学习的过程中，系统指标往往容易被忽视，或者相关信息零散难以系统获取。实际上，深入了解系统指标对于射频芯片的设计和应用具有极大的帮助。

在这一章中，我们将把焦点放在射频系统的设计与应用上，特别是那些与芯片设计紧密相关的系统指标。我们的目标是通过结合实际应用场景，对常见的系统指标进行详细的探讨。通过本章，希望可以抛砖引玉，讨论射频系统指标与芯片设计之间的紧密关系，掌握常见的系统指标及其分析方法，并能够在实际应用中灵活运用这些知识，进行射频系统的设计和优化。

5.1 EVM 详解

作者：彭洋洋 博士

在现代数字通信系统的核心评估体系中，EVM（Error Vector Magnitude，误差向量幅度）是一项至关重要的性能指标。几乎在所有现行的通信标准中，都明确规定了 EVM 的衡量标准，并且这一指标还在随着通信协议的不断演进而持续发展。

举例来说，5G NR 蜂窝系统和 Wi-Fi7 系统对发射通路 EVM 的要求如下表所示 [1][2]。EVM 在两种系统中都占据着重要位置，尤其是在 Wi-Fi 通信系统中，EVM 甚至可能被视为 Wi-Fi 系统射频性能评估中的首要指标。

被广泛使用的 EVM 究竟代表什么物理含义？为什么对于现代通信系统如此重要？为什么 EVM 有% 和 dB 两种表示单位？EVM 指标又是如何计算及如何使用的？

我们带着这些问题，试着重新详细认识“EVM”这一既常见又略显神秘的重要性能指标。

Table 6.4.2.1-1: Requirements for Error Vector Magnitude

Parameter	Unit	Average EVM Level
Pi/2-BPSK	%	30
QPSK	%	17.5
16 QAM	%	12.5
64 QAM	%	8
256 QAM	%	3.5

公众号 · 慧智微电子

图 5.1: 5G NR 发射通路的 EVM 要求

Table 36-65—Allowed relative constellation error versus constellation size and coding rate

Modulation	Coding rate	Relative constellation error in an EHT MU PPDU (dB)	Relative constellation error in an EHT TB PPDU when transmit power is larger than the maximum power of EHT-MCS 7 (dB)	Relative constellation error in an EHT TB PPDU when transmit power is less than or equal to the maximum power of EHT-MCS 7 (dB)
BPSK	1/2	–5	–13	–27
QPSK	1/2	–10	–13	–27
QPSK	3/4	–13	–13	–27
16-QAM	1/2	–16	–16	–27
16-QAM	3/4	–19	–19	–27
64-QAM	2/3	–22	–22	–27
64-QAM	3/4	–25	–25	–27
64-QAM	5/6	–27	–27	–27
256-QAM	3/4	–30	–30	–30
256-QAM	5/6	–32	–32	–32
1024-QAM	3/4	–35	–35	–35
1024-QAM	5/6	–35	–35	–35
4096-QAM	3/4	–38	–38	–38
4096-QAM	5/6	–38	–38	–38

公众号 · 慧智微电子

图 5.2: Wi-Fi7 发射通路对星座误差（EVM）的要求

5.1.1 EVM 的来源

5.1.1.1 左手数字，右手模拟：重要的数字调制技术

要了解 EVM，还需要从数字通信技术说起。

在通信的早期时代，模拟通信系统占据主导地位。信号作为模拟量，在幅度、频率上进行转换及传输。然而，模拟信号抗干扰能力差、保密性差、传输距离受限，且难以支撑复杂的通信系统。于是，在 20 世纪 80 年代，随着数字通信技术和集成电路技术的迅猛发展，通信技术逐渐从模拟向数字转型。



图 5.3: 模拟通信技术向数字通信技术的转换

在这一转型过程中，数字调制技术成为了关键。

在数字通信中，尽管信号处理已经数字化，但数字“0”和“1”无法直接漂浮在空中，因为自然界中传输的仍然是具有特定物理量值的模拟信号，转换后的数字信号仍需以某种形式映射到模拟信号上。而实现数字信息和模拟信号传输桥梁和映射的，正是数字调制技术。



图 5.4: 数字调制技术完成的功能

常见的数字调制技术包括 ASK（Amplitude Shift Keying，幅度键控）、FSK（Frequency Shift Keying，频移键控）、QAM（Quadrature Amplitude Modulation，正交幅度调制）等。这些技术通过特定约定，建立起数字信号与模拟信号的对应关系，然后通过改变模拟信号的不同参数来传递数字信息。

以 16QAM 举例。16QAM 信号数字信息和模拟信号的对应可以由星座图直观反应出来。图5.5为 16QAM 调制下信号的星座图表示。

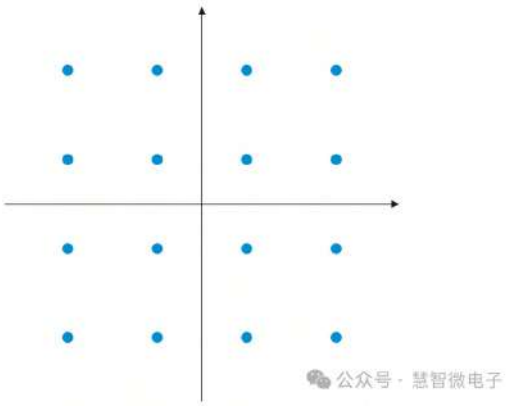


图 5.5: 16QAM 信号的星座图表示

在 16QAM 的星座图中，一共有 16 个点，所以可以代表 16 个不同的数字信息。每个点，就称为一个符号（Symbol）位。在 2 进制中，16 可以由 2 的 4 次方来表示。所以 16QAM 中的每一个点，可以代表 4 位不同的 2 进制信息。也就是常说的，16QAM 中的每个符号代表 4 个 bit 的信息。

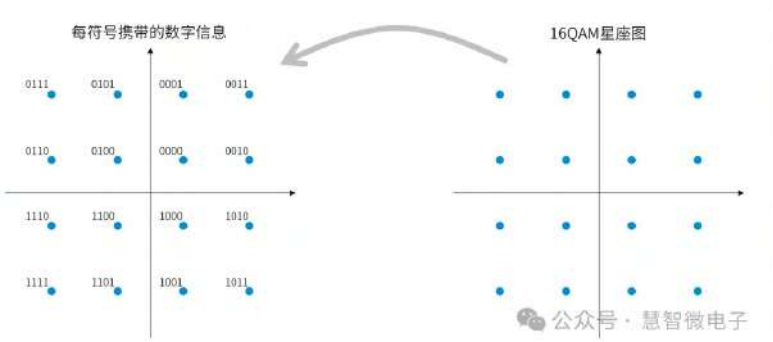


图 5.6: 左手数字：16QAM 中每个符号携带 4 个 bit 数字信息

在星座图上，也可以显示出来每个符号所对应的模拟信息。比如对于位于 A 位置的信号，其幅度就为 a，相位就为 α。对于位于 B 位置的信号，其幅度就为 b，相位就为 β。有了幅度和相位，就可以相应得出来唯一确定的模拟信号。

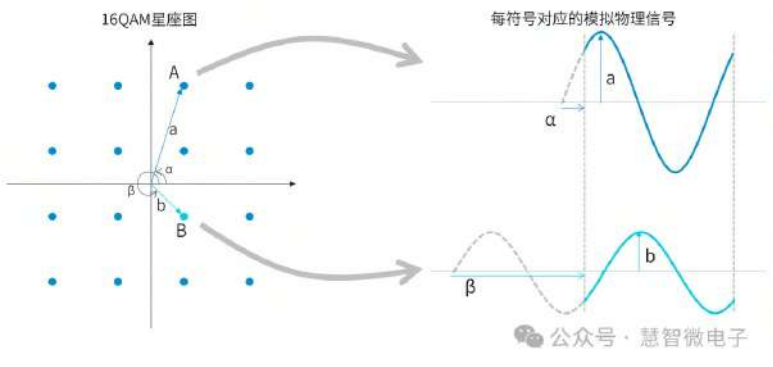


图 5.7: 右手模拟：16QAM 中的符号与模拟物理信号的对应

在数字信号的调制中，每一个符号都建立起数字和模拟的严格对应关系。左手数字，右手模拟，通过数字信号调制技术，数字世界和模拟世界就此连接。

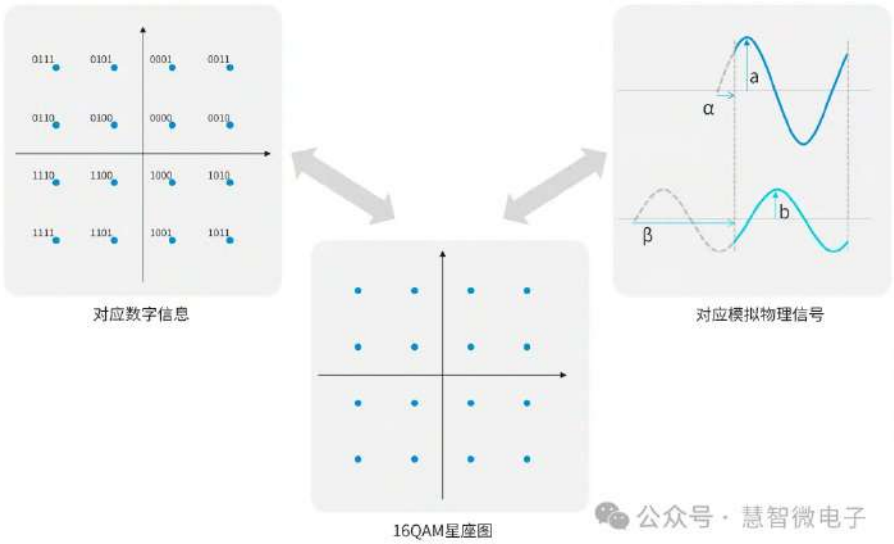


图 5.8: 数字信号调制技术将模拟世界和数字世界相连

5.1.1.2 EVM：数字信号调制效果的裁判

在数字信号调制中，每个符号的所在位置都非常重要。虽然数字通信有较强的抗干扰能力，符号位置发生轻微偏移并不影响通信。但如果偏移位置较多，就可能造成信号传输的误码，造成信号传输的错误。

如何来衡量数字信号调制的实际效果是好还是坏呢？这时就引入了误差向量（Error Vector）指标。误差向量反应实际信号和理想信号之间的差别。

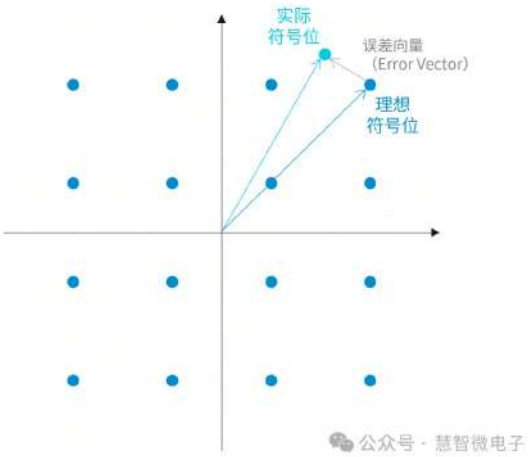


图 5.9: 误差向量的引入

5.1.1.3 EVM 的定义

在星座图中，任何一个符号都有其对应的理想位置，也即有确定的幅度及相位。但在实际信号中，总是会存在误差。

误差主要分为两种，分别是幅度误差（Magnitude Error），以及相位误差（Phase Error）。幅度误差反应信号幅值的大与小，相位误差反应信号的超前与滞后。

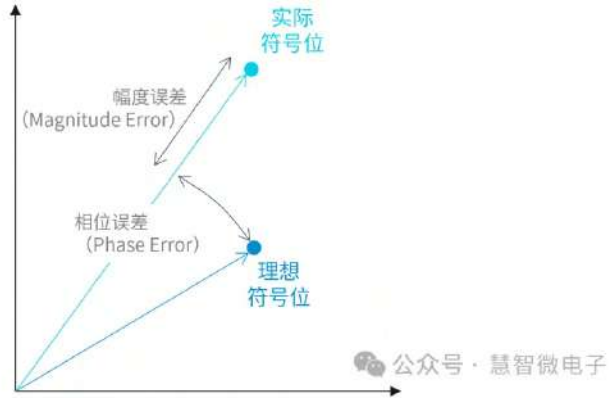


图 5.10: 实际信号与理想符号之间的误差

在实际应用，可以用一个连接理想信号和实际信号的向量，把幅度误差和相位误差统一表示出来。这个向量就叫“误差向量（Error Vector）”。

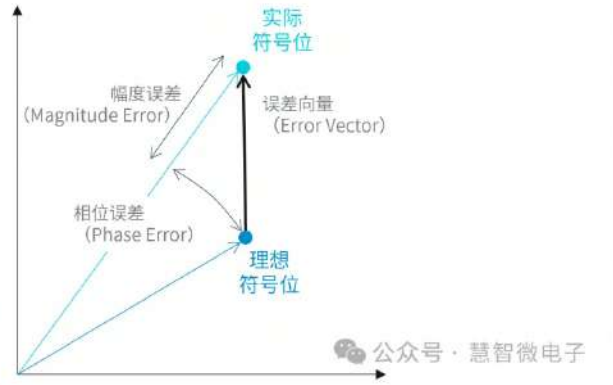


图 5.11: 误差向量（Error Vector）

在实际应用中，实际信号和理想信号之间的距离更能反应信号质量的好坏，而不是两个信号之间的角度差。所以会单独把误差向量的幅度（Magnitude）提取出来分析。这个幅度就是误差向量幅度，英文名称为 Error Vector Magnitude，简称 EVM。

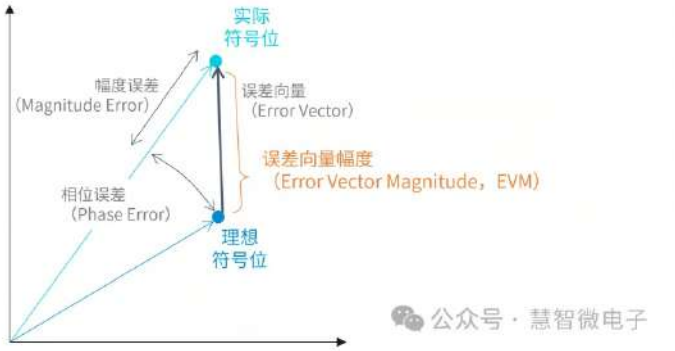


图 5.12: EVM 的定义

需要注意的是，EVM 是“误差向量幅度（Error Vector Magnitude）”，和前述的幅度误差（Magnitude Error）虽然名称中有两个英文单词一样，但却不是一回事。幅度误差只反应两者幅度之间的差别，而 EVM 反应的是误差向量的幅值。

5.1.2 EVM 的计算

在对 EVM 进行定性了解之后，就可以对 EVM 进行计算。在计算中，有几个细节需要注意。了解了这些细节，才能对 EVM 指标有清晰 de 地了解。这些细节包含：

- 选峰值还是选均值？EVM 中的参考向量选择
- “%”和“dB”：EVM 的两种单位
- 不同调制的 EVM 需求计算
- 器件级连的影响：链路中的 EVM

5.1.2.1 选峰值还是选均值？EVM 中的参考向量选择

虽然在前述定义中，EVM 是一个反应误差向量幅度的“绝对值”，但在实际使用中，误差向量幅度和有用信号幅度之间的“比值”比误差向量幅度的绝对值更为有意义。所以在使用应用中，EVM 经常会做归一化使用。

EVM 的归一化有两种方式，一种是对星座图中的最大信号做归一，另外一种是对信号的均方值（RMS）做归一。两种归一化时，所参考的信号如图5.13所求。

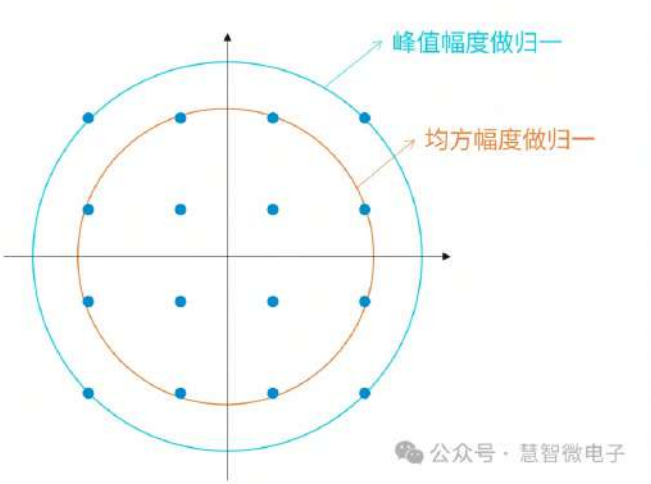


图 5.13: EVM 归一幅度的两种参考方式

对于 EVM 计算来说，以上两种选择方式均可，但需要在计算时明确做好约定。例如在 Wi-Fi7 协议中，就会明确约定归一化选择的功率为星座的“平均”功率，而非最大功率 [2]。

$$Error_{RMS} = \frac{1}{N_f}\sum_{i_f=1}^{N_f}\sqrt{\frac{\sum_{i_s=1}^{N_{SYM}}\sum_{i_{ss}=1}^{N_{SS}}\sum_{i_{sc}=1}^{N_{SD,total}}\left(I_e(i_f,i_s,i_{ss},i_{sc})-I_0(i_f,i_s,i_{ss},i_{sc})\right)^2+\left(Q_e(i_f,i_s,i_{ss},i_{sc})-Q_0(i_f,i_s,i_{ss},i_{sc})\right)^2}{N_{SYM}N_{SS}N_{SD,total}P_0}} \quad (36\text{-}102)$$

where

$I_0(i_f,i_s,i_{ss},i_{sc})Q_0(i_f,i_s,i_{ss},i_{sc})$ denotes the ideal symbol point in the complex plane in data subcarrier i_{sc} of the RU under test, spatial stream i_{ss}, and OFDM symbol i_s of frame i_f.

$I_e(i_f,i_s,i_{ss},i_{sc})Q_e(i_f,i_s,i_{ss},i_{sc})$ denotes the equalized observed symbol point in the complex plane of the data subcarrier i_{sc} of the RU under test, spatial stream i_{ss}, and OFDM symbol i_s of frame i_f.

P_0 is the average power of constellation.

N_f is the number of tested frames.

$N_{SD,total}$ is the total number of data tones of the occupied RU.

N_{SS} is the number of spatial streams of the data.

N_{SYM} is the number of data OFDM symbols.

公众号 · 慧智微电子

图 5.14: Wi-Fi7 协议中 EVM 计算的定义

5.1.2.2 “%”和“dB”：EVM 的两种单位

将 EVM 信号和实际信号进行比较，就得到 EVM 的值。归一化后的 EVM 值为一个百分比。如图5.15所示。

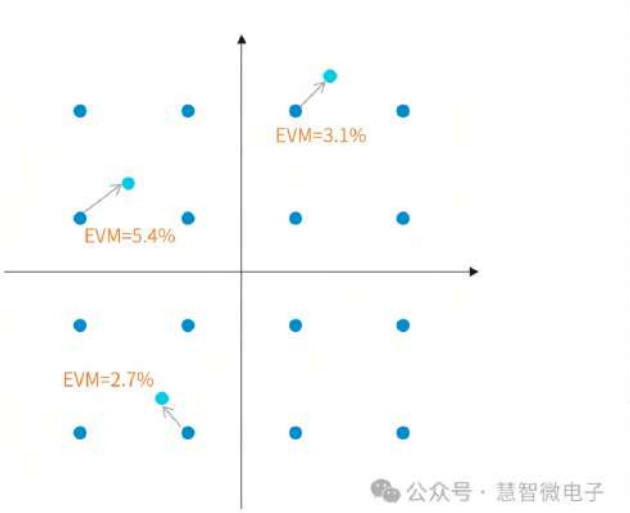


图 5.15: 归一化后的 EVM，用百分数表示

在射频的计算中，为了方便计算，方便地进行数字的观察处理，经常会用 dB 为单位来表示倍数或者比例关系。比如对于功率来说，说 3dB 代表两倍，10dB 代表 10 倍，30dB 代表 1000 倍等。这一方法同样可以在 EVM 的表征中使用。EVM 的 dB 和倍数之间的转换关系如下公式所示。

$$EVM_{dB} = 20 \times log_{10} EVM_{\%} \tag{5.1}$$

需要说明的是，EVM 的% 计算时，所使用的是信号“幅度”之间的比值，而非功率。所以在进行 dB 的转换时，需要用到的是 20*Log 来转换，而不是 10*Log。在转换时需要特别注意。

采用 dB 表示之后的 EVM 如图5.16所示。

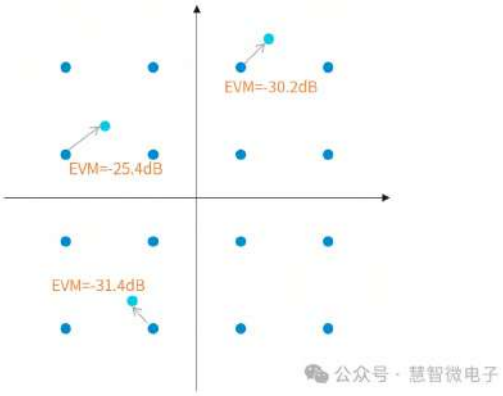


图 5.16: 采用 dB 表示的 EVM

在不同使用中，为了使用方便，EVM 的表征经常会在% 和 dB 两种表示之间切换。为方便使用和转换，整理二者对应关系如表5.17所示。

EVM (%)	EVM (dB)	EVM (%)	EVM (dB)	EVM (%)	EVM (dB)
0.1	-60.0	3.0	-30.5	11	-19.2
0.2	-54.0	3.3	-29.6	12	-18.4
0.4	-48.0	3.6	-28.9	13	-17.7
0.6	-44.4	4.0	-28.0	14	-17.1
0.8	-41.9	4.5	-26.9	15	-16.5
1.0	-40.0	5.0	-26.0	17	-15.4
1.2	-38.4	5.5	-25.2	20	-14.0
1.4	-37.1	6.0	-24.4	25	-12.0
1.6	-35.9	6.5	-23.7	30	-10.5
1.8	-34.9	7.0	-23.1	35	-9.1
2.0	-34.0	7.5	-22.5	40	-8.0
2.2	-33.2	8.0	-21.9	50	-6.0
2.4	-32.4	8.5	-21.4	60	-4.4
2.6	-31.7	9.0	-20.9	70	-3.1
2.8	-31.1	9.5	-20.4	80	-1.9
3.0	-30.5	10.0	-20.0	90	-0.9

公众号 · 慧智微电子

图 5.17: 采用% 和 dB 表示的 EVM 相互转换关系表

5.1.2.3 不同调制的 EVM 需求计算

了解了 EVM 的计算方法，我们就可以尝试对不同调制的 EVM 需求进行计算。以下以 64QAM 为例，讨论计算过程。

对于 64QAM 信号，若以符号与符号之间的间隔为定为 a，则如果一个符号偏离自己位置 0.5a 时，将会发生信号的误码。所以极限的 EVM 绝对值为 0.5a。同时，计算 64QAM 信号的幅度均方值，可得其均方值为 3.24a。此时计算 64QAM 的 EVM 为：0.5a/3.24a=15.4%，转换为 dB 表示为-16.2dB。

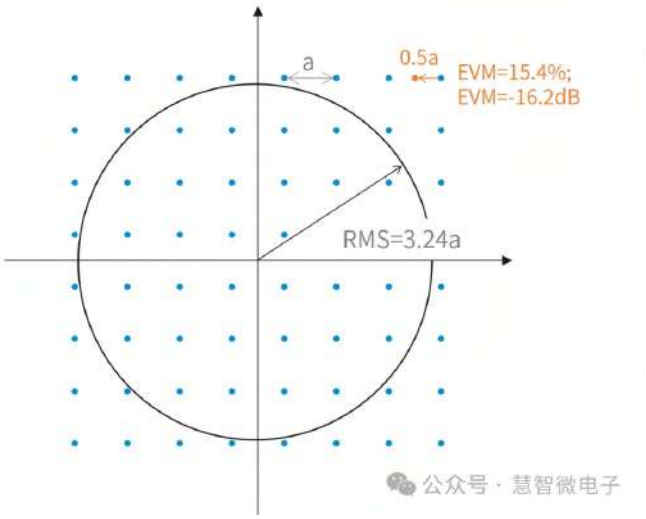


图 5.18: 64QAM 的 EVM 计算

同理，可以计算 256QAM、1024QAM 以及 4096QAM 下的 EVM 需求，计算结果如下表5.19所示。

调制	理论极限EVM要求	
	%	dB
16QAM	31.6	-10.0
64QAM	15.4	-16.2
256QAM	7.7	-22.3
1024QAM	3.8	-28.3
4096QAM	1.9	-34.3

公众号 · 慧智微电子

图 5.19: 不同调制下的 EVM 极限需求计算

通过上表可以看到，对于越高阶的信号，其对 EVM 的要求也就越高。在理解上也相对直观：高阶调制下各符号在星座图上的位置也更加靠近，就需要各符号更紧密的围绕在理想符号周围，来确保信号不出现误码。

5.1.2.4 器件级连的影响：链路中的 EVM

在射频发射及接收链路中，不同器件的非理想特性都会对 EVM 产生影响。对于多级链路，EVM 影响可以由以下公式计算 [7]：

$$EVM_{total} = \sqrt{\sum EVM_k^2} \tag{5.2}$$

其中 k 表示不同的 EVM 来源序号。

采用以上公式时需要注意，以上要求计算各链路 EVM 的影响之间互不相关。另外，各 EVM 单位均为%。如果 EVM 已经用 dB 表示，需要先将 dB 转为%，再进行计算。

在实际电路中，电路的非理想影响还可能表现在交调、相位误差、LO 泄露等多个方面，也有专门论文讨论和分析各种非理想特性与 EVM 的对应关系。有需要可以下载相应的论文来参考 [7]。

一些论文也对 EVM 的来源进行了分析和总结，比如在文章 [3] 中，在 Wi-Fi7 的射频收发机芯片设计中，发射链路的 EVM 中的 48% 是由 IMD 所贡献，20% 由相位噪声贡献，32 由带内噪声贡献；在接收链路中，60% 由热噪声贡献，30% 由相位噪声贡献，10% 来自于其他。

另外，在观察上一部分对 EVM 的理想计算，以及系统协议对发射 EVM 的要求，就会发现二者有明显的差别。比如对于 4096QAM 信号，理论计算 1.9%（-34.3dB）的 EVM 就可以完成传输，但 Wi-Fi7 协议中却要求达到 1.3%（-38dB）。

调制	理论极限EVM要求		Wi-Fi7的发射要求	
	%	dB	%	dB
16QAM	31.6	-10.0	11.2	-19.0
64QAM	15.4	-16.2	4.5	-27.0
256QAM	7.7	-22.3	2.5	-32.0
1024QAM	3.8	-28.3	1.8	-35.0
4096QAM	1.9	-34.3	1.3	-38.0

公众号 · 慧智微电子

图 5.20: EVM 的理论要求及 Wi-Fi7 协议中的发射要求

二者存在区别的原因是还要考虑接收通路的解调对 EVM 的恶化。并且保持一定的余量来保证信号的误码率不受影响。

5.1.3 影响系统 EVM 的因素

EVM 的定义和计算更多是从系统综合表现维度来看，是多重影响下的表现结果。在一个通信系统中，可能会对 EVM 产生影响的原因比较多，常见的有以下影响。

- 幅度影响
- 相位影响
- IQ 不平衡影响
- 其他影响

5.1.3.1 幅度影响

通过星座可以看到，当信号幅度发生变化时，将影响信号与理想信号之间的偏移，进而影响 EVM。

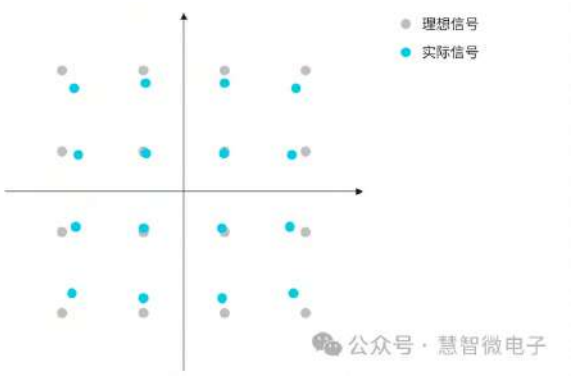


图 5.21: 幅度对 EVM 的影响

射频通信中，功率放大器（PA）作为发射链路的重要单元，若其幅度出现非理想特性，将会对系统 EVM 的影响。

PA 的幅度响应曲线一般用 AM-AM（Amplitude Modulation -Amplitude Modulation，幅度-幅度调制）来表示。AM-AM 曲线的横轴为输出功率，纵轴为幅度的放大量，也就是增益。对于放大器，其 AM-AM 曲线一般形状如图5.22所示。

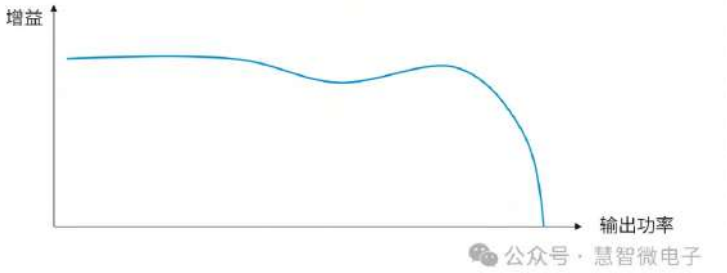


图 5.22: 放大器的 AM-AM 曲线

在以上曲线中，放大器有一定的增益，并且由于非理想效应，或者所选用放大器的特殊架构原因，其增益会在一定范围内发生变化。另外，在达到一定输出功率时，由于放大器会出现功率饱和，无法无限制的输出功率，所以对应增益曲线也会下折下来。

由于调制信号不同符号位的幅度不同，所以就有可能对应到放大器不同的输入信号幅度上。如果这两个符号位对应的放大器的增益差别过大，就造成经过放大器之后幅度相对关系发生变化。也就是 EVM 性能发生了恶化。

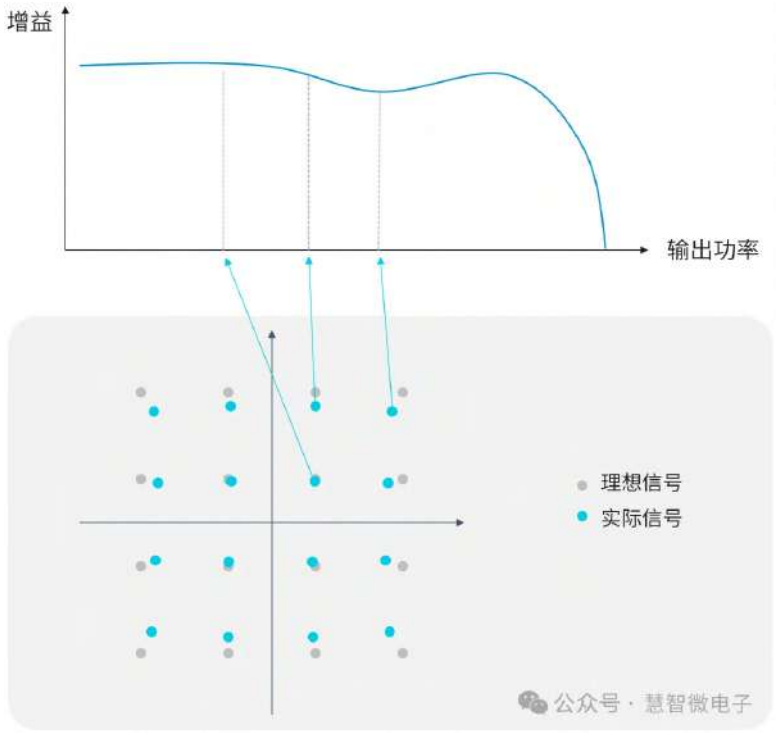


图 5.23: 放大器的 AM-AM 特性对 EVM 的恶化

正是因为这个原因，越先进的通信制式对 PA 的 AM-AM 特性要求越高，越需要足够平缓的 AM-AM 曲线。以 64QAM 到 4096QAM 的演进为例，对 PA 的 AM-AM 曲线要求可表示为如下关系。

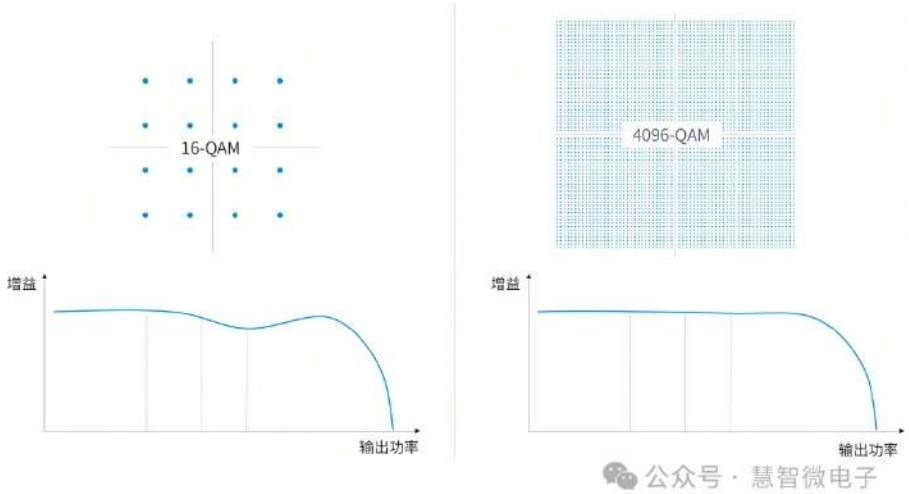


图 5.24: 不同调制对 PA AM-AM 曲线的要求

5.1.3.2 相位影响

在整个通信系统中，经过非理想器件时，产生变化的不止是信号的幅度，还有信号的相位。同样以功率放大器来说，其不同功率下相位的变化一般用 AM-PM（Amplitude Modulation-Phase Modulation，幅度-相位调制）曲线来表示。和 AM-AM 曲线相同，AM-PM 曲线变化的越剧烈，其对 EVM 的影响就会越明显。



图 5.25: AM-PM 曲线

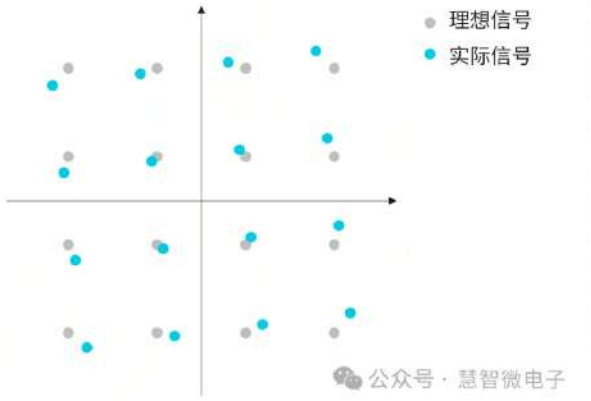


图 5.26: 相位对 EVM 造成的影响

5.1.3.3 IQ 不平衡影响

在通信系统的构建中，特定向量的产生一般通过 I 路和 Q 路两个信号的加合来形成。合成中，均是假定 I 和 Q 路完全对称。但在实际电路中，电路之间可能存在误差。若 I 和 Q 路出现不平衡，则将使合成的信号位置产生误差。例如，假定 I 路信号的幅度产生 90% 的偏移时，产生的 EVM 如图5.27所示。

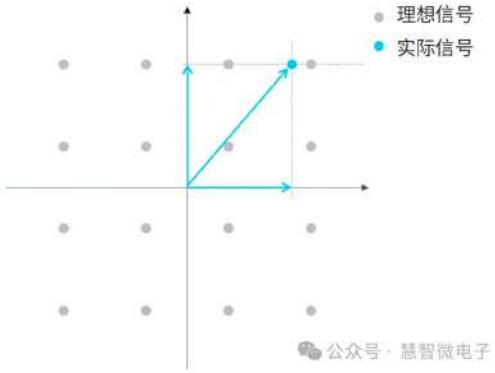


图 5.27: IQ 不平衡引起的 EVM

5.1.3.4 其他影响

除了上述幅度影响、相位影响以及 IQ 不平衡影响，还有其他多种因素会对 EVM 产生影响。比如：相位噪声、LO 泄露、信号交调、其他频率干扰等。影响 EVM 的因素可能是多方面的，在实际分析中，需要根据实际情况加以分析。

5.1.4 EVM 的应用

了解了 EVM 的物理意义、产生机理之后，就可以在应用中使用 EVM 来指导电路设计，或者是问题解决。

在应用中，EVM 还可以和其他参数结合，以得到更多的信息。常见的和 EVM 结合分析的参数主要有：

- EVM 和功率的结合：分析器件最大功率
- EVM 和时间的结合：分析器件时域表现
- EVM 和频率的结合：分析系统干扰频率

5.1.4.1 EVM 和功率的结合：分析器件最大功率

EVM 和功率结合的分析方法经常应用在发射通路器件的评估中。

发射通路的重要器件是功率放大器（PA），在 PA 分析中，经常会用到 EVM 和输出功率结合的方式，来观测 PA 的最大输出能力。

图5.28为慧智微某款 Wi-Fi7 FEM 的 EVM 与功率曲线。可以看到，在整体应用区间，慧智微 Wi-Fi7 FEM 有着良好的 EVM 特性。随着功率的升高，在达到标称功率之后，由于器件的饱和，EVM 曲线开始恶化上升。当 EVM 曲线达到指标临界时对应的功率点，就为此 FEM 能工作的最大功率。如下图慧智微 Wi-Fi7 FEM 产品，在维持-46 ~-47dB EVM 性能时，可以达到 17dBm 输出功率。

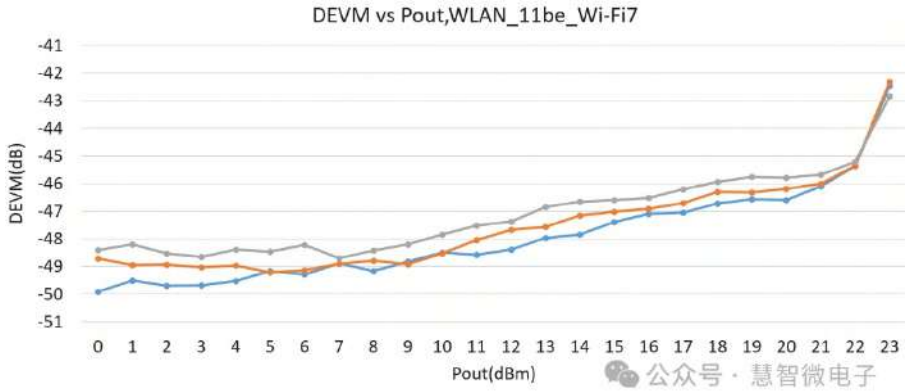


图 5.28: 慧智微 Wi-Fi7 FEM EVM 曲线

EVM 和功率结合的分析方法也可以用来分析射频收发机（Transceiver）的性能。图5.29为典型射频收发机输出表现能力图示，收发机 EVM 能力随功率变化呈“浴盆曲线”形状。在低工作功率水平下，EVM 性能主要由系统的噪声性能决定。在高工作功率水平下，EVM 主要由系统的非线性决定。系统的最低 EVM 水平（EVM Floor）通常根据所有误差源（包括相位噪声）的组合来决定 [4]。

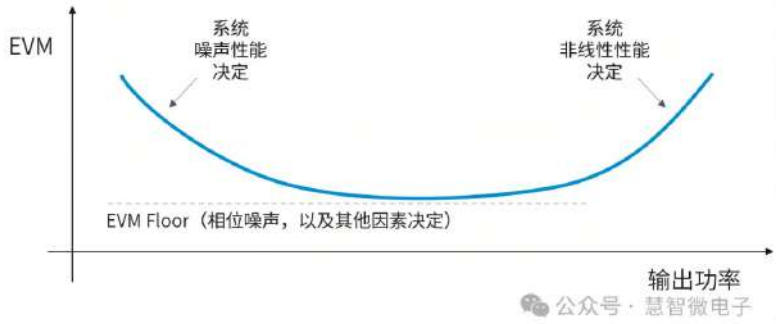


图 5.29: 典型射频收发机输出 EVM 能力的“浴盆曲线”

为了适配更大的功率范围，射频收发机会采用不用的功率模式进行设计，以使工作功率的 EVM 值达到最优状态。图5.30为某 ADI 收发机多档位下 EVM 与功率的对应曲线 [4]。

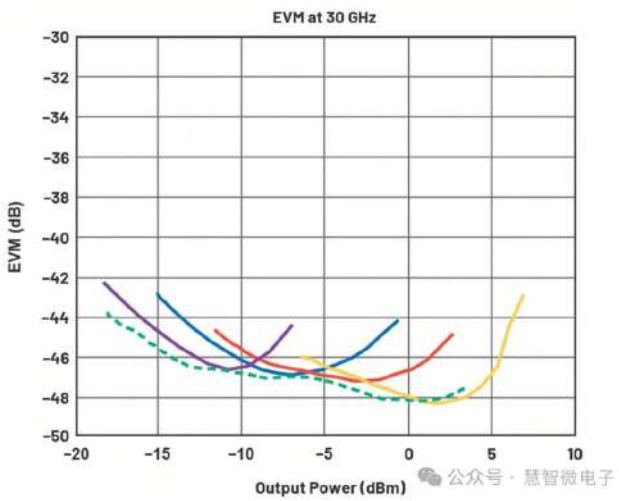


图 5.30: ADI 收发机 EVM 随功率变化特性

5.1.4.2 EVM 和时间的结合：分析器件时域表现

通过 EVM 与时间的结合，可以看出器件特性随时间的变化。

EVM 随时间的变化曲线在 Wi-Fi 系统中尤其重要。Wi-Fi 是一个 TDD 系统，在 Wi-Fi 信号传输过程中，发射和接收信号在快速进行切换。但由于 PA 的稳态建立需要一定的时间，在这个稳定时间到来之前，不希望系统传输有用信息。而 Wi-Fi 系统的帧结构复杂，在头部的 Preamble、header 帧中带有同步符号、样本信号、传输参数等信息，这就需要 PA 快速建立，用来响应帧头部的重要信息。

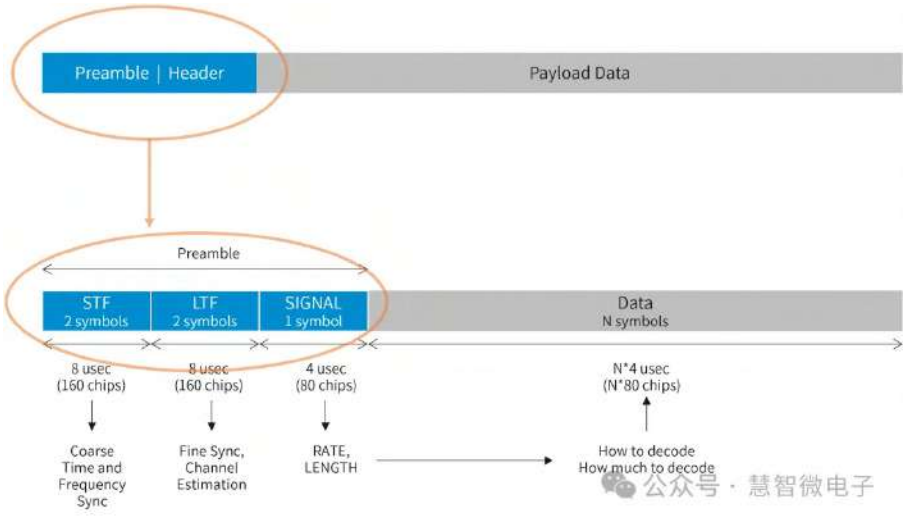


图 5.31: Wi-Fi 的帧结构信息

但对于 PA 电路，PA 的偏置电流随 TDD 脉冲不断打开和关闭，PA 电流的建立本身需要一段时间来响应。另外，电流稳态建立之前，PA 的温度也随着电流的变化而变化，造成 PA 本身的增益、相位特性均在变化。因为 Wi-Fi 系统的 EVM 参考信息在 Preamble 中进行传输，PA 在 Preamble 及后续数据传输中状态的不一致，造成系统 EVM 的恶化。所以在 Wi-Fi 系统中，除了查看静态 EVM 之外，还会重点查看 DEVM（Dynamic EVM，动态 EVM）。而器件随时间变化的特性，也是 DEVM 恶化的来源。

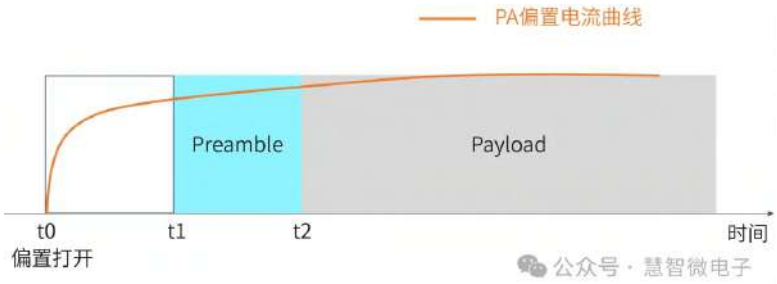


图 5.32: PA 特性随时间的变化，造成系统 DEVM 恶化

图5.33为文献 [5] 中，实际测试中 PA 输出功率随时间变化的曲线，及其对 EVM 的影响。可以看到 PA 输出功率随时间的变化，以及对 DEVM 产生的影响。

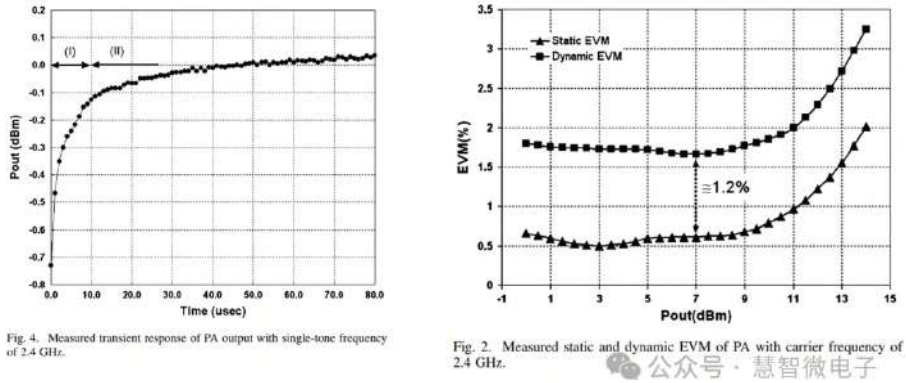


图 5.33: 实际测试中 PA 输出功率随时间变化的曲线，及其对 EVM 的影响

为了改善以上问题，一般会加入补偿电路。补偿电路的存在，可以提供一个瞬态电流补偿，将 PA 的偏置电流瞬间拉到所需要的静态值，然后再慢慢衰减下来，确保 PA 的总偏置电流一直维持固定值。从而保证 EVM 不随时间变化而变化。

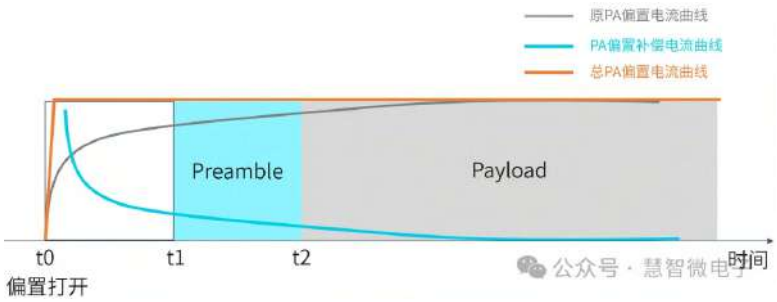


图 5.34: DEVM 的改善技术

5.1.4.3 EVM 与频率的结合：分析系统干扰频率

不止可以从时域角度来观察 EVM，将时域 EVM 信号进行傅立叶变换，转化到频域之后，同样可以看到众多信息。

比如当 EVM 恶化时，在频域对 EVM 进行分析，可以看到 EVM 恶化的频率点，从而有助于找到干扰来源，对系统进行改善 [6]。

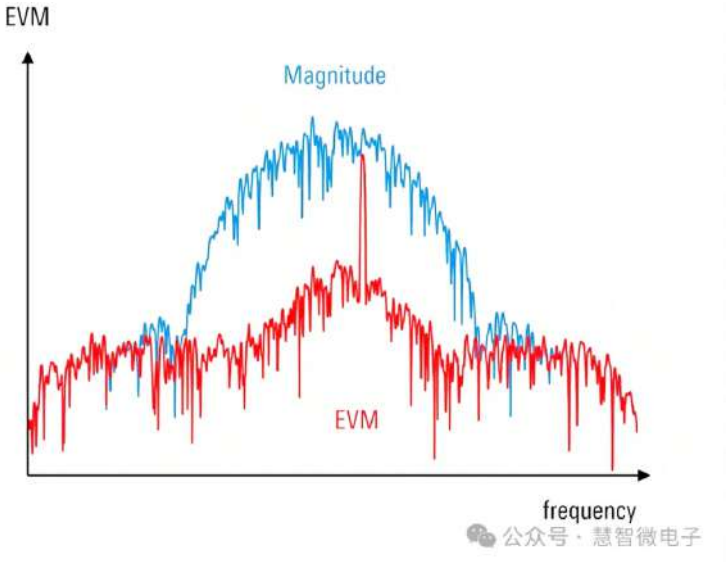


图 5.35: EVM 对频率做图观测干扰信号频率

5.1.4.4 总结

EVM 是现代数字通信中的重要指标之一，深入理解 EVM 指标对于电路设计、系统设计，以及问题解决有重要帮助。

本篇尝试从 EVM 的起源背景、精确定义、计算方法以及它对通信系统性能的具体影响等方面，分析 EVM 指标的理论基础、实际应用中的挑战与系统优化和故障诊断。

您还有哪些角度，可以更好的理解 EVM 这一指标，欢迎留言讨论，我们将对精彩留言赠送《无线重构世界》一本。

参考文献：

1. 5G NR 协议，3GPP TS 38.101-1 V18.6.0, 2024-06
2. Wi-Fi7 协议，IEEE P802.11be™/D3.0, 2023-1
3. J. Lee et al., ”A Tri-Band Dual-Concurrent Wi-Fi 802.11be Transceiver Achieving – 46 dB TX/RX EVM Floor at 7.1 GHz for a 4 K-QAM 320 MHz Signal,” in IEEE Journal of Solid-State Circuits.
4. ADI, ”How Error Vector Magnitude (EVM) Measurement Improves Your System-Level Performance,” , 2022
5. S. -W. Yoon, ”Static and Dynamic Error Vector Magnitude Behavior of 2.4-GHz Power Amplifier,” in IEEE Transactions on Microwave Theory and Techniques, vol. 55, no. 4, pp. 643-647, April 2007.
6. Understanding Error Vector Magnitude, ROHDE & SCHWARZ
7. Q. Gu, ”RF System Design of Transceivers for Wireless Communications”, New York: Springer-Verlag New York Inc., 2005.

在线阅读

本文摘自慧智微微信公众号，扫描以下二维码，在线阅读



5.2 除了调匹配，射频人还要掌握系统知识

作者：彭洋洋 博士

在《5G 射频前端模组的前世与今生》中，我们讨论了射频前端模组化这一趋势。在与射频工程师们讨论这一话题的时候，有些射频工程师问到：既然芯片都全集成了，那我们射频人是不是要失业了？

我觉得答案肯定是：不会。随着无线通信的发展，射频会越来越复杂，需要射频人掌握的知识范围会越来越广，射频人会越来越紧俏。

只不过，系统越来越复杂、集成度越来越高之后，对射频人的能力也提出了更高的要求。射频人除了对射频匹配这项基本功熟悉之外，**要把复杂的射频系统玩的转，还必须具备足够强的系统知识。**

为此，我们借 5G 手机中的射频系统构成为例，抛砖引玉，与大家讨论射频人日常工作中可能遇到的系统类知识。

5.2.1 手机射频系统演进

随着 1973 年第一部手机被发明，开启了人类的移动通信时代。在过去 50 年时间里，移动通信以大约每 10 年为一个周期快速演进，从 1G 到 5G，手机成为了连接万物的智能中枢。

	1G	2G	3G	4G	5G
手机					
时间	1980s	1990s	2000s	2010s	2020s
主要功能	模拟语音	数字语音	无线数据	宽带数据	万物互联
通信能力	2.4 Kb/s	64 Kb/s	2 Mb/s	100 Mb/s	1 Gb/s
频段数目	1	<5	<10	<70	<100*
射频方案	晶体管 分立方案	分立 芯片方案	分立 芯片方案	多频多模 分立芯片方案	集成模组 芯片方案

*定义中

图 5.37: 移动通信的演进之路

移动终端演进的革命性突破就是无线通信能力。如图5.38，从 3G 至 5G 的演进过程中：

- 主要通信频段范围不断展宽，由最高频段的 2GHz，已扩展至 5.0GHz，在部分毫米波终端中，已经扩展至 40GHz；
- 信号带宽不断增加，单载波信号带宽由 3G 时代的 3.84MHz，扩展至 100MHz，信号带宽增加近 30 倍；
- 载波聚合、多天线方式进一步提升通信能力；

无线通信能力的提升要借助射频系统的演进来实现。伴随着通信能力的提升，射频前端方案也变的越来越复杂。图5.39为 3G、4G 及 5G 系统方案的简要框图 [1][2]。需要说明的是，为了呈现简洁，以下方案已经做了非常大程度的简化。但依然可以看到，当前 5G 方案非常复杂。

5.2.2 5G 射频前端方案架构

上一节提到的 5G 射频前端方案已经比较复杂，但在实际应用中，伴随着不同项目间细节需求的差异，实现方式的不同，方案只会更加复杂，常常让人无从下手。图5.40为较为典型的 5G 系统方案图示。

此方案只是较为简单的标准系统框图，并且做了大量简化。可以看到，即使这样，大量的电路模块还是让系统理解起来产生了困难。为了理解以上系统方案，我们将系统分为如下几个层次讨论：

- 层 1：天线层
- 层 2：天线合路层

		3G	4G	5G
主要通信频段	低频	824~915MHz	617~960MHz	617~960MHz
	中频	1.71~2.02GHz	1.71~2.02GHz	1.71~2.02GHz
	高频	-	2.3~2.7GHz	2.3~2.7GHz
	超高频	-	-	3.3~5.0GHz
	毫米波	-	-	26.5~40GHz
最大信号带宽		3.84MHz	20MHz	100MHz
载波聚合		不支持	支持	支持
多天线		不支持	支持	支持
最大通信速率		~2 Mb/s	~100 Mb/s	~1 Gb/s

图 5.38: 从 3G 到 5G 的通信能力演进

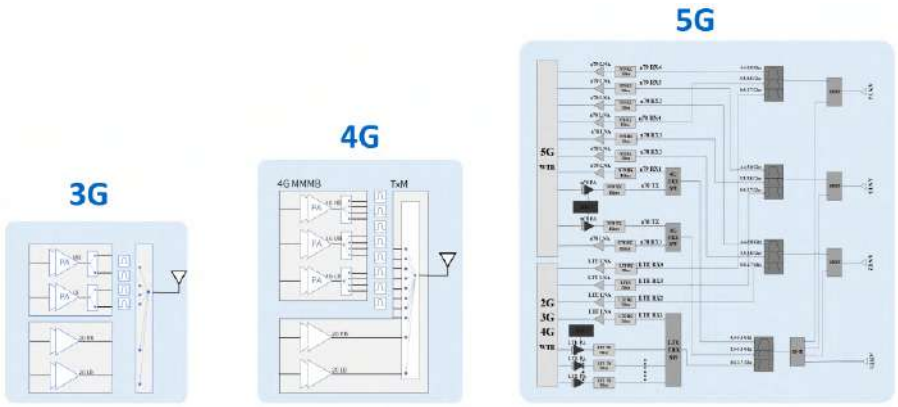


图 5.39: 射频方案的演进

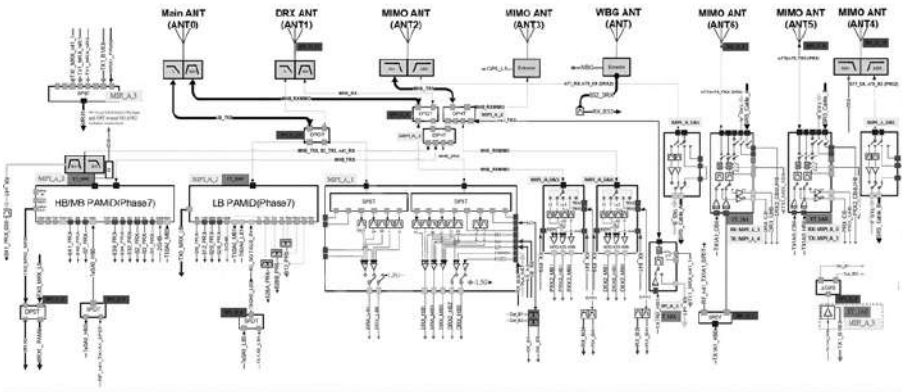


图 5.40: 典型 5G 射频系统方案

	LB	MHB	UHB
全称	Low Band	Mid/High Band	Ultra High Band
频率范围	617MHz~960MHz	1.7GHz~2.7GHz	3.3GHz~5GHz
MIMO需求	2x2	2x2或4x4	4x4

图 5.41: 天线的分组设计

	功能需求	天线设计
LB	2x2	ANT0: LB PRX ANT1: LB DRX
MHB	4x4	ANT2: MHB PRX ANT3: MHB DRX ANT4: PRX_MIMO ANT5: DRX_MIMO
UHB	4x4	ANT6: n77/79 PRX ANT7: n77/79 DRX ANT8: n77/79 PRX_MIMO ANT9: n77/79 DRX_MIMO
n41	4x4 (由于EN-DC，需要独立实现)	ANT10: n41 PRX ANT11: n41 DRX ANT12: n41 PRX_MIMO ANT13: n41 DRX_MIMO

图 5.42: 完整的天线设计需求

- 层 3：天线切换层
- 层 4：频段开关层
- 层 5：子路径实现层

层 1：天线层

之所以把天线作为第一层来理解，原因有两个，分别是：

- 在手机终端中，天线设计往往是受限的，与手机外型设计相关性强
- 手机通信能力的定义，如频段、MIMO 数目、CA、EN-DC、多卡共存及通信系统共存等，与天线设计直接相关

5G 手机中，一般根据天线的频率特性和系统需求，将天线以 LB、MHB 及 UHB 进行分组。各天线的需求和特性如图5.41。以 NSA（n41/77/79 与 LMH 频段做 EN-DC）、LTE 4x4 MIMO、NR 4x4 MIMO 为约束下，可设计天线如图5.42。

可以看到，以上设计中合计需要 14 根天线，这个天线数目在手机中是很难实现的。这还不包含 Wi-Fi，GPS 等天线。于是，就需要用到“天线合路器”，将可以合路的天线合路起来。

层 2：天线合路层

如果每个天线分拆设计，可以得到最为简洁、互不干扰的天线方案。但手机终端中寸土寸金，无法放置这么多天线。并且天线数目增多之后，每个天线得到的净空变小，天线性能也会变差。于是就需要将可以合并的天线进行合并。

天线的合并主要有三个原则：

- 合并的频率在天线可覆盖的范围内
- 合并带来的隔离度满足系统需求（EN-DC、CA 等）
- 整体的合路损耗满足系统要求

	LB	MHB	UHB	n41
ANT0	PRX	PRX	-	-
ANT1	DRX	DRX	-	-
ANT2	-	PRX_MIMO	DRX	-
ANT3	-	DRX_MIMO	-	DRX
ANT4	-	-	PRX	PRX
ANT5	-	-	PRX_MIMO	PRX_MIMO
ANT6	-	-	DRX_MIMO	DRX_MIMO

图 5.43: 经过合并之后的天线分组

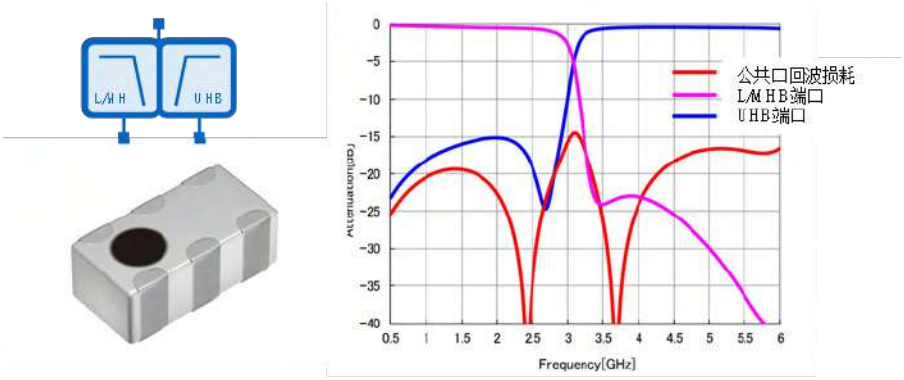


图 5.44: Diplexer 的图标、LTCC Diplexer 产品照片，以及频率响应

如《5G 移动终端天线设计》所述，典型的天线合并设计如图5.43所示 [3]。

天线合并主要通过 Diplexer 来实现。Diplexer 中文名一般翻译为天线共用器或者双讯器，是一种将不同频率信号合成后接入天线的装置。在实现上，一般采用 LTCC 工艺通过设计低通与高通两个滤波器来实现。图5.44为典型的用于 L/MHB 与 UHB 合路的 Diplexer 框图、LTCC Diplexer 照片，以及频率响应曲线。

在实际应用中，除了 Diplexer 之外，还有 Triplexer（三合一），和 Diplexer 的功能类似，可以合成三路。

Diplexer 和 Triplexer 在相近频带之间的隔离度和带外抑制较差，所以一般只用于频率间隔较大的频段，比如这四种最常用的频率范围：617 ~960MHz，1710 ~2200MHz，2300 ~2690MHz，3300 ~5000MHz，基本都可以任意组合进行合路。

但是有一点要注意：2200MHz（B66 RX）和 2300MHz（2300MHz 是 B40 TRX 频点，2305MHz 是 B30 TX 频点，一般北美市场中高端机型有 B66+B30 CA 需求）由于频率间隔只有 100MHz，普通的 LTCC Diplexer 或者 Triplexer 带外抑制性能不是太优，一般都用 LTCC+SAW（或者 BAW）混合工艺实现。

还有一点，很多频带间隔更小，比如 Wi-Fi 5G 和 n79 合路，或者 Wi-Fi 2.4G 和 B40（或者 B41）合路，这样根本无法用一般 LTCC 工艺实现，所以另一种形态的器件也诞生了，Extractor （抽取器）。Extractor 的功能和 Diplexer 或者 Triplexer 一样的，但是每一路都是用 SAW 或者 BAW 工艺实现的，这样以保证每一路都有更好的隔离度和带外抑制性能，当然，这样的工艺也会导致 Extractor 的成本更高，所以一般都是用于中高端设备。

天线合路完成之后，就可以结合手机外观进行天线的排布放置。图5.45为典型的天线放置图例。

层 3：天线切换层

经过层 2 的天线合并之后，天线基本上可以固化下来了。但固化下来的天线并不是一成不变的，一些功能需要多天线之间有相互切换的特性。这些功能主要有：

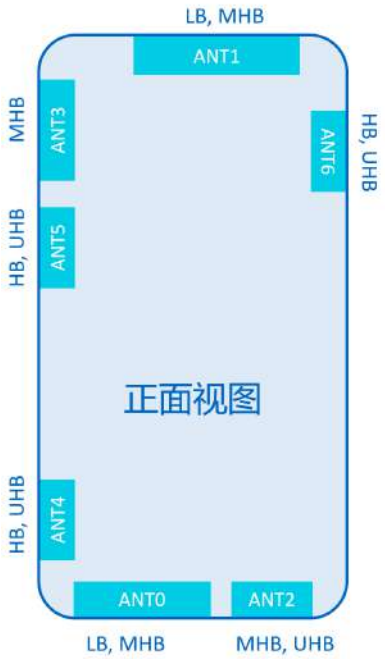


图 5.45: 典型的手机天线放置图例



图 5.46: 人手持握状态下对手机下天线的遮挡

- 智能天线切换
- SRS 天线轮发
- 天线的临时占用

智能天线是指发射天线可以进行主/副天线的切换，当一根天线信号不好时，手机信号就可以切换到另外一根天线进行发射。手机因为握持、放置的原因，天线很容易受到影响，如图5.46，智能天线功能可以大大保证发射信号的可靠性。在一些平台中，智能天线切换的功能被称为 TAS（Transmit Antenna Selection，发射天线选择）或者 ASDiV（Antenna Switch to Diversity）。

SRS 是 Sounding Reference Signal 的缩写，中文名为“探测参考信号”。SRS 是利用 TDD 系统中发射与接收同频互易性，靠发射上行探测信号，来探测中最优的接收路径与接收信道，完成接收性能的优化。与 TAS 功能一样，SRS 也需要用到发射信号在天线间的切换，不同的是，TAS 功能生效后，天线会较长时间驻留在当前天线上，而 SRS 是根据协议需求所做的轮发侦听动作，所以在平台应用中，SRS 的优先级一般高于 TAS。

除了 TAS、SRS 等天线切换外，还有部分功能会临时占用天线。比如，双卡双待方案中两张 SIM 卡对部分天线的切换使用；或者 SUL（Supplementary Uplink，辅助上行）中辅助上行频段借助部分天线完成上行性能的

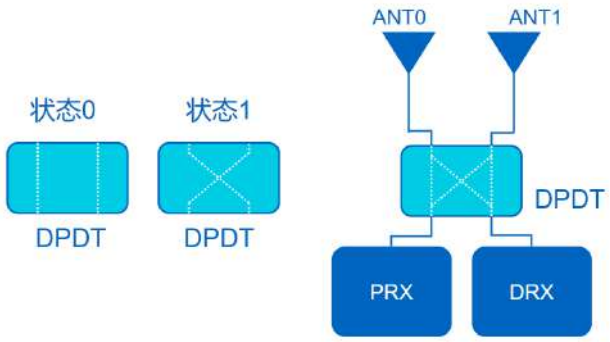


图 5.47: DPDT 的两种状态，及在系统中的位置示意

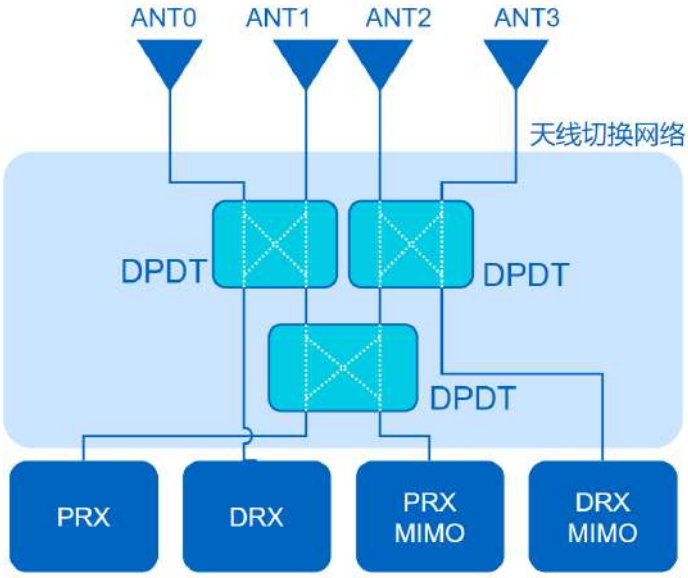


图 5.48: 利用 DPDT 实现 4 天线 SRS 轮发功能

补充。这些都需要用到天线切换的功能。

以上这些功能都需要信号在不同天线间切换。实现天线所连接信号切换的一般是多刀多掷开关，最常见的是 DPDT（双刀双掷）开关。DPDT 常见的符号表示及其在系统中的位置如图5.47所示。

利用多个 DPDT，或者 NPNT 开关，就可以完成天线切换网络，使射频信号在多个天线间切换。下图为天线切换网络示意图。图5.48可完成射频信号在四天线之间的 SRS 轮发功能。

层 4：频段开关层

进入 5G 之后，手机支持的频段数目急剧增加，根据 Mobile Expert 统计，进入 2019 年之后，高端手机支持频段数目将达到 30 个以上 [4]，如图5.49。

如此多的射频频段无法直接连接进入天线，也无法进入天线切换网络进行切换。这个时候就需要将这些射频通路先进行一次合路，再接入到天线选择开关、天线合路器、或者直连天线中。

通路合路的功能一般通过 ASM（Antenna Switch Module，天线开关模组）实现。在功能上，ASM 一般为单刀多掷开关，具体的掷数和分组根据系统方案确定。在连接上，单刀多掷开关的多端口侧连接各频段的发射及接收通路，另一端口端连接后续的天线侧。如图5.50。

层 5：子路径实现层

到了这一层，5G 手机射频系统才拆解到了我们教科书上所学到的射频前端子模块：PA、LNA、滤波器及开关。子路径实现层利用这些子模块，将每个频段的射频通路搭建起来。



图 5.49: 手机支持上行频段数目（Mobile Expert, 2022）

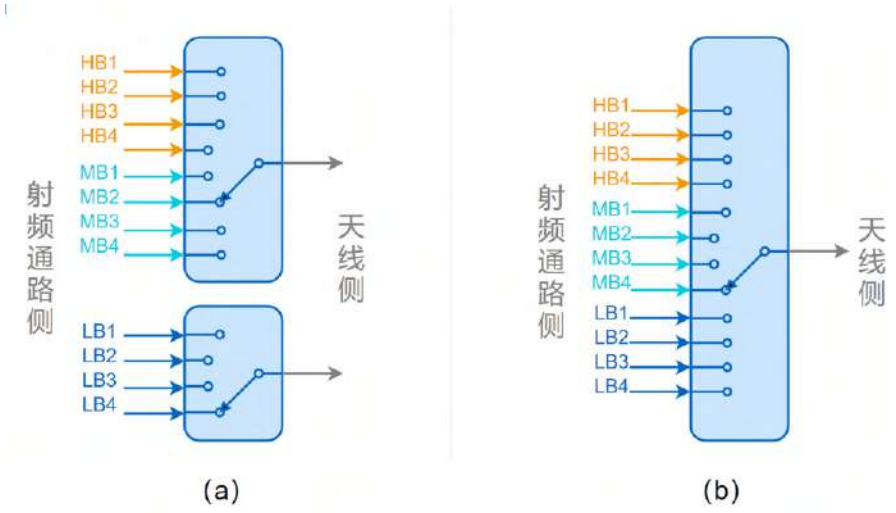


图 5.50: 频段选择开关及不同的合路方案

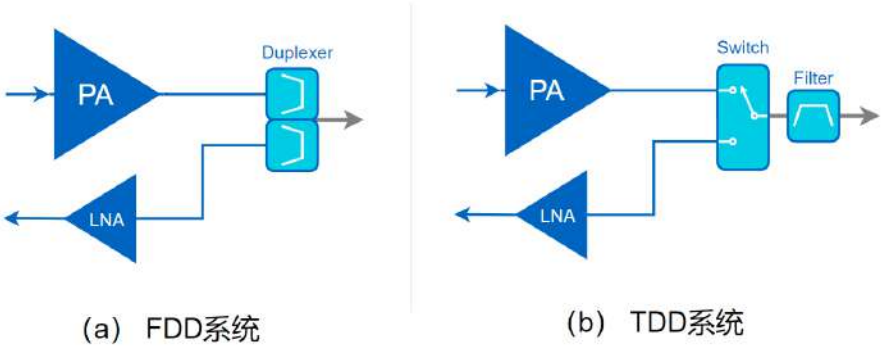


图 5.51: FDD 与 TDD 系统的构成

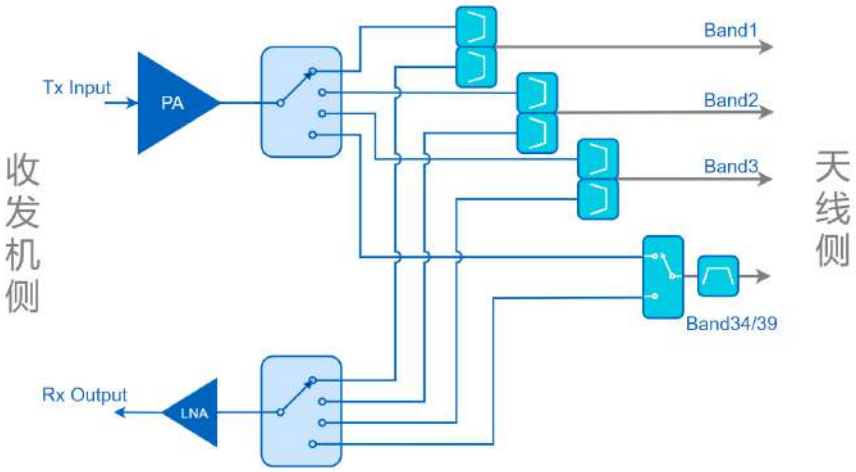


图 5.52: 中频段射频通路实现方式示意

在通路实现上，根据双工方式的不同，可以分为 FDD（Frequency Division Duplexing，频分双工系统）结构和 TDD（Time Division Duplexing，时分双工系统）结构。在 5G 协议中，以上两种双工方式均有定义，并且在不同频段上做了区分。FDD 与 TDD 系统在实现上的结构如图5.51所示。

另外需要说明的是，由于滤波器带宽较窄带，一般一个频段就需要用到一个滤波器或双工器。而 PA 和 LNA 作为有源电路，可以覆盖较宽的频率范围。所以在实际应用中，会将相近频段的 PA 用一个宽带 PA 来实现，再用开关分开，接到不同的滤波器或者双工器上，达到节省 PA 和 LNA 通路的目的。以中频段 Mid Band 中的 Band1/2/3/34/39 频段为例，具体射频通路实现如图5.52所示。

在子路径实现上，主要有“分立方案”和“模组方案”两种形式。

分立方案是指 PA、滤波器、LNA 等各采用不同的芯片进行实现，再在手机板上组合成完整方案。

模组方案指的是芯片厂商利用集成化设计整合能力，将 PA、LNA、滤波器及开关等子模块采用 SiP（System in Package，系统化封装）的形式整合在一起，达到节省布板面积，优化射频性能的目的 [5]。如图5.53。

其他考虑

基于以上解构，可以将 5G 射频系统分解如图5.54模块来理解。

需要说明的是，以上只是较为标准的系统考虑方式，在实际项目中，还会遇到多频段 CA、EN-DC 的考量，Wi-Fi、GPS、蓝牙及 UWB 等共天线的需求，以及要综合考虑的系统干扰等问题，实际系统会比以上示例系统复杂很多。

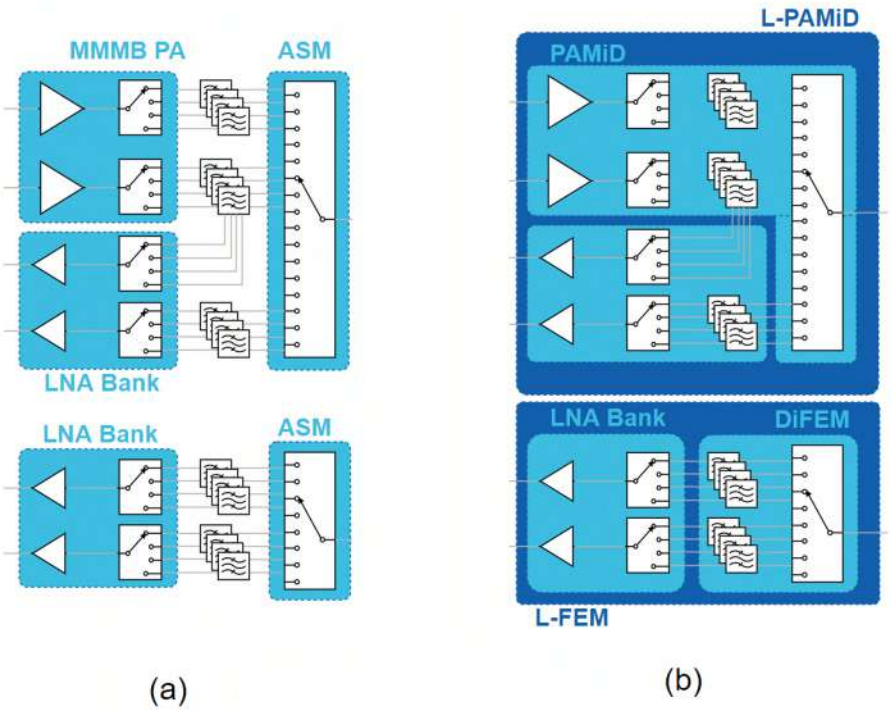


图 5.53: 分立方案（a）与模组方案（b）实现的射频前端系统

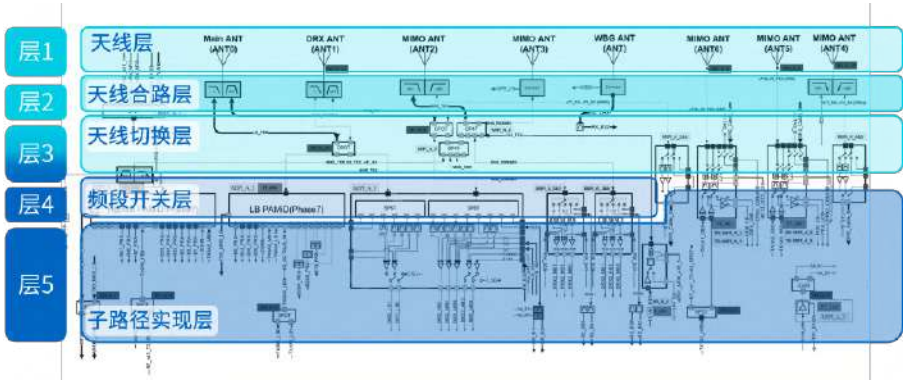


图 5.54: 典型 5G 射频系统方案

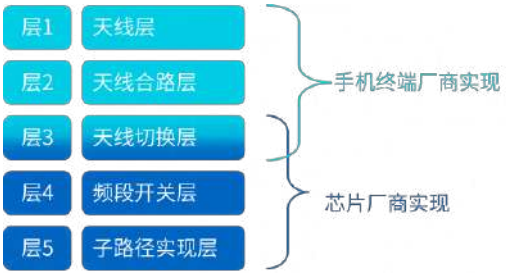


图 5.55: 射频前端方案的实现

5.2.3 5G 射频前端方案的实现

通过以上结构，可以看到不同层的模块有着不同的功能与作用。在具体实现上，一般天线层、天线合路层及天线切换层由手机终端厂商来实现；频段开关层、子路径实现层由芯片厂商实现。不过近年来由于射频前端方案集成度越来越高，一些高集成模组中开始集成天线切换开关，方便系统应用时做开关切换需求使用。如图5.55所示。

过去十年，射频前端不断演进，不管是和天线密切相关的层 1 至层 3，还是与具体通路实现相关的层 4 与层 5，都发生了巨大的变化。

在芯片实现侧，不仅有 PA、滤波器等各个子电路的性能演进，还出现了各个子模块多种组合带来的方案变化。为了帮助大家了解射频前端发生的变化，把握最新的演进趋势，慧智微将层 4 至层 5 中的芯片方案、具体实现做了梳理，详情可参考以下材料：

- 有关射频前端方案演进：《从 4G 到 5G，射频前端方案 10 年之路》
- 有关模组化方案的趋势：《5G 射频前端模组的前世今生》
- 有关 PA 技术的演进：《5G 射频 PA 的架构》《高效率 PA 设计：从 Class A 到 Class J》《5G 射频 PA 的 Load-line 与 Load-pull》
- 有关滤波器技术：《5G 射频前端模组中的滤波器》

欢迎系统平台厂商、手机终端厂商射频工程师一起，分享层 1 ~ 层 3 的技术及演进。

5.2.4 总结

除了射频前端芯片侧一直有各种各样技术频繁更新，随着 5G 应用的增多，设计一部 5G 手机还需要我们射频工程师从需求收集、方案评估、设计实现、调试及 debug、到量产等每个环节都要深度参与，每个环节都在考验着我们射频人对射频系统性知识的掌握是否扎实，对射频器件及指标理解是否透彻和射频技术经验是否丰富。

- 3G 手机开发：总共频段数是个位数，口算即可完成链路预算，主要工作是 PCB 设计及调试匹配；
- 4G 手机开发：单个项目一般十几个频段，较复杂应用场景是 DLCA 和 MIMO，链路预算简单，主要工作是 PCB 设计及调试匹配和优化 MIPI 参数；
- 5G 手机开发：单个项目二十个频段以上，需支持 EN-DC、CA 和 UL MIMO 等场景，链路预算复杂且工作场景繁杂。主要工作是：需求理解、RFFE 器件理解、系统框图梳理，链路预算、PCB 设计（要考虑很多因素）、单频段性能调试，共存频段性能测试及调试，debug 时间长难度大等，整个工作难度及工作量都呈指数级增加。

正如 5G 的愿景所期待，万物互联的智能世界正在到来。而万物互联的智能世界一定是基于无线连接的，移动通信技术由 2G 演进到 5G，再到未来的毫米波通信、卫星通信，射频技术不断在发生变革，无线射频技术是移动终端的核心关键技术。

在这个演进过程中，射频工程师的重要性毫无疑问会越来越突显。期待和大家一起，解构射频演进中的系统知识。

本文撰写过程中，得到慧智微应用团队大量专业的讨论与分析支持，在此表示感谢！

参考文献

1. 从 4G 到 5G，射频前端方案演进十年之路, 慧智微，2021
2. RF Implementation of Multi-mode Multi-band 5G UE, vivo, 2018
3. 5G 移动终端天线设计，林辉等，中国工信出版社，2021
4. RF Front Ends for Mobile Devices, Mobile Expert, 2022
5. 5G 射频前端模组的前世今生，慧智微，2021

在线阅读

本文摘自慧智微微信公众号，扫描以下二维码，在线阅读



5.3 射频中的“Bypass”与“Decoupling”

作者：Gareth

射频电路工作频率高，与数字、模拟电路相比，射频电路非理想特性众多，寄生效应复杂。实际应用中，需要对射频信号串扰、不同器件间隔离等做良好处理，稍有不慎，便会出现干扰、震荡等问题。在射频电路的测试中，实测结果与仿真结果不一致是家常便饭，实测被仿真完美预测的情况寥寥无几。

“旁路（Bypass）”与“去耦（Decoupling）”处理不好是造成射频实测异常的重要原因之一。虽然对于这两个问题在设计时有一些经验可循，但如果对问题成因理解不深，很有可能造成问题无法本质解决，严重时还会引起其他更严重的后果。

而且“Bypass”和“Decoupling”很多时候结构类似，很多人也将“Bypass 电容”与“Decoupling 电容”用做同义词，不过二者的使用目的和设计考虑是不同的。

本文尝试对“Bypass”和“Decoupling”的概念做一个讨论，理解二者设计的基本理念，同时给出二者设计时注意的要点。

5.3.1 Bypass”与”Decoupling” 的定义

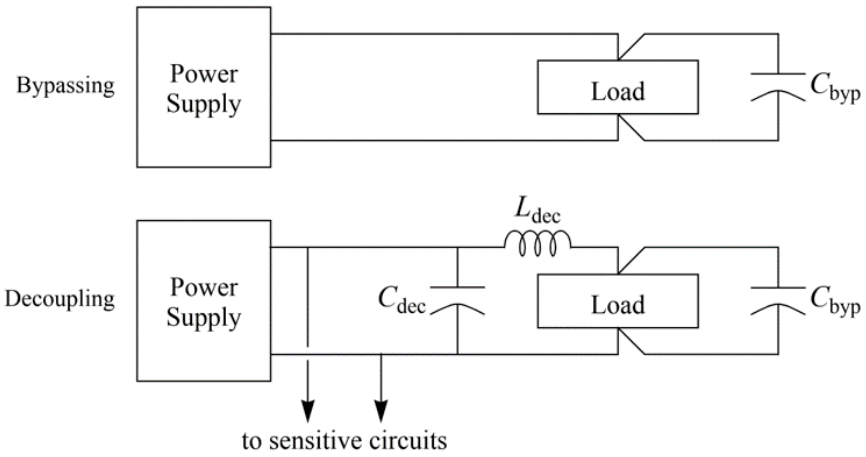


图 5.57: Bypass 与 Decoupling

- Bypass（旁路）
  是指为射频信号提供一个低阻通路（通常使用电容，如上图中的 C_{byp}，使射频信号沿此低阻通路流动，减少射频信号向其他高阻路径的流动。
- Decoupling
  目的是将共用连接线上的不同电路做隔离，一般由低通网络构成。隔离的目标是不让电路间的噪声互相传播干扰。

在图5.57中 Decoupling 电路设计中，Bypass 电容 C_{byp} 和 Decoupling 电路 L_{dec}、C_{dec} 共同完成 Decoupling，这种设计方法在射频电路实现中经常使用。以下将对 Bypass 和 Decoupling 电路做详细讨论。

5.3.2 Bypass 电路

理想情况下，电压源提供低阻阻抗，但实际应用并非如此。电压源的阻抗随着频率的升高而升高；另外由于电源走线的寄生电感效应，以及电源上连接其他器件的寄生效应，射频芯片端口的射频阻抗已经不再保证是低阻，这个时候，就需要用 Bypass 电容来提供低阻路径。如图5.58所示。

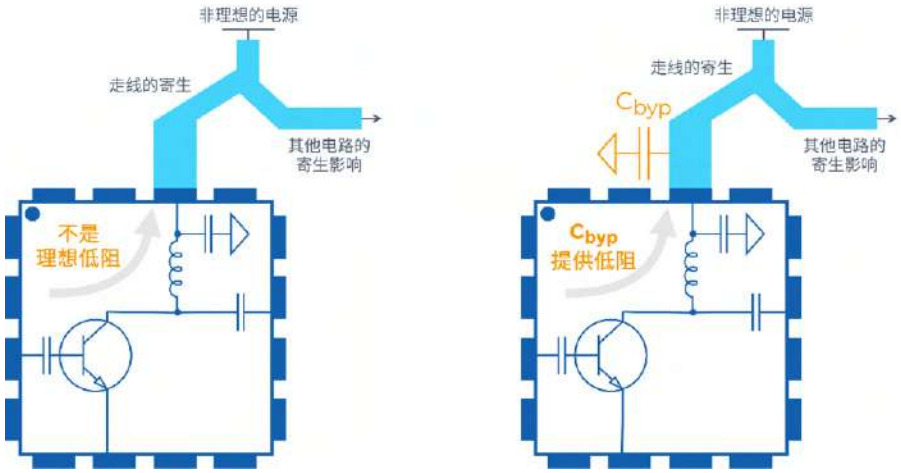


图 5.58: 射频芯片应用中，需要片外 Bypass 电容提供低阻路径

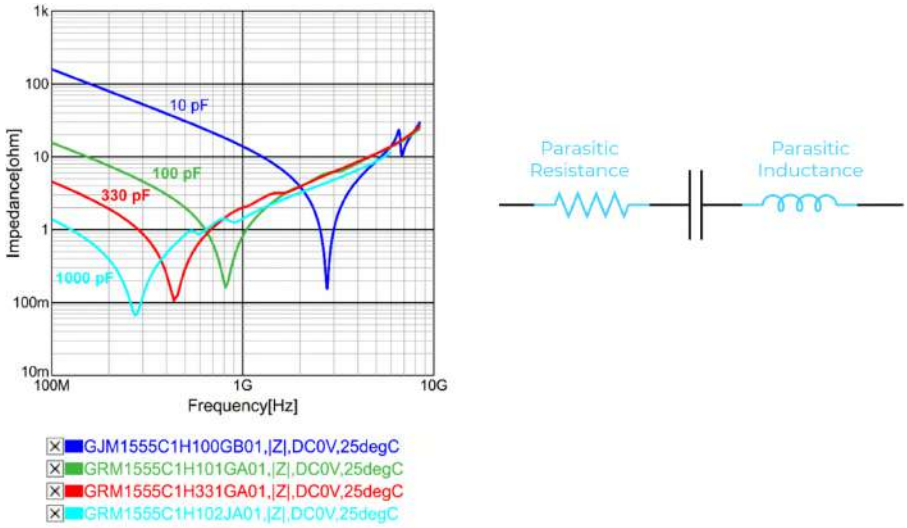


图 5.59: MLCC 阻抗随频率的变化及其等效电路

在 Bypass 电容设计中，需要注意的是，由于寄生电感的存在电容无法在全频段内保持低阻。随着频率升高，寄生电感与电容谐振，超过谐振频率之后，电容表现为感性，频率继续升高阻抗增大。

一般 MLCC 电容寄生电感为 0.3nH，可以以此来计算不同容值电容的谐振频率。图5.59为 Murata 不同容值电容串联阻抗的测试值 [1]，可以看到，1000pF 电容低阻区域（< 3Ω）在 1GHz 以下，而 10pF 所覆盖的范围在 2~3.5GHz 附近。

为了使全频段得到较低的阻抗，需要并联多个电容，容值从小到大排列。在 4G/5G PA 设计和应用中，一般采用 100pF、100nF 及 1F 的 Bypass 电容，分别覆盖 GHz、几十 MHz 及十 MHz 以下频率。其中第一个 100pF 量级电容大多在芯片内部实现，100nF 及 1F 级别电容需要在芯片外部实现。图5.60为慧智微 5G L-PAMiF 产品 S55255-12 应用手册中的 Vcc Bypass 电容设计建议 [2]，片外建议就近放置 100nF 及 1F 两颗 Bypass 电容。

除了在电源提供低阻路径外，Bypass 电容还可以在其他射频通路提供低阻路径。例如图5.61分别为在发射极提供射频旁路路径，减少发射极负反馈电阻 R_E 对射频增益的影响；以及利用电容 Bypass 特性，将直流路径与射频路径分开，在一路直流供电的 Casecode 连接叠管中, 实现两级共源放大器 [3]。

5.3.3 Decoupling 电路

如果需要将不同的电路模块做隔离，防止他们之间的相互干扰，那么最好的方法是减少电路模块之间的线路的共用。不过，这通常在设计中是不现实的。不仅如此，在一些复杂电路设计中，甚至一些共用走线时的准

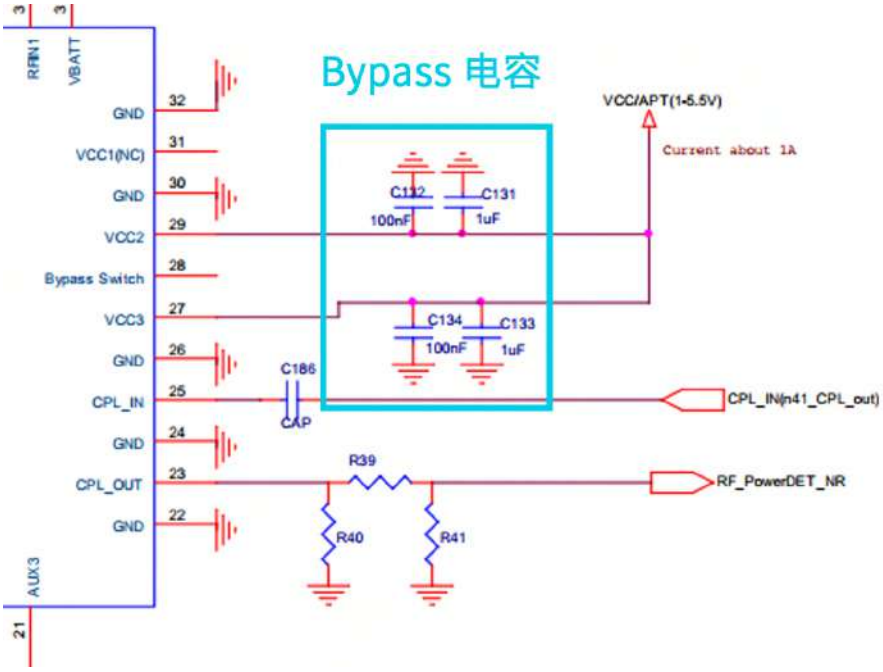


图 5.60: 慧智微 5G L-PAMiF S55255-12 产品 Bypass 电容应用建议

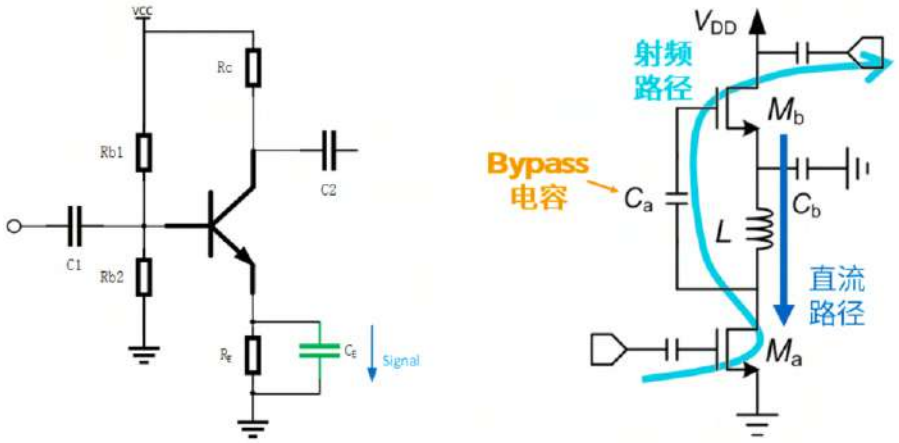


图 5.61: 射极电阻 Bypass 及采用 Bypass 电容电流复用放大器

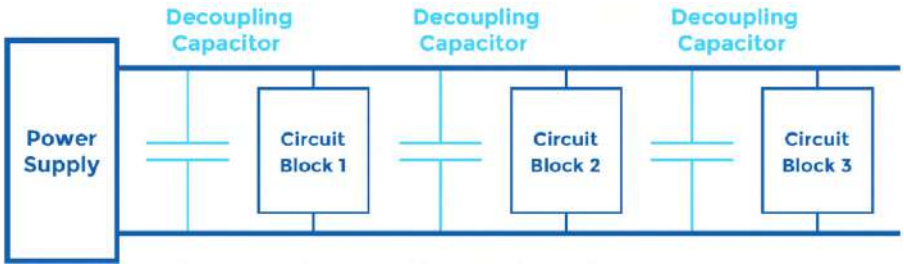


图 5.62: Decoupling 电路在系统中的位置

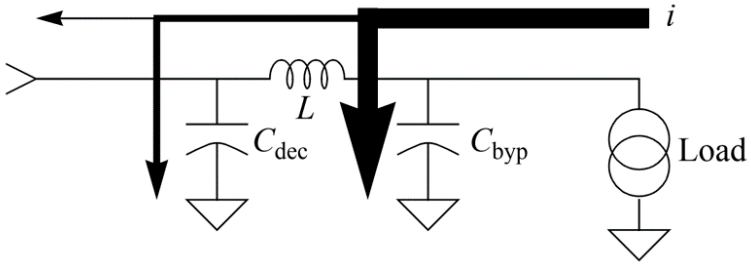


图 5.63: 典型 Decoupling 电路中的电流路径

则，如“电源要星型走线”等也无法实施，这时，就要设计好 Decoupling 电路，完成各电路模块间的 Decoupling。图5.62为 Decoupling 电路在系统中的框图。

典型的电源/射频 Decoupling 电路 [4] 如图5.63所示。在电路中，串联通路上采用了串联电感使串联通路在高频时表现出高阻，这可以帮助 Bypass 电容发挥旁路特性，确保高频噪声可以充分的旁路到地。其次，电路传输特性表现为低通特性，这样流经网络时高频电流将被抑制，只有低频电流可以流过，保持了供电的稳定。

如果说 Bypass 电路提供的低阻通路是为自身电路服务，那 Decoupling 电路就是为整体系统设计的去耦网络，需要将整个系统联合考虑。在 Decoupling 电路设计中，核心是要找到耦合线路以及信号环路，并采用有效的措施切断耦合线路，最小化信号的干扰环路。

在 Decoupling 电路设计中，有以下要点需要注意：

1. 定位干扰源，并识别耦合环路。如果存在多个环路，则每个环路都需要做 Decoupling 处理。例如有些电路需要正压、负压及数字电路控制，在这些 pin 脚连接上都需要做去耦处理。
2. 在 Layout 时，将 Decoupling 电路尽量靠近干扰源放置，以防止过长走线时走线上产生互感效应。例如芯片电源、数字控制端口的 Decoupling 电路应该靠近芯片端口就近放置。
3. Decoupling 电路必要的时候需要加入损耗型元件，如电阻等。
4. Decoupling 电路中所使用的器件在关心的频率有良好的射频特性，例如 Bypass 电容必须在所关心的频率呈现低阻（如 Bypass 章节分析），串联电感要注意其谐振频率。一般 Decoupling 电路使用的是贴片陶瓷电容，而不用长引出脚的陶瓷电容、电解电容。如图5.64所示。

在 Decoupling 电路设计中，需要特别注意的是 Decoupling 电路接地点的选取。一定要识别并选取有效的 Decoupling 节点。图5.65针对高速运算放大器展示了两种负电源 V-的 Decoupling 的设计方式 [5]。

在第一种 Decoupling 中，选取节点在 V-电源的某个位置中，这样选择之后，高频信号环路需要经过 V-电源以及较长的地线才连接到 Decoupling 电容，电源及地线上接收到了大量的高频干扰，不是有效 Decoupling。

第二种 Decoupling 中，将 Decoupling 电容直接跨接在在芯片 V-电源出口及信号参考地上，高频信号保持在较少环路，对电源线及地线的低频影响小，是有效 Decoupling。

讨论了这么多 decoupling 的设计，那么在实际电路中 decoupling 到底有什么影响呢？我们以 ADI 高速低噪声运放 AD9631 为例进行说明。图5.66是 AD9631 驱动 100Ω 负载时的输出频谱（输出信号 20MHz，$2V_{P-P}$[6]。左图采用了合适的 decoupling，频谱很干净，在 2 次谐波 40MHz 处谐波失真为-70dBc 左右；右图是去掉 decoupling 后的频谱，可以看到明显的噪声。因此合适的 decoupling 电路十分重要。

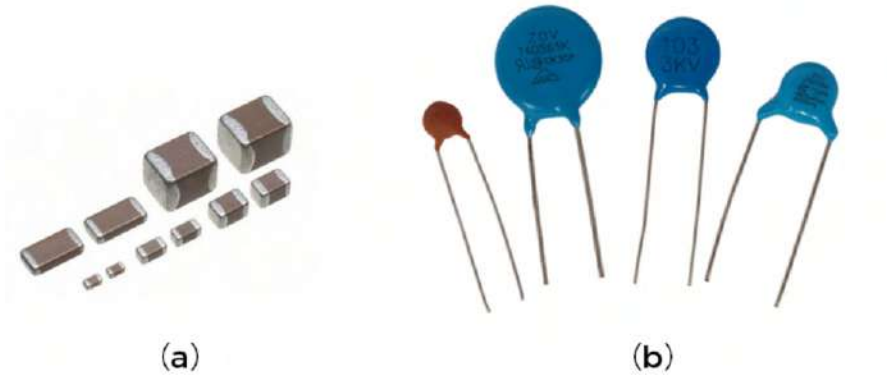


图 5.64: （a）适合做高频 Decoupling 的贴片电容；（b）不适合做高频 Decoupling 的长引出脚电容

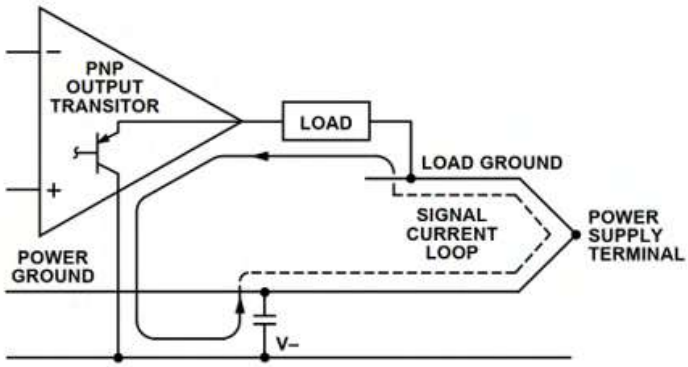


Figure 3a. Decoupling for Negative Supply Ineffective

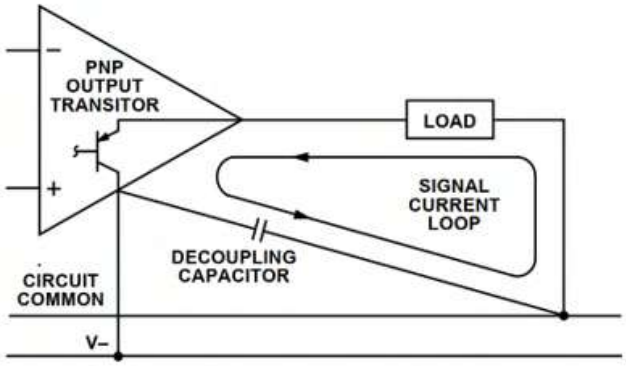


Figure 3b. Decoupling Negative Supply Optimized for "Grounded" Load

图 5.65: 非有效的 Decoupling 及有效的 Decoupling

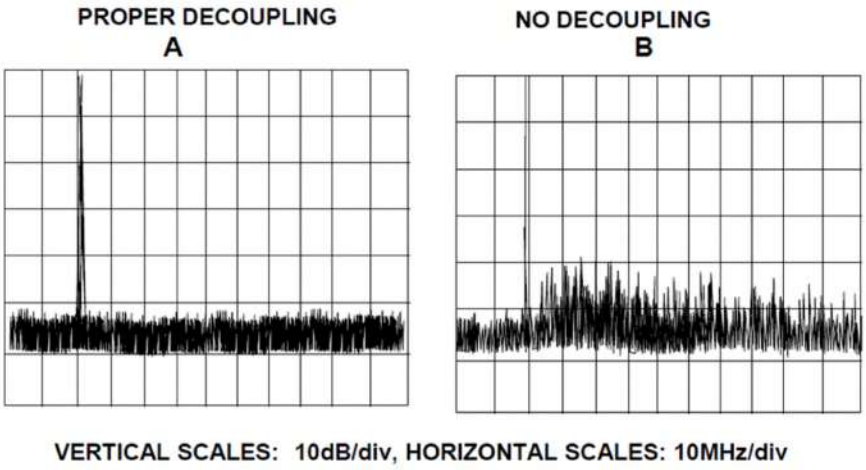


图 5.66: Decoupling 对 AD9631 性能的影响

	Bypass	Decoupling
设计目的	为射频提供低阻路径	隔离不同模块或电路
设计思路	从器件角度出发设计低阻路径	从系统角度设计低通去耦网络
关注指标	节点阻抗Z11	传输系数S21
注意要点	1. 寄生电感带来的谐振 2. 需要多个电容实现宽带低阻	1. 识别射频环路，并在最有效的地方放置 2. 尽可能离干扰近的地方放置

图 5.67

5.3.4 Bypass 与 Decoupling 的比较

通过以上比较，可以看到因为设计目标的不同，Bypass 与 Decoupling 在设计方法和注意事项中有所不同。二者对比总结如图5.67：

您在应用中是否遇到过因为 Bypass 不理想，或者 Decoupling 不充分带来的问题？欢迎扫码关注公众号在线留言讨论。

参考文献

1. analog filters frequency response.” 2019 International Seminar on Electron Devices Design and Production (SED). IEEE, 2019.
2. S55355-12 Application Note, SmarterMicro.
3. Lin, Yi-Jing, et al. ”A 3.1–10.6 GHz ultra-wideband CMOS low noise amplifier with current-reused technique.” IEEE Microwave and Wireless Components Letters 17.3 (2007): 232-234.
4. Ken Kundert, “Power Supply Noise Reduction”, Designer’s Guide Community, 2019.
5. Paul Brokaw. ”An IC Amplifier User’s Guide to Decoupling, Grounding, and Making Things Go Right for a Change.” Analog Devices Application Note, AN-202, Rev.B.
6. Hank Zumbahlen. ”Basic Linear Design”, 2007 Analog Devices, Inc.

在线阅读

本文摘自慧智微微信公众号，扫描以下二维码，在线阅读



5.4 协议再演进：从 5G-Advanced 到 6G

作者：彭洋洋 博士

5G 可能是人类历史上发展最快的通信协议。

自 2019 年 5G 开始部署以来，5G 在全球迎来快速发展。根据 GSMA（全球移动通信系统协会）的统计，从部署元年开始计算，3G 经过了 11 年达到了 10 亿用户，4G 用了 8 年，而 5G 只用了 4 年时间。到 2023 年，全球 5G 用户已达 15 亿。

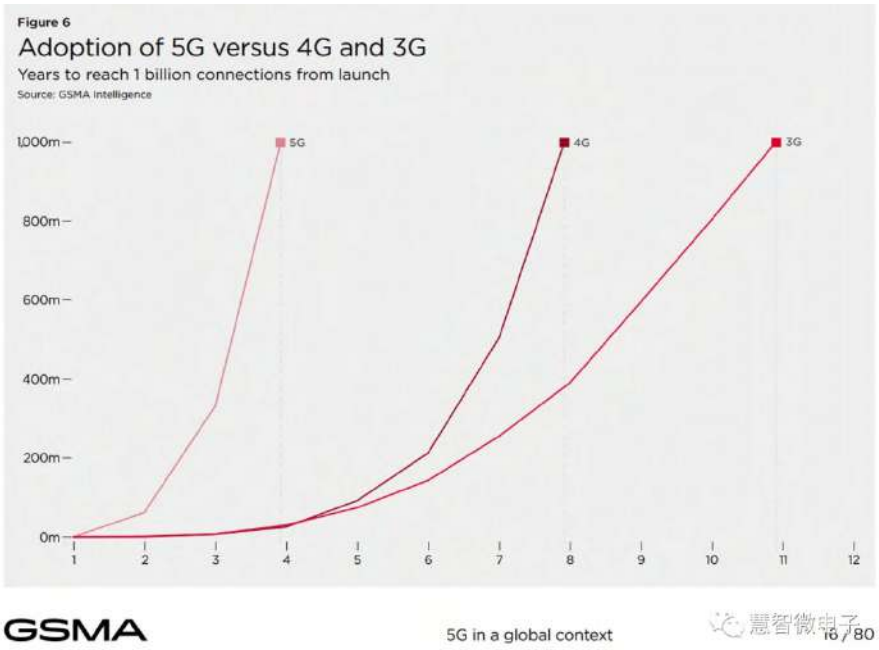


图 5.69: 不同协议达到 10 亿用户所用时间

伴随着 5G 的快速部署，5G 已完成第一阶段商用使命已完成，5G 也进入到深度优化阶段。2021 年 4 月，3GPP 正式将 5G 演进的名称命名为“5G-Advanced”，2021 年 12 月，作为 5G-Advanced 的首个协议版本 Rel-18 首批项目正式立项，5G 发展进入新阶段。

同时，在 5G 部署及商用成功后，行业领导厂商又马不停蹄地开始 6G 协议的制定。2023 年 6 月，在 ITU-R（国际电信联盟-无线电通信部门）发布 IMT-2030（6G）总体目标建议书，首次向全世界正式发布了 6G 部署时间、6 大应用场景以及主要能力提升。6G 协议有条不紊开始推进。

5G-Advanced 究竟先进在哪里？6G 还会做哪些改进？带着以上问题，我们一起来对 5G 通信协议的演进做一个梳理，与大家一起探讨未来无线通信可能带来的变化。

5.4.1 持续演进的 5G



图 5.70: 2030 年之前，全球移动通信发展主要节点

根据 GSMA 对未来 7 年的预测，在接下来的 7 年时间里，全球蜂窝移动终端连接数目将由 85 亿继续增长至 95 亿。其中 5G 连接数目增长迅速，将从 2023 年的 15 亿连接数，增长至 2030 年的 50 亿。占总连接数的比例从 20%，增长至超过 50%。

为了满足移动通信的不断发展，自 3G 时代起，协议发布之后不再是一成不变，而是不断演进。例如 WCDMA 协议在后期迭代出现 HSPA（High Speed Packet Access），以提供更调整的上下行速率。这一做法在 4G 时代得到了充分的使用。

ITU-R 于 2008 年发布 4G 协议技术要求和性能目标，定义 4G 需要满足 1Gbps 的峰值下载速率和 500Mbps 的峰值上传速率。这在当时 3G 时代有几十 Mbps 的上传下载速率来说提升是不可想象的。3GPP 一直在做对 3G 协议的改进和提升，自 2004 年开始，3GPP 不断采用新的信号处理技术、调制方式、多天线技术等提升无线数据网络的容量与速度。3GPP 认为无线通信技术的提升是一个长期工作，于是将这个项目的名称称为“长期演进”（Long Term Evolution），简称为 LTE。

LTE 并不是完全满足 ITU-R 4G 需求的协议，在 2009 年 LTE 的第一个版本（R8 版本）商用时，其只有峰值 150Mbps 的下行速率和 50Mbps 的上行速率。为了达到 4G 协议的要求，LTE 进行了多次版本升级，最终在 2011 年发布的 R10 版本上，才达到下行 1Gbps，上行 500Mbps 的峰值速率，真正达到了 ITU-R 的需求，成为被 ITU-R 认可的 4G 协议。3GPP 也给这了这一版本协议一个新的名称：LTE-Advanced。所以严格意义上讲，LTE 并不是 4G 协议，真正的 4G 协议是从 LTE-Advanced 开始。

ITU-R 对 5G 的需求定义是满足 20Gbps 以上的峰值下行速率，这一次 3GPP 并没有打折，在 5G 的第一版本协议 R15 中，协议所能达到的峰值速率即达到了 20Gbps。但 5G 协议并没有止步于此，为了达到 5G 定义之初“铁三角”的应用场景，5G 首版协议 R15 完成 eMBB 场景的主要需求定义后，R16 版本又逐步完善。

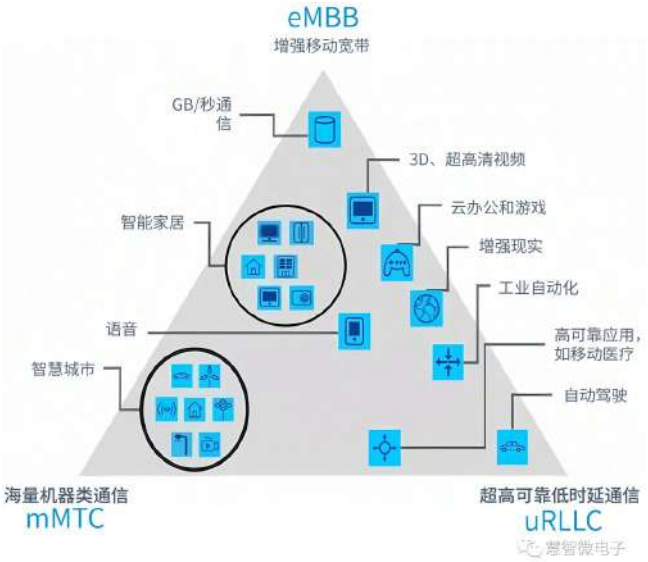


图 5.71: ITU 所定义的 5G 应用场景“铁三角”

其中，在 eMBB 场景中完善了毫米波的定义，并加入了 SA（独立组网）的定义，使 5G 成为真正的独立宽带网络。

在 uRLLC 场景中，定义也逐步完善，并加入了 NR V2X 的定义，完成基础车联网的定义，使 5G 低延时特性得到更大发挥。

在 mMTC 场景中，将 NB-IoT、eMTC 协议纳入到 5G 核心网的支持中来，补齐海量物联场景的定义。并且，在网络基础能力定义中，定位、节能、智能化功能也不断增强。5G 标准从基础走向标准。

2022 年，5G R17 协议冻结，完整的 5G 协议再走向增强。初步的 NTN（Non-Terrestrial Network，非地面网络）功能被定义进来，5G 网络从地面走向天空。同时车联网、物联网协议得到增强，定位、人工智能的结合持续升级。随着 5G R15 标准的基础定义，R16 的完善，R17 的增强，5G 第一阶段的使命已基本完成。



图 5.72: 5G 协议中的演进

2021 年 12 月，3GPP 完成 R18 协议首批项目立项，立项项目涵盖 eMBB、uRLLC、mMTC 多个场景，涉及 27 个项目，这些项目对 5G 无线通信的容量增强、人工智能应用、空天一体、轻量化终端等多个方面进行增强和完善。3GPP 对这些项目寄予厚望，认为这些项目将是 5G 完成第一阶段之后的性能再次提升，并且是未来 6G 演进中的重要铺垫。3GPP 也将 R18 及之后的协议定义了新的名称，即“5G-Advanced”，用来区分 5G 的新阶段，展示 5G 的技术演进与创新。



图 5.73: 5G-Advanced 的演进

虽然 5G-Advanced 仍在研究中，一些技术还并未成熟商用，不过有几项特色的项目仍值得关注。

比如 5G 与人工智能的融合，5G-Advanced 有望通过人工智能对物理层的优化、对资源管理的优化、对故障检测与预测的优化，来提高物理层的性能与效率，提高网络的资源利用率与分配效率，提高网络可靠性与稳定性。

再比如对双工技术的研究上，全双工的概念再次被提出并展开研究，通过对自干扰的消除，使得上下行信号可以同时、同频进行传输，极大地提高了频谱利用效率。

另外，5G 轻量终端相关的 5G RedCap 协议，空天一体相关的 5G NTN 协议，都将继续探讨与发展。

5.4.2 6G 应用场景正式发布

随着 5G 的正式商用，6G 的技术讨论也悄然到来。虽然不断有声音说，6G 技术还为时尚早，因为 5G 商用还未全面普及，5G 的优势还没有完全体现，5G 需要的提升也未有定论。不过随着无线通信管理组织的探讨，行业领先公司的研究，6G 的目标逐渐清晰。

一般在全球性新协议的开发过程中，主要分为三个步骤进行，第一个步骤是国际电联等全球无线电管理机构发出愿景性标准，定义新协议的主要实现目标及主要时间点，第二个步骤是协议组织（如 3GPP、IEEE 等）制定技术方案，将协议进行标准化，第三个步骤是行业组织推动商业落地。目前的 6G 协议在第一个步骤完成，到第二个步骤启动的切换阶段。



图 5.74: 新协议的开发过程

经过近两年的讨论，ITU 在 2022 年 6 月确定了 6G 的主要时间表。在 ITU 的指引中，6G 发展将分为三个阶段进行。

第一阶段是愿景定义阶段，这个阶段在 2023 年 6 月之前完成，将完成愿景定义，新应用趋势与技术趋势的识别，新场景、新频谱的定义。

第二阶段是确定需求和评估方法，这一阶段主要是指定技术性能要求，并且收集候选技术方案，确定评估方法，这一阶段将在 2026 年之前完成。

第三阶段是输出规范阶段，这个阶段主要目标是输出规范，发布标准，将在 2030 年之前完成。也就是说，将在 2030 年实现 6G 的商用。射频无线协议的 10 年一更迭的节奏仍然在继续。

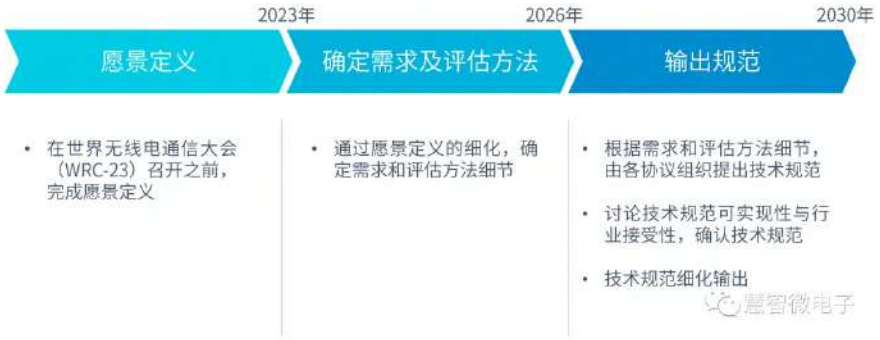


图 5.75: 6G 的主要时间点

2023 年 6 月，ITU 将 6G 场景、能力需求做了发布，顺利实现 6G 第一阶段的愿景定义。根据 ITU 的定义，6G 应用场景归为 6 大类。其中 3 大类为原来 5G 场景的增强，3 大类为 6G 的新增。

3 大类增强的场景是将 eMBB 升级为沉浸式通信，将 mMTC 升级为海量通信，将 uRLLC 升级为超级高可靠低延时通信。3 个新增的场景为：无处不在，AI 集成，感知集成。

另外，6G 还定义了四大设计原则，分别为持续性、连接未联，无处不在的智能，安全/隐私/弹性。

ITU 所定义的 6 大场景和 4 大设计原则相互协同，共同构成 6G“场景之轮”（Wheel Diagram）。

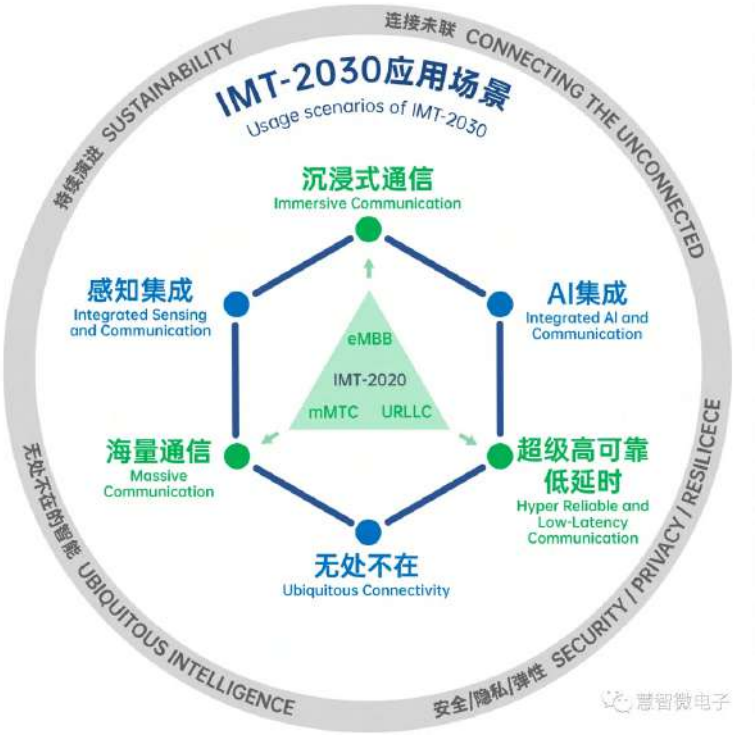


图 5.76: ITU 定义的 IMT-2030 “6G”场景之轮

ITU 也对 6G 的能力做了定义，射频无线通信网络的 15 大能力再次得到提升或重新定义。

在这些能力提升中，有 9 项能力是在原有网络能力中做的提升，比如在过去几代协议演进中被不断提及的上峰值速率、连接密度、延时等，这些能力将再次得到数量级的提升。

6G 也重新定义了新的 6 大能力，这 6 大能力分别是 AI、持续演进、感知、覆盖、交互、定位。在 6G 的新能力定义下，6G 必将实现更智能化的射频无线网络。

ITU 将 6G 的能力称为“6G 能力调色板”，9 个能力提升与 6 大能力增强使 6G 应用更值得期待，多姿多彩。

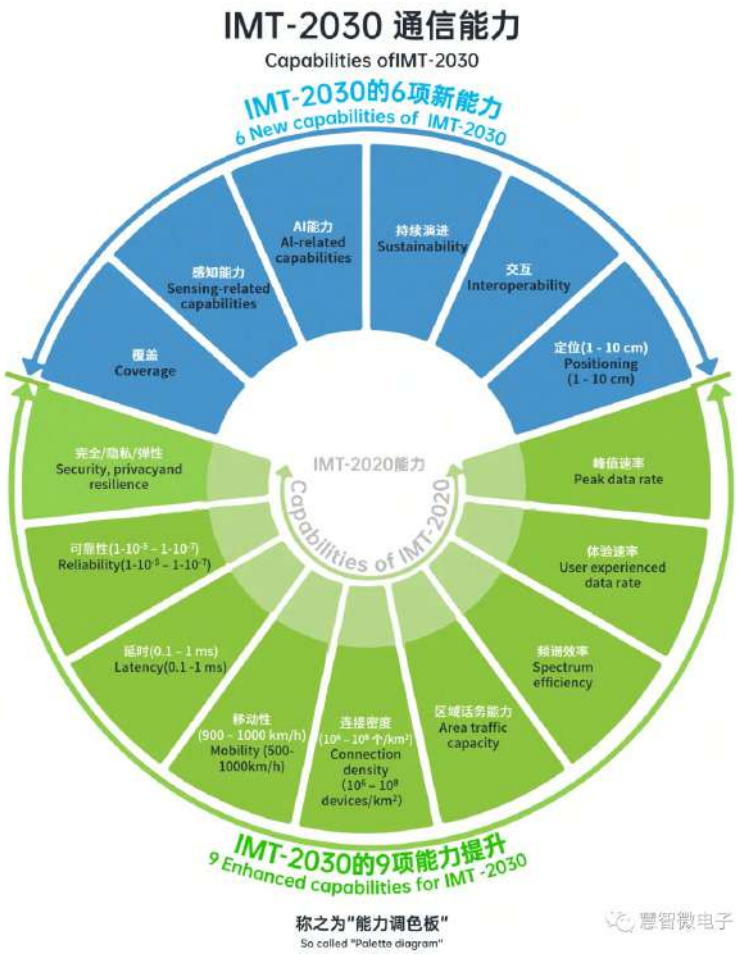


图 5.77: 6G 能力展望之“6G 能力调色板”

在 6G 的定义下，无线连接的能力再次得到提升，如果说 5G 应用场景的“铁三角”让无线网络从人的互联走向万物互联，那么 6G 的“六边花”就将实现万物互联下的能力加持。

6G 的到来，使得射频无线网络不再是连接的延伸，而成为能力边界突破的助力。6G 强大的通信能力将使物理世界、数字世界界限不再明显，两个世界相互交织，实现智能的全息网络。6G 无处不在的连接将使人类、机器、海量传感器实现互联互通，共同构建起一个高效协同的体系。6G 还将使卫星通信、地面通信、空中通信等多种接入方式实现完美融合，实现连接的全域覆盖。6G 也将充分利用人工智能、边缘计算等新兴技术，让连接变地更加强大与智能。6G 不仅继续实现万物互联，6G 还将带来万物突破能力边界的新机遇。

5.4.3 总结

很少有一门技术像射频技术一样，依靠不到 200 年前的四个公式，就将射频技术的发展牢牢框死。1873 年，麦克斯韦在他伟大著作《电磁理论》中系统总结了他的方程，这个方程用四个公式揭示了电磁波的本质。射频技术在接下来 150 年快速发展至今，但仍未超脱麦克斯韦方程的范围。射频技术好像没有本质更新。

随着信息论、集成电路技术在射频技术中的应用，人类对于无线连接的需求极大地得到了满足。几乎人手一部的手机、无处不在的网络、越来越多的互联设备，好像人类已经将所有能想象的物体连接上了网络。

再也没有本质更新的技术，应用需求好像也已经被全部满足。这不禁让人提问，我们还需要更强的射频技术吗？

这个问题在每个代际的无线通信协议发展中都会被人提及。在手机诞生之前，大部分人类觉得有广播就够了；第一代移动通信诞生后，很多人觉得人手一部能随时打电话的手机已是神话；第三代移动通信让人类体会到了“数据连接”；第四代移动通信将互联网变成了“移动互联网”；第五代移动通信开始实现汽车、卫星等的

万物互联……无线通信技术不断用它不断强大的能力，刷新人类对于射频技术的认识，更不断拓展人类对这个世界认识的方式与边界。

6G 到来将继续开拓人类认知边界。6G 带来的无处不在的连接、更加智能网络、通信与感知的集成、沉浸式的通信，将使这个世界再次得到改变。6G 不仅是无线连接的功能实现，而将成为万物能力拓展的手段。未来更加强大的互联世界正缓缓到来。

您对 6G 到来的世界怎么看？欢迎留言讨论。

参考文献：

1. Global Mobile Trends 2023, GSMA
2. ITU towards “IMT for 2020 and beyond”, ITU-R, https://www.itu.int/en/ITU-R/study-groups/rsg5/rwp5d/imt-2020
3. Framework and overall objectives of the future development of IMT for 2030 and beyond, ITU-R, https://www.itu.int/en/ITU-R/study-groups/rsg5/rwp5d/imt-2030

在线阅读

本文摘自慧智微微信公众号，扫描以下二维码，在线阅读

5.5 手机“射频增强”的技术方案

作者：彭洋洋 博士

尽管有声音说近年来手机革命性变革已不多，但仍然有厂商的突破让人眼前一亮。

近日，在荣耀 Magic5 系列手机的发布会上，荣耀 CEO 赵明表示：“Magic5 Pro 和至臻版搭载自研、业界首颗射频增强芯片 C1，实现对多个传统信号弱势场景的全面优化”[1]。从发布会信息看，Magic5 Pro 手机抓住了出电梯后快速回网、地库弱信号增强等痛点场景进行优化，给用户带来良好的 5G 信号体验。荣耀对于射频性能的关注，使大家重新审视“通信”这个手机里本该最基本的功能。

虽然智能手机是以无线通信技术为基础的智能设备，但通信技术往往只是“背后英雄”。华为、vivo、OPPO、小米等厂商都发布过射频信号优化技术，不过相比于看起来高高在上的射频技术，让用户更愿意为之买单的，可能还是那颗几 nm 工艺的 CPU。

从荣耀的新机发布看，手机厂商好像找到了更好的优化方法：**与其堆砌用户听不懂的高深技术，还不如在用户使用的场景中做增强和优化，切实改善用户使用体验。**

在用户的这些体验中，“连网不掉线”无疑是最重要的体验之一。于是，我们就围绕“射频增强”这一话题，谈谈有什么思路可以实现“射频增强”。

5.5.1 无线通信的“链路”

在分析无线射频信号时，通常采用“链路预算（Link Budget）”的方式来计算链路中各路径的损耗及各节点的信号强度。无线通信的“链路”是指无线信号传输过程互相连接的各段路径。

在手机到基站的通信过程中，射频信号主要通过以下路径：

- 手机中收发机发出的射频信号被功率放大器放大，并通过滤波器、天线开关等通路损耗后，在手机天线测试口达到协议所规定的传导功率；
- 传导信号经过天线损耗后，经天线辐射发出；
- 手机发出的信号经过一定的空间距离后，到达基站天线处；
- 由于基站天线可提供一定的定向增益，所以信号等效被天线增益所放大，形成基站侧传导功率；
- 传导功率再经过一定的路径损耗后，进入基站接收通路进行处理。

以上路径可以用图5.79表示。

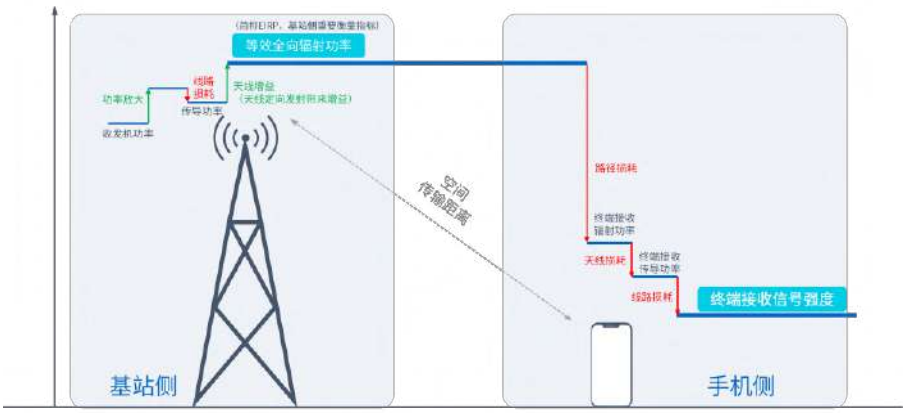


图 5.79: 手机发射信号被基站所接收的通信链路示意图

在以上路径中，基站所发出的 EIRP 功率，及终端所接收的信号强度功率是手机接收路径的两个主要功率值。在基站 EIRP 指标确定、路径距离确定的情况下，可通过减少天线损耗、线路损耗，来增强终端接收信号强度。

以上为手机与基站通信时的基本路径。虽然在实际应用中，会有多个频段和模式同时工作、5G/4G/3G 不同模式对接收信号强度需求不同等具体实现和计算，但传输路径与上述过程基本相同。

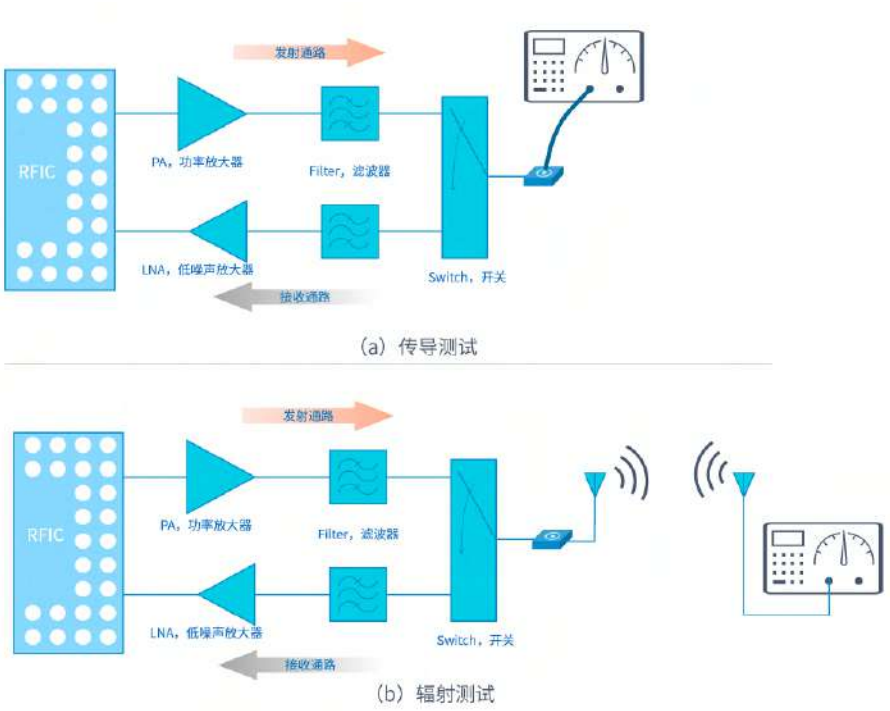


图 5.80: 传导和辐射测试

5.5.2 “信号增强”的方法

了解了以上链路，就可以针对性的做“信号增强”。以手机角度考虑，在信号增强的处理中，主要分为以下三种处理方式：

- 发射信号增强
- 接收信号增强
- 传输匹配增强

5.5.2.1 发射信号增强

观察图 1 所示手机发射路径，手机在发射时，传导功率与辐射功率是两个重要的功率指标。

传导功率通常是指，将手机发射功率直接接入仪表这种方式得到的功率值，目的是了解手机发射功率的能力。由于传导测试通路稳定，所以指标清晰，3GPP 等协议组织所发布的众多指标也是以传导测试方法作为标准。

辐射功率是指手机在接入天线之后，向外辐射的功率大小。辐射功率测试更接近于实际使用环境，但由于测试受天线设计、天线阻抗的影响较大，同时，影响辐射功率测试准确度的因素也比较多，比如测试环境中的探头补偿，转台精度，暗室吸波材料质量等，这些都导致辐射功率的测试实现起来更加复杂。传导和辐射测试如图5.80。

既然传导功率和辐射功率是影响手机发射信号的两个重要指标，那如何才能增强这两个功率的大小呢？

传导功率的增强

虽然传导功率提升会对手机发射信号的传输距离有直接的提升，但由于功率提升会带来对人体的辐射、对其他通信设备的干扰等问题，协议一般对手机的传导功率进行严格的限制。

以 5G 手机为例，3GPP 规定手机有 PC3、PC2、PC1 功率，PC 为 Power Class 的缩写。图5.81为 3GPP 协议中对不同频段可以发射的功率等级，及这些功率等级的功率定义 [2]。

以中国三大运营商均推荐支持的 n3 频段为例，此频段手机发射的传导功率在 23dBm，容限是 +/- 2dB，即此频段手机传导功率需要限制在 21 ~25dBm 之间，想发射更大的功率是不可能的。所以，是没办法通过大幅度

Table 6.2.1-1: UE Power Class

NR band	Class 1 (dBm)	Tolerance (dB)	Class 1.5 (dBm)	Tolerance (dB)	Class 2 (dBm)	Tolerance (dB)	Class 3 (dBm)	Tolerance (dB)
n1							23	±2
n2							23	±2[3]
n3							23	±2[3]
n5							23	±2
n7							23	±2[3]
n8							23	±2[3]
n12							23	±2[3]
n13							23	±2
n14	31[6]	+2/-3					23	±2
n18							23	±2
n20							23	±2[3]
n24							23	+2/-3[3]
n25							23	±2[3]
n26							23	±2[3]
n28							23	+2/-2.5
n30							23	±2
n34							23	±2
n38							23	±2
n39							23	±2
n40					26	+2/-3	23	±2
n41			29[5]	+2/-3[3]	26	+2/-3[3]	23	±2[3]

图 5.81: 3GPP 规定的 5G 手机发射的功率值

的提升传导功率，来让手机的发射信号变得更强的。

不过，即使是 3GPP 所定义的功率范围，也是有容限的。比如在上表中，大部分频段的上下限之间差别在 4dB。以 3dB 就是功率差一倍简单估计，同样是 PC3 功率，靠近上限的手机是靠近下限的手机发射功率的两倍，也是一个较为可观的值了。

但手机发射大功率需要较大的电流支撑，这对电池是一个压力。对于手机厂商来说，不得不在强信号和长待机时间做个折衷。**射频器件供应商开发的高效率 PA 可以来解决这个折衷问题**。高效率 PA 可以在足够小的电流下，发射出足够大的功率来，从而使手机发射功率可以更加靠近于协议上限，发射性能得以增强。

另外，优化路径损耗，可以减小 PA 到传导天线测试口之间损失的能量，同样可以在电流限定的情况下，使传导能量有所增加。

辐射功率的增强

辐射功率是手机中的重要性能参数，决定了真实场景下手机的通信能力，所以移动终端厂商、运营商以及某些强制认证中都会着重要求此指标能力。

辐射功率的重要衡量参数是 TRP（Total Radiated Power，全向辐射功率），TRP 的计算为传导功率乘以天线辐射效率 η_{Ant}，计算公式为：

$$P_{TRP} = P_{Tran} \times \eta_{Ant}$$

当传导功率一定时，辐射功率主要由天线效率决定。

随着 5G 手机支持频段越来越多，并且随着屏幕面积的进一步缩小，天线的辐射效率也受到明显影响。下图为 Qorvo 展示的 5G 手机天线辐射效率仿真，可以看到，随着天线与地之间净空距离由 12mm 降低到 5mm，天线辐射效率的峰值也从 85% 降低到 35%，辐射能量减少到原来的一半。在当前 5G 手机中，天线净空基本已经到 5mm 以下，给天线设计提出了巨大挑战，同时也影响到手机辐射功率性能。如图5.82所示。

另外，天线带宽的增加，也会引起天线效率的降低。为了保障天线辐射效率，天线设计一般会将频段分组进行支持。根据《5G 移动终端天线设计》所述，典型的天线设计分组如图5.83所示 [4]，采用 7 天线的方式，进行 5G 全频段的支持。

为了进一步增强手机辐射功率性能，还可以对天线进一步拆分，使每个天线性能在较窄的频段内优化。不过会带来成本及手机布局压力的增加。在 2019 年华为发布的 Mate 30 Pro 手机中，天线进行了更细的拆分，采用 14 根天线对 5G 信号进行支持 [5]。如图5.84所示。

在天线净空受限、天线数量受限的情况下，天线辐射效率将受到影响。为了增强天线辐射性能，可以增加调谐器件，如天线 Tuner，以使天线性能可以根据频率和应用场景做针对性的优化，以增强辐射功率性能 [4]。有关

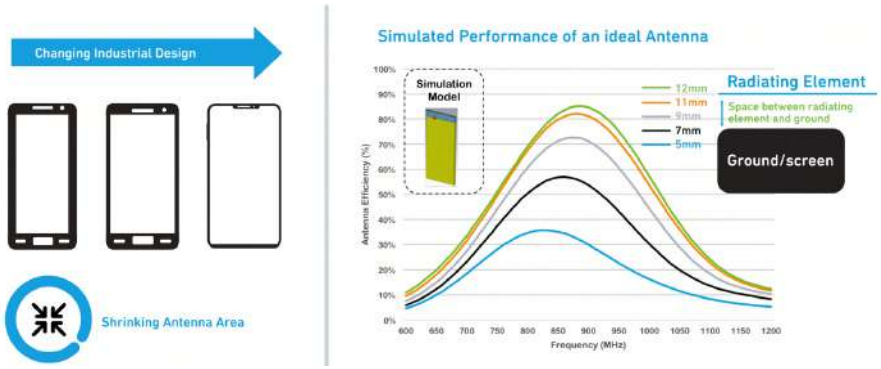


图 5.82: 天线辐射效率与天线净空的关系

	LB	MHB	UHB	n41
ANT0	PRX	PRX	-	-
ANT1	DRX	DRX	-	-
ANT2	-	PRX_MIMO	DRX	-
ANT3	-	DRX_MIMO	-	DRX
ANT4	-	-	PRX	PRX
ANT5	-	-	PRX_MIMO	PRX_MIMO
ANT6	-	-	DRX_MIMO	DRX_MIMO

图 5.83: 经过合并之后的天线分组



图 5.84: 2019 年华为 Mate 30 Pro 手机中的天线设计

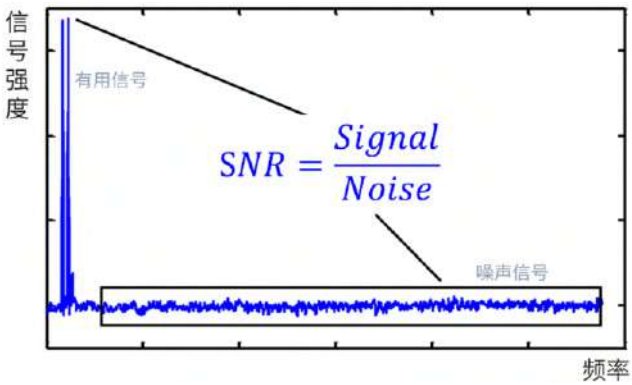


图 5.85: SNR 的定义

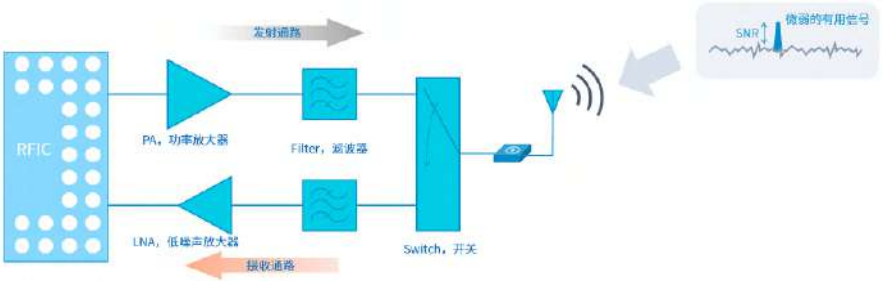


图 5.86: 经过长途跋涉，手机接收到的信号已经非常微弱

天线 Tuner 的设计考虑在第三节的“传输匹配”增强中与其他调谐方式共同讨论。

5.5.2.2 接收信号增强

在如图 2 所示手机接收链路中，基站 EIRP 和传输距离一定时，到手机天线口的无线信号功率也确定。于是，增强手机接收能力的方式主要就剩下减少天线损耗，及减少线路损耗。

由于天线收发的对称性，接收通路减少天线损耗的方式与发射侧相同。减少线路损耗的方式也是尽量缩短接收天线至接收电路之间的距离，减少损耗性器件的数量。以上两个考虑与前一节发射通路的考虑相同，在此不再赘述。

在接收信号增强中，还有一个重要考虑，就是减小其他干扰信号对接收信号的影响。在干扰信号中，“白噪声”和其他频段的干扰噪声是主要的干扰源。以下分别对二者进行讨论。

“白噪声”的处理

白噪声是由电子热震动引起的，在自然界无处不在，是不能够被消除的噪声。白噪声的存在决定了系统通信能力的上限。在室温环境中，白噪声的功率谱密度约为-174dBm/Hz。

由于白噪声的存在，有用信号需要尽可能的远离白噪声噪底，这样接收电路才可以尽可能的将信号分辨出来。为了表征信号强度距离噪底的大小，SNR（信噪比，Signal Noise Ratio）被引入进来，如图5.85。SNR 的计算方式是信号强度除以噪声强度。SNR 是通信中的重要指标，根据通信中基本香农定理的计算，信道的通信容量和 SNR 成正比。

在手机接收信号时，尤其在距离基站较远场景下，接收信号经过长途跋涉，到达手机时已经非常小，甚至接近噪底。这时就需要注意在处理时尽量不再引入额外的噪声，保持已经微弱的信噪比 SNR 不再受到破坏。如图5.86所示。

在放大过程中，尽量保持 SNR 不受破坏的方式就是引入“LNA”，低噪声放大器（Low Noise Amplifier）。

低噪声放大器并不是将噪声减小，事实上，放大器并没有办法区分信号和噪声，二者会同时被放大。低噪声放大器在放大过程中尽量的不去引入额外的噪声，尽量少的破坏信号与噪声之间的比值。低噪声放大器与普通

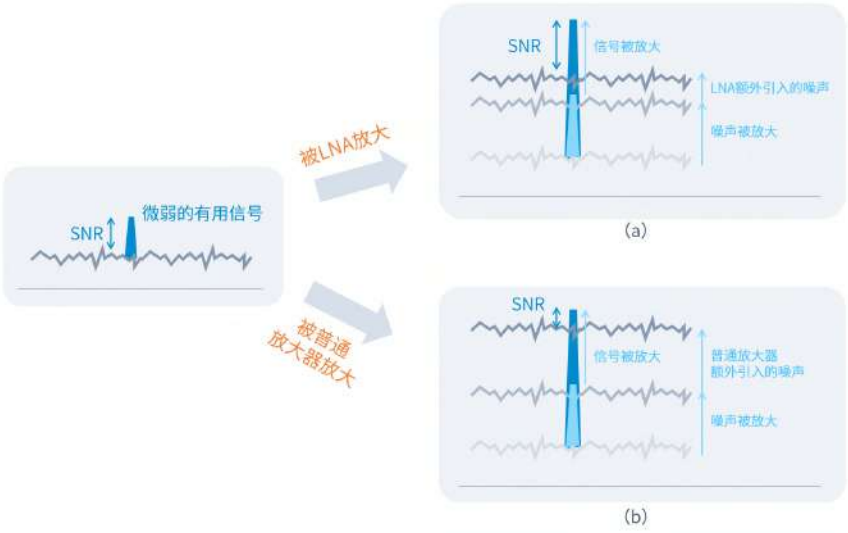


图 5.87: LNA 与普通放大器对微弱信号的处理效果



图 5.88: 现实生活中干扰信号无处不在

放大器对信号的放大效果如图5.87所示。低噪声放大器引入的信噪比恶化要远小于普通放大器。

为了表征放大器对 SNR 恶化的程度，噪声系数指标被提了出来。噪声系数的定义是进入放大器的 SNR，除以经过放大器后输出的 SNR。噪声系数越大，说明信号 SNR 被破坏的越多。

$$F = \frac{SNR_{in}}{SNR_{out}}$$

噪声系数是接收信号增强中重要考虑的指标，对噪声系数的优化和控制，有助于增强手机的接收能力。

“干扰噪声”的处理

除了白噪声外，通信系统中还有其他各种干扰，如图5.88，这些干扰一定要被仔细处理或屏蔽，否则也会对接收信号造成干扰，影响手机的接收性能。

在干扰信号中，有一类干扰信号非常特殊，即 FDD（频分双工）系统中手机自己的发射信号。

FDD 是 3G 时代通信的主流制式，到了 4G 和 5G 时代仍然是手机中重要通信模式。FDD 模式工作下，发射信号和接收信号同时工作在不同的频率上，相互不需要时间切换，信号传输速率高。

但 FDD 也有一个重要缺点，就是一定要处理好发射信号对接收信号的干扰。由于发射信号是大功率信号，通常伴有强烈的接收频段干扰噪声。而接收信号通常较小，如果与发射信号之间的隔离没有做好，则非常容易引起 Desense（接收灵敏度恶化）。

在系统路径中，一般通过双工器来进行发射与接收频段间的噪声滤除。滤除过程如图5.89所示。

以上路径为信号主要传输路径，在其他路径传输时，发射通路与接收通路之间的隔离应优于以上路径。否则发射信号就会从其他路径干扰到接收信号，以上滤波特性将不起作用。

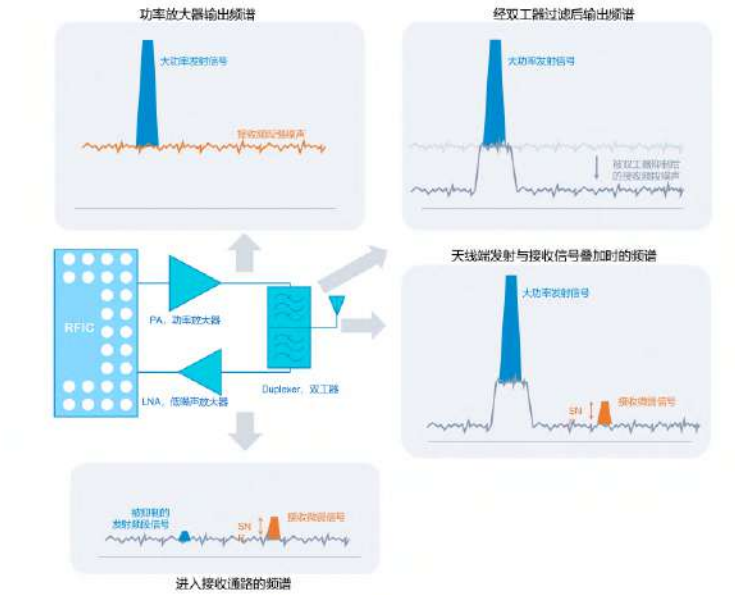


图 5.89: FDD 系统中，对发射噪声的处理

除了发射噪声直接进入接收信号频段外，还有可能发生交调干扰，即同时在发射的两个或多个大功率信号，经过交调之后处于接收频段，影响接收性能。交调干扰分析更为复杂，需要具体问题具体分析，保障接收信号不被干扰。

严格来说对接收信号所做的以上防干扰处理，只能算是保障接收信号不被恶化。不过由于当前手机系统相当复杂，在方寸之间有全球的频段和多种制式，保障接收信号完全不被干扰并不是容易的事。不断的将信号解耦，尽量让接收信号受干扰的场景少出现，也算是在为接收信号的“增强”做的努力。

5.5.2.3 传输匹配增强

射频阻抗的窄带特性

尽管理想的射频系统是各界面间均为 50 Ohm，此时，信号可以达到没有反射的向前传输。但这一理想目标在手机中较难取得，主要原因是：

- 手机需要支持的频段从 0.6GHz 一直到 6GHz，很难有射频器件在这么宽的范围内始终保持 50 Ohm;，用划分子频段支持的方式分而治之，但受限于手机面积，仍然有器件（如天线）需要在较宽频范围内复用；
- 人手持握等外部环境对天线阻抗影响极大，造成内部器件的负载阻抗始终在不断变化中。

阻抗失配带来的影响

阻抗失配会带来传输效率的损失。以 50 Ohm 的源向 50 Ohm 的负载传输效率定为 100% 计算，当负载阻抗由 5 Ohm 变化至 500 Ohm 时，传输效率的变化如图5.90所示。

可以看到，当负载阻抗变化时 100 Ohm 时，只有 89% 的功率有效传输，剩下的 11% 功率被界面反射，当负载阻抗变化至 200 Ohm，传输效率降低至 64%。当负载阻抗由 50 Ohm 向低阻变化时，传输效率同样变低。

阻抗变化除了会引起传输效率的在恶化外，还会恶化通信系统中子电路的性能。如《5G 射频 PA 的 load-line 与 load-pull》中提到，PA 输出特性与输出负载相关，当负载阻抗偏离最优阻抗点时，射频 PA 的输出功率、线性度等指标都会恶化。PA 的最优负载 load-pull 示意图如图5.91所示。

收敛阻抗，增强传输

既然阻抗不匹配对射频传输有明显影响，那是否有办法改变阻抗匹配特性呢？答案是有的，天线调谐、匹配

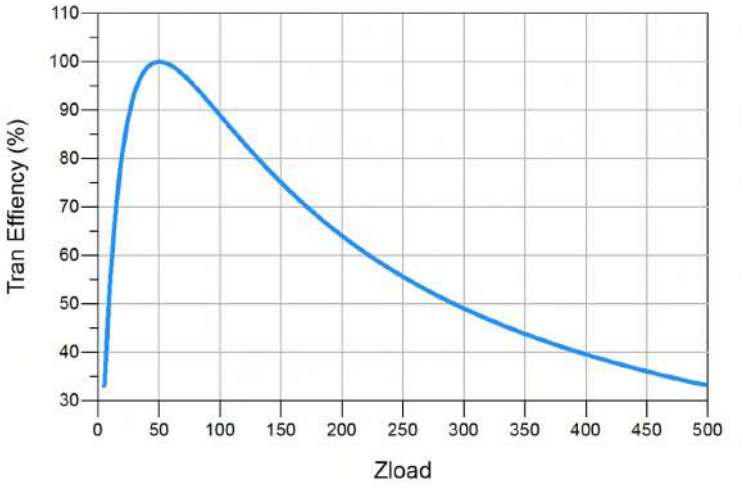


图 **5.90:** 负载阻抗变化时传输效率的变化

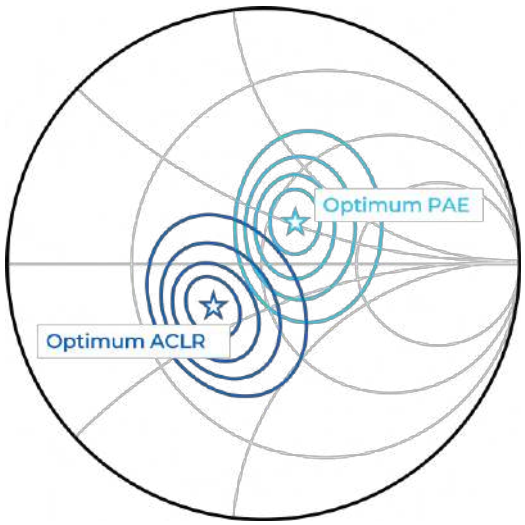


图 **5.91:** PA 的最优负载 load-pull 示意图

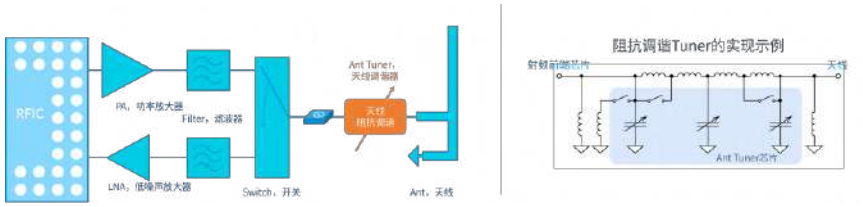


图 5.92: 阻抗调谐天线 Tuner 放置的位置，及实现示例

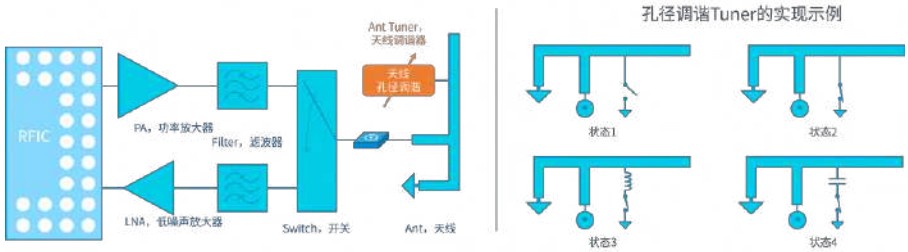


图 5.93: 孔径调谐天线 Tuner 放置的位置，及实现示例

调谐就是其中的代表。

天线调谐器：主动调整天线阻抗

由于天线阻抗无法在较宽的范围内控制在 50 Ohm，那可以在天线处放置一个阻抗调谐器（Antenna Tuner），动态的根据频率变化、天线外部环境变化来调整天线特性，尽可能的达到匹配的状态，使失配状态下的信号得到增强。

天线调谐器有两种设计方法，分别为阻抗调谐（Impedance Tuner）和孔径调谐（Aperture Tuner）。

天线阻抗调谐器放置于天线和射频前端芯片之间，是一个可以根据状态调整的匹配电路，根据天线阻抗的不同，匹配电路可以调整至不同状态，保持左右两边的阻抗匹配。阻抗调谐天线 Tuner 放置的位置及实现示例如图5.92所示。

天线孔径调谐器是利用调谐器件，改变天线的谐振特性，以达到天线阻抗及辐射效率随频率的变化。孔径调谐一般放置在天线结构以内，直接对天线特性进行调整，在天线效率调整上有更大的调整范围。孔径调谐天线 Tuner 放置的位置及实现示例如图5.93所示。

在传导功率被协议限制的情况下，合理设计天线 Tuner 就可以实现天线辐射效率的明显提升，达到射频增强的目的。图5.94为 Qorvo 所展示的某孔径调谐 Tuner 效果 [6]。通过调谐，850MHz 的效率由基线版本中的-3.5dB 左右，提升至-1.2dB；810MHz 的效率由基线版本的-11dB，提升至-2dB，提升效果达 9dB。以 3dB 为两倍来估算，810MHz 的辐射效率提升了 8 倍之多，可见增强效果的明显。

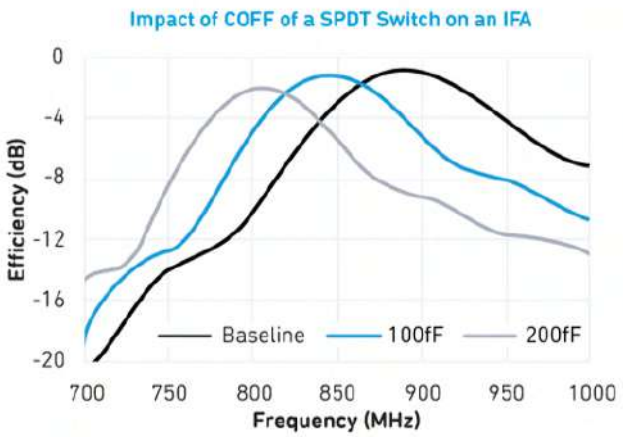


图 5.94: Qorvo 公司展示的某孔径调谐 Tuner 调谐效果

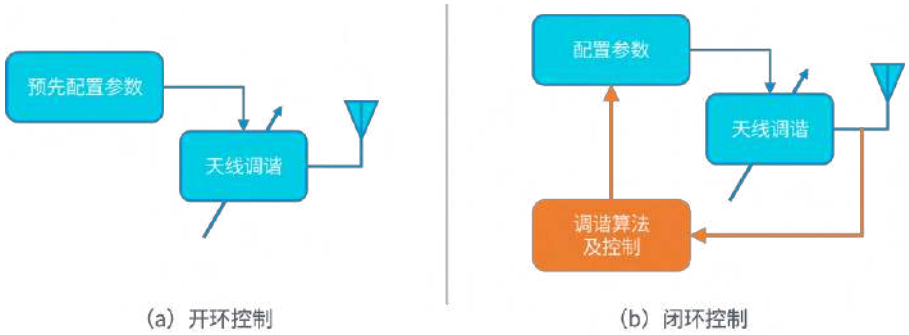


图 5.95: 天线调谐器的两种控制方式

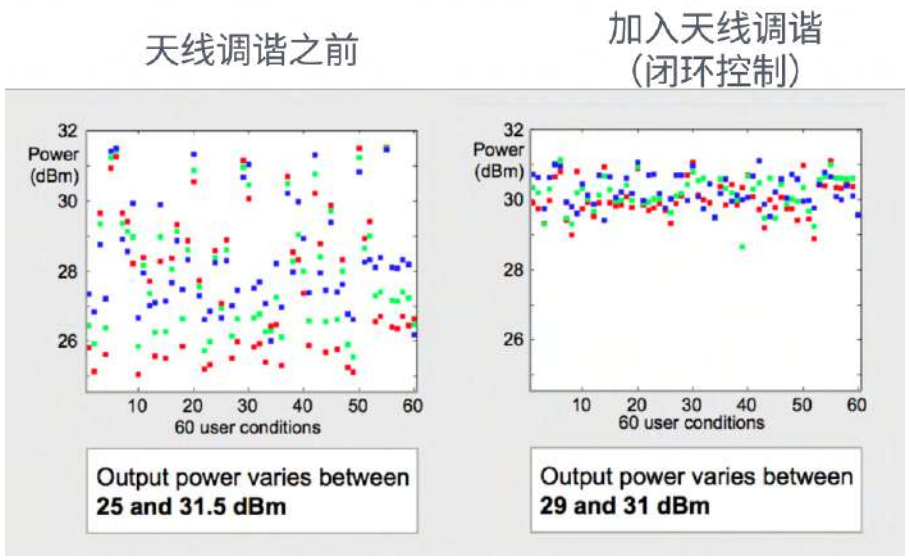


图 5.96: 闭环天线调谐对于射频发射信号的增强

天线调谐器对于天线阻抗的调节，必须是基于天线场景做的针对性调节，所以天线信息的获取就变得尤其重要。根据获取天线信息的不同，天线调谐器在使用中分为开环控制和闭环控制两种方式。如图5.95所示。

开环控制是将天线调谐控制参数预先写入，在使用时直接调用。比如预先针对工作频率的不同，设置不同的匹配参数等。开环控制的好处是控制简单，只需要预先写入一定参数，在后续使用中就可以调用这个确定性的信息。其缺点是无法根据使用场景的不同重新优化，调节场景也限定在预先写入的几个状态范围之内。

闭环控制是构建控制环路，实时检测天线信息，根据算法计算出当前需要的匹配及天线状态后，直接控制天线控制器进行实现。闭环控制的优点是控制场景覆盖广泛，性能优化更明显。缺点是控制复杂，需要闭环算法进行计算。

在当前使用中，各手机平台对于天线 Tuner 的开环控制已经有了较好的支持，预留了相应的控制接口。但天线 Tuner 的闭环控制需要天线控制器和平台做深入的联动设计，在应用中还有很大的挑战。

图5.96为采用闭环控制算法实现的天线 Tuner 调谐效果 [7]，可以看到，不使用天线调谐的情况下，手机输出功率在 25dBm 到 31.5dBm 进行波动。而使用了闭环控制的天线调谐器后，功率的波动范围缩小到 29dBm 至 31dBm 之间。天线性能得到了明显的稳定，多种场景下发射信号得到了明显的增强。

匹配调谐：保证通路始终匹配

与天线匹配调节的思路类似，还可以在不同电路模块之间加入匹配调谐网络，实现可以随状态配置的匹配网络。如图5.97所示。

慧智微公司所开发的 AgiPAM® 系列射频前端系统中，就采用了阻抗匹配调谐网络技术，通过匹配网络的调谐配置，慧智微射频前端产品可以在较宽的频率范围内实现性能窄带优化，增强射频性能。慧智微可重构射频

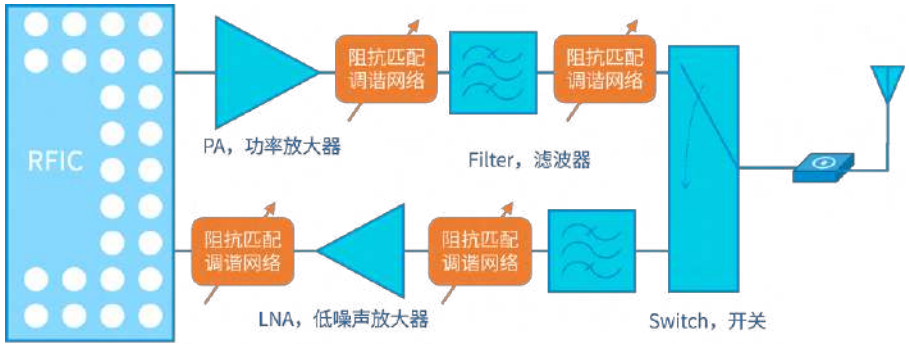


图 5.97: 包含阻抗匹配调谐网络的射频前端系统

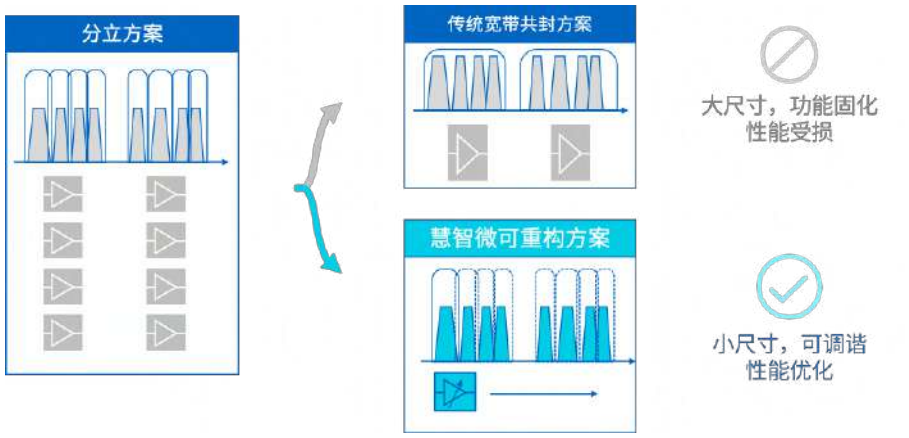


图 5.98: 慧智微可重构射频前端设计理念

前端设计理念如图5.98所示。

慧智微自 2011 年投入进行射频前端可重构技术平台的研究，并于 2014 年开发出世界首款商用可重构射频前端技术平台 AgiPAM®，经过近 10 年的迭代发展，目前可重构技术平台已发展至第四代，实现包含 4G/5G 产品在内的数亿片产品出货。

在具体产品实现上，基于 AgiPAM® 3.0 技术平台实现的 5G Phase5N 多频多模 PA 模组产品 S55643，通过阻抗匹配调谐和 PA 性能调谐，可实现 n7、n40（PC2）、n41（PC2）等高频关键频段优化。实测结果显示，S55643 系列产品可实现电流 13 ~28% 的降低。性能对比如图5.99所示。

与天线调谐类似，阻抗匹配调谐也需要根据应用场景进行配置。慧智微 AgiPAM® 平台目前采用开环的方式对阻抗匹配调谐进行控制。如图:ssrgert5uj176 所示，进入 AgiPAM® 5.0 后，慧智微期待与平台及终端客户深入合作，实现与应用场景结合更紧密的射频性能优化，实现手机射频性能的继续增强。

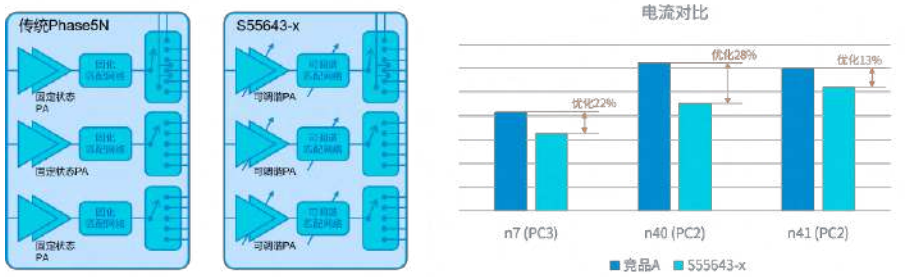


图 5.99: 传统 Phase5N 产品与基于可重构射频前端技术产品 S55643-x 对比

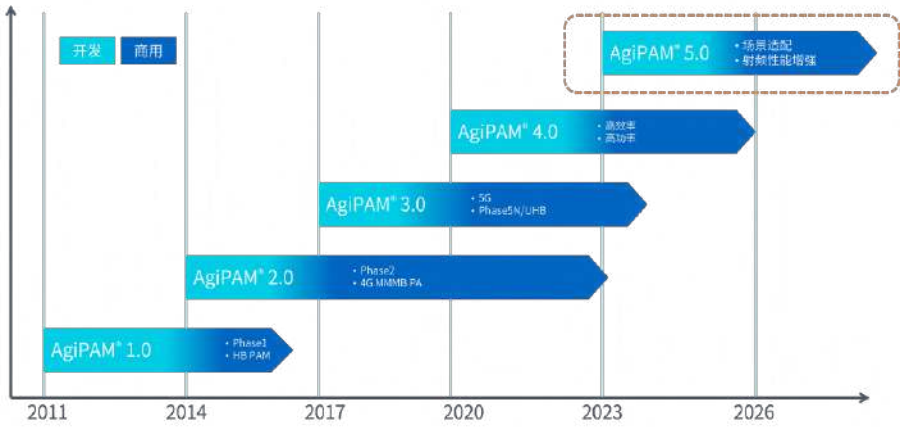


图 5.100: AgiPAM® 系列产品路线图

5.5.3 总结

近年来，因为手机年出货量的饱和，以及芯片供应商的增加，“卷”字成为手机行业的主题。随之而来的就是成本导向，价格雪崩，整个行业陷入不健康竞争。

用户不愿意为高价产品买单的原因，还是因为产品无法提供解决差异化的价值。如果用户的痛点得以解决，产品是可以卖出溢价的。

具体到射频前端芯片行业，可能的研究方向有：

- 用户关心的信号问题，可以通过优化射频前端性能、干扰，甚至采用新的天线方案、系统架构来进一步增强；
- 用户关心的功耗问题，可以通过高效率的 PA 内核，与平台 DPD 的配合等技术来优化；
- 用户关心的手机厚度问题，可以通过自屏蔽技术，减少额外屏蔽罩的应用来解决。

以上问题解决起来并不容易，需要大量投入与长期积累，但却是形成差异化价值的必经之路。慧智微期待利用技术的演进与优化，使手机射频性能不断增强。

您在手机射频增强方面有什么思考，欢迎扫描文章末尾二维码，在线留言讨论。

参考文献

1. 荣耀发布业界首颗射频增强芯片 C1, http://tech.china.com.cn/phone/20230307/394582.shtml, 2023
2. 3rd Generation Partnership Project, Technical Specification Group Radio Access Network, NR, User Equipment (UE) radio transmission and reception，3GPP, 38.101
3. Abhinay Kuchikulla, Aperture Tuning: An Essential Technology in 5G Smartphones, 2020, Qorvo
4. 5G 移动终端天线设计，林辉等，中国工信出版社，2021，
5. Huawei launches the Huawei Mate 30 Pro, https://drscmedia.eu/5475, 2019
6. https://www.qorvo.com/design-hub/blog/mobile-5g-device-antenna-tuning-demystified
7. RF SOI for Front-End Tunability Components, Dr. YuminLu, SITRI

在线阅读

本文摘自慧智微微信公众号，扫描以下二维码，在线阅读

5.6 如何评价一部手机的通信能力？

作者：彭洋洋 博士

自 2007 年 iPhone 初代发布以来，智能手机快速发展，屏幕、拍照、运算能力均得到大幅度提高，现在已经成为我们生活中不可缺少的移动智能中枢。

不过在过往 16 年的发展中，智能手机作为一个无线通信最有代表性的载体，其“无线通信”能力反而像是一个被忽略的技术，并不被用户熟悉和感知。即使有 2019 年 5G 商用的宣传加持，但完成 4G 至 5G 的切换后，“无线通信”继续从台前走到幕后，默默在背后提供着支持。

不过这一情况最近正在变化，越来越多的手机厂商关注到手机通信能力优化带来的收益，越来越多的用户寻求连接体验的提升。在材料《手机“射频增强”的技术方案》中，我们讨论了有关“射频增强”可能的技术方案。

在材料发出后，一些运营商、手机/芯片厂商以及手机用户问到：这只是一些改进方案，**那有没有一些方法，可以对手机通信能力做一个评价呢？**

我们觉得这是一个非常好的问题，于是我们就把关系到手机通信能力的技术要点做一个总结，与大家讨论“如何评价一部手机的通信能力”。

5.6.1 引言 · 手机通信能力

手机的通信能力是指手机在移动网络中无线信号的接收和发送的能力，通信能力的高低会影响手机的上网速率、通信质量、网络覆盖甚至手机功耗等性能，是一部手机里重要的能力指标。

在手机参数介绍中，都会明确的标识手机所支持的频段和通信制式支持。如图5.102为 iPhone 14 Pro 5G 手机中所标识相关信息 [1]。虽然手机都会将通信的基本信息标注出来，但即使对从业者来说，这些信息对了解手机通信能力都太过寥寥，很难判断出手机通信能力的强弱。

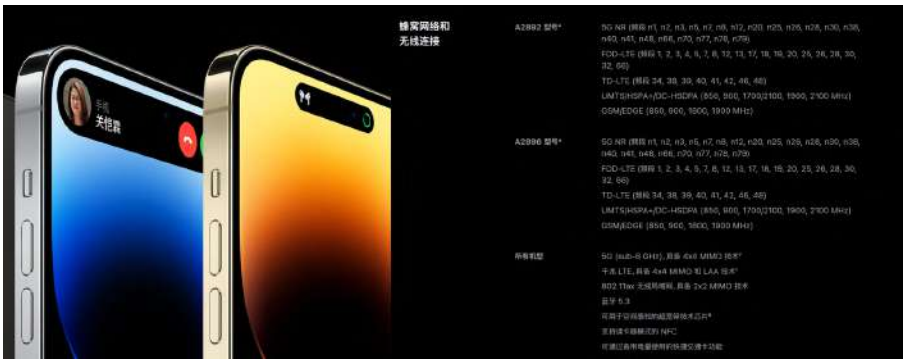


图 5.102: iPhone 14 Pro 5G 手机支持的通信频段和制式

但可以继续深挖一下，一般来说，手机通信能力主要与以下几个技术和优化有关，分别是：

- 通信技术能力
- 射频技术能力
- 实际优化能力

接下来，我们就从以上三方面讨论如何评价一部手机的通信能力。

5.6.2 通信技术的评价

“通信技术”是电子工程里重要分支，是现代信息传输和信号处理的基础，可以说没有通信技术，就没有现在的电话、互联网、卫星等有线/无线通信网络。通信技术是通信行业原理性的技术，在此我们只从实用的角度，对影响到手机通信能力的相关技术做一个讨论。

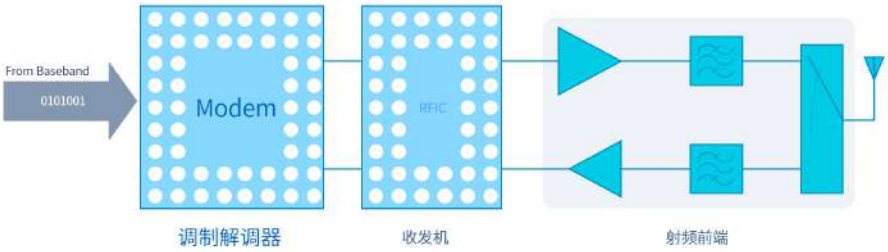


图 5.103: 调制解调器在手机系统中的位置

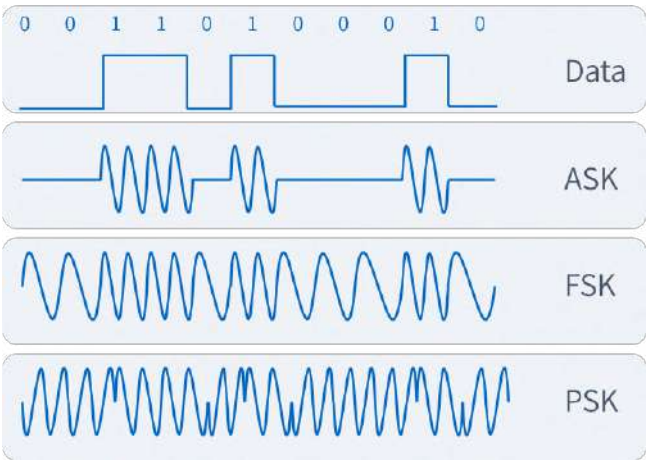


图 5.104: ASK、FSK、PSK 调制方式

在通信技术中，影响到手机通信能力的主要有高阶调制能力、香农定理极限。

5.6.2.1 高阶调制能力

调制（Modulation）是一种信号变换的过程，调制通过改变数据的频率、相位、幅度等特性，使信号更好的在信道上传输。

实现调制功能的电路是调制解调器，也就是大家常说的 Modem（Modulation-Demodulation）。在 2000 年左右，不少家庭利用 Modem 实现了电话线上网功能，大家也从这个时候开始听说“猫”可以用来上网。其实在当前的手机中，都集成了 Modem 功能，实现信号转换发射。图5.103为调制解调器在手机系统中的位置。

在通信中，必须要将数据进行调制的原因是：

- 如果直接传播 0、1 的数字信号，相当于是在传播方波，这将占据大量带宽，所以需要将数字信号调制成高频模拟信号（也就是射频信号），来传输
- 将数据调制到不同的载波，也有利于信道的充分利用
- 利用高阶调制，可以实现更有效的数据传输

在调制中，最简单的调制方式就是调幅或者调相等直接调制，这些方式主要有：

- 移幅键控（ASK，Amplitude Shift Keying）
- 移频键控（FSK，Frequency Shift Keying）
- 移相键控（PSK，Phase Shift Keying）

三种调制的工作方式如图5.104所示。

在以上调制中，每个调制符号仅带有单个比特的信息，要么是“0”，要么是“1”。为了提升信号的传输效率，更高阶的调制方式被发明了出来，这就是经常被提到的 QPSK、16QAM、256QAM，甚至 1024QAM。图5.105为 QAM 星座图。

高阶调制的理念是把多个比特放在直角坐标系中，对于 16QAM 来说，总共有 16 个坐标点，每个点就代表

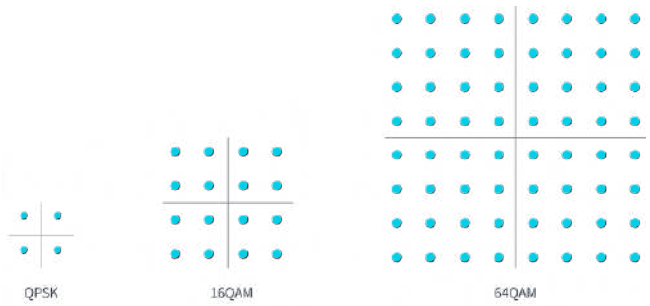


图 5.105: QAM 星座图

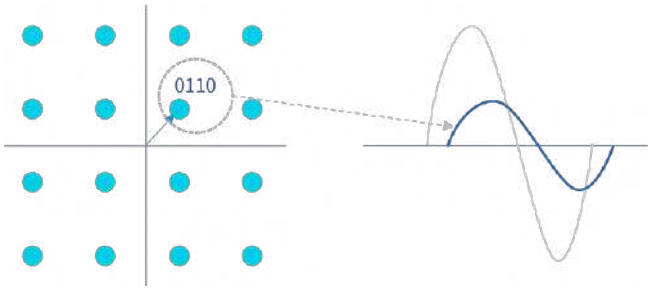


图 5.106: 16QAM 调制下的信息传输

0000-1111 这 16 个信息中的一个。对于选定要传输的信息，只需要挑选对应坐标点的幅度和相位进行传输就可以，这样可以将传输效率大大提升。对于 16QAM 信号，一个调制符号位携带了 4 个比特的信息，所以 16QAM 调制的信道利用率就是 4。图5.106为 16QAM 调制下的信息传输。

同理，如果是利用 1024QAM 调制进行数据传输，单个符号位就可以携带 10 比特信息（$2^{10} = 1,024$），信道利用率提升到了 10，传输效率再次加强。1024QAM 信号传一次，等于原来的 FSK 方式传 10 次，这就是数字通信不断追求高阶调制的原因。

由于调制方式是决定通信速率的重要指标，在手机中一般会对此项能力做严格规定。在 2019 年版本的 4G 手机白皮书中，上行调制方式仅需要支持 64QAM[2]。而在中国移动 5G 手机白皮书中，这一能力做了提升，要求 5G 手机必须支持上行及下行的 256QAM 调制方式 [3]。手机在调制方式上的能力做了增强，通信能力也相应提升。图5.107为中国移动 5G 手机白皮书中对 5G 手机调制方式做的规定。

在手机设计中，Modem 一般集成于 SoC 芯片中，所以一旦 SoC 芯片选定，最大的高阶调制能力也确定下来。如图5.108为联发科推出的天玑 7200 5G 手机 SoC 芯片，其集成的 R16 5G Modem 可支持 256QAM 的高阶调制。

SoC 芯片将最高阶调制实现后，还需要射频前端通路完成高阶调制信号的不失真传输。慧智微 5G 射频前端系列产品均经过高阶调制信号严格测试，确保信号高阶调制传输不失真。

5.6.2.2 “香农定理”决定极限

既然高阶调制可以如此有效的提升传输效率，那是否可以无限的增强调制的 QAM 数，来使通信能力得到无限增强呢？答案是否定的，这个限制的来源就是“香农定理”。

香农（Claude Elwood Shannon，1916-2001）是信息论的奠基人，曾就职于贝尔实验室。1948 年，32 岁的香农在其著名的论文《通信的数学理论》中，提出了如下经典公式：

	下行	上行
n41	256QAM	256QAM
n79	256QAM	256QAM
n28	256QAM	256QAM

图 5.107: 中国移动 5G 手机白皮书中对 5G 手机调制方式做的规定



图 5.108: 联发科天玑 7200 5G 手机 SoC 芯片

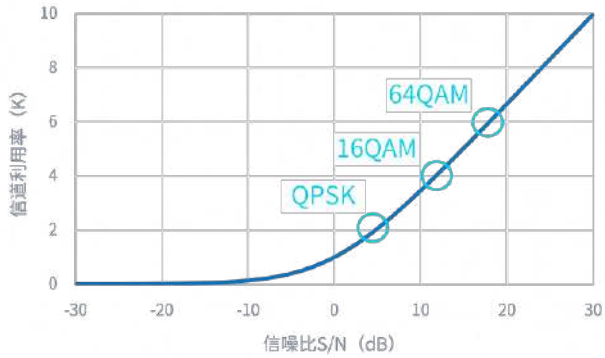


图 5.109: 信道利用率与 S/N 之间的关系

$$C = B\log_2\left(1+\frac{S}{N}\right)$$

其中，C 代表可传送的最大信息速率，又称信道容量；B 代表信道带宽；S/N 代表信号与噪声的能量比值。这就是大名鼎鼎的香农公式，也被认为是通信行业最著名的公式。

虽然数学公式总是让人望而生畏，但香农公式却可以用很简单的直觉来理解：**通信能力和信号带宽呈正比，和信号质量呈正比**。以上理解已经可以帮我们解释日常生活中的大多数问题，比如为何要用更大的带宽来进行信号传输，为什么要追求更强的信号质量。这些在香农公式看来，都是为增强通信能力所做的必要努力。

香农公式的力量绝不只是帮我们理解通信这么简单，香农公式还可以帮我们计算出给定信道下的通信速率极限，以帮我们确认如高阶调制等技术努力的极限在哪里。比如将香农公式略做变形，将信道带宽 B 移到等式左边，就可以得到信道利用率 K（定义为 C/B），与信噪比之间的关系：

$$K = \log_2\left(1+\frac{S}{N}\right)$$

对于 16QAM 信号来说，如前节所述，其信道利用率是 4，代入上式可以求出，S/N=15，即 12dB。所以对于 16QAM 的信号传输来说，理论上必须要达到 12dB 以上的信噪比，才可以进行有保障的传输。

将信道利用率 K 与信噪比的关系在同一图表上表示如图5.109。

通过上图可以看到，高阶信号的传输对信号质量也提出了要求。这就说明手机信号强度对通信速率的重要性。在手机信号强的区域，信号信噪比高，手机就可以用高阶调制进行数据传输；而当手机信号变弱，就必须要回退到低阶调制来保障传输可靠，通信速率自然下降。

所以对一部手机来说，要增强通信能力，就需要保障信号的质量，以使手机尽可能的工作在高调制阶数区域。

需要说明的是，上图只是理论计算的理想情况，在实际应用中，由于信道编码的效率问题、冗余校验比特需

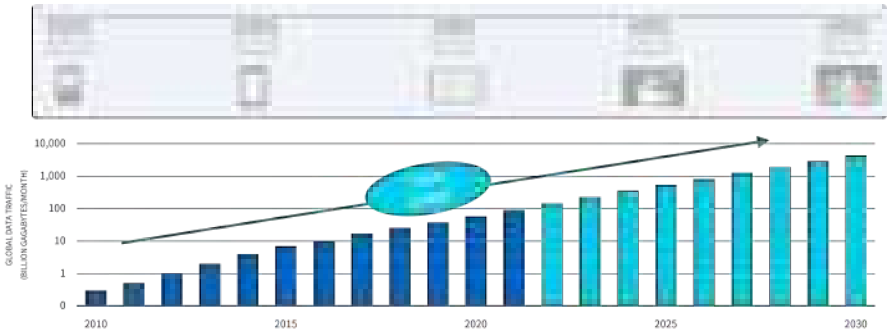


图 5.110: 全球移动通信数据

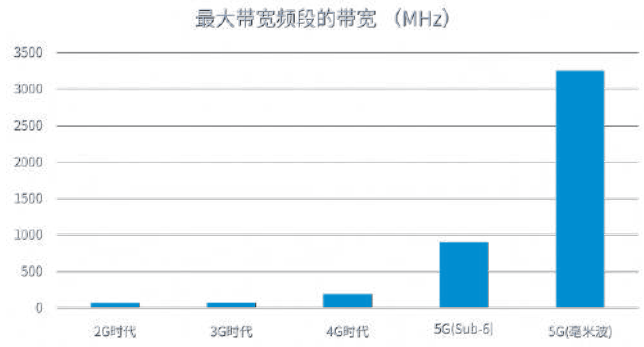


图 5.111: 最大带宽频段的带宽演进

求等，传输效率会低于理论计算。但通信工程师们也在不断努力，通过如 Turbo 编码等更高效的信道编码方式不断接近理论极限的信道利用率 [6]。

5.6.3 射频技术的评价

通过过去 80 多年的研究，通信技术在不断的逼近香农定理的极限，信道的利用率已接近上限。

但人类对通信能力的要求是永无止境的，根据国际电信联盟 ITU 的统计和预测，如图5.110，在 2010 至 2020 的 10 年中，全球月移动通信数据由 5 亿 GB 每月，提升至 700 亿 GB 每月，数据增加 100 倍。并且在至 2030 年的下一个 10 年中，还将再提升 100 倍，至近十万 GB 每月 [7]。

香农定理逼近极限，无线需求能力持续增加，这给手机工程师们提出了难题。目前在手机中，再次将通信能力提升的方法主要是依靠射频能力的增强。主要用到的技术为：

- 更大带宽的通信频段
- 载波聚合（CA）
- 多入多出（MIMO）

手机的通信能力，也和以上技术的支持程度息息相关。

5.6.3.1 更大带宽的通信频段

根据香农定理，信道的通信容量与带宽呈正比。所以就意味着找到更大带宽的通信频段，就可以实现更大速率的无线通信。

自 2G 以来，移动通信就致力于寻找更大带宽的通信信道进行信号传输。在低频段频率资源受限情况下，无线通信也向更高频率发展。目前在 5G Sub-6GHz 频段，n77 和 n79 频段分别可以提供最大 900MHz 及 600MHz 频率资源，到了毫米波频段，更是可以提供 3GHz 以上频率资源，这是 4G/3G 时代是不可以想象的。如图5.111所示。



图 5.112: iPhone 14 Pro 手机支持的 5G 频段列表

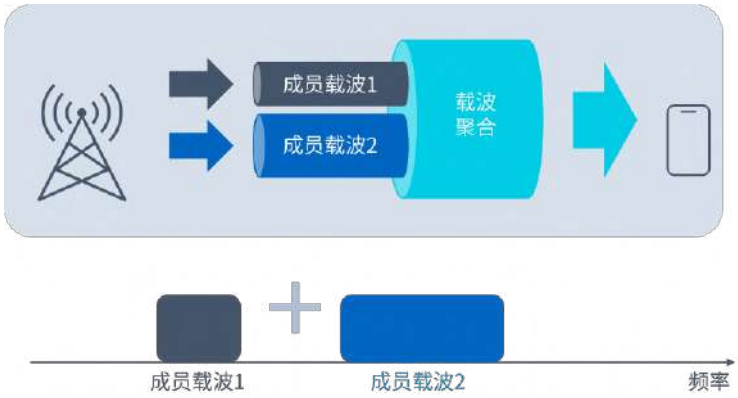


图 5.113: 载波聚合的支持

对于手机来说，可以支持更多的频段，也意味着有更强的通信能力。在一般的 5G 手机中，都可以对中国移动、联通、电信要求的 n1/3/5/8/28/41/78 频段进行支持，部分高端手机还可以对大带宽的 n79 进行支持。

手机会在其规格参数中明确显示支持的频段，如图5.112为 iPhone 14 Pro 手机支持的频段列表，手机支持大部分的国内及国际 5G 频段，并且支持 n79 频段 [1]。

手机对新频段的支持不仅需要 SoC、Transceiver 芯片的支持，还需要射频前端芯片，如 PA、滤波器、开关等芯片的配套支持。

5.6.3.2 载波聚合能力

虽然增大带宽可以有效提升通信速率，但频谱资源对于任何一个国家来说都是稀缺资源。怎么利用有限的频谱资源，取得更大速率的传输成为手机通信行业努力的方向。

于是，载波聚合（CA，Carrier Aggregation）就被提了出来。载波聚合的理念是将多个窄带的成员载波（Component Carrier，CC），聚合在一起，多个信道同时传输，单车道变多车道，通信能力自然增强。如图5.113所示。

载波聚合的支持需要在射频通路建立同时打开的通道，并且需要解决好两个频段间的隔离问题，对射频前端设计提出了更多的挑战，并且会增加射频前端的成本。所以对载波聚合功能的支持，成为不同价位带手机重要的差异化功能。手机对 CA 功能的支持如图5.114所示。

以高通平台为例，以欧洲版本手机方案为例，对于中阶 4 系列平台，其支持的 CA 组合为 63 个，而对于其高阶 6 系列平台，支持的 CA 组合数量增加至 112 个。如图5.115所示。

同新频段的支持一样，载波聚合功能的支持需要 SoC、Transceiver 与射频前端芯片的共同协作才能支持。

手机价位带	CA支持
入门级	不支持
中阶	带内上行/下行CA，及带间简单下行CA
高阶	复杂上行/下行CA

图 5.114: 手机对 CA 功能的支持

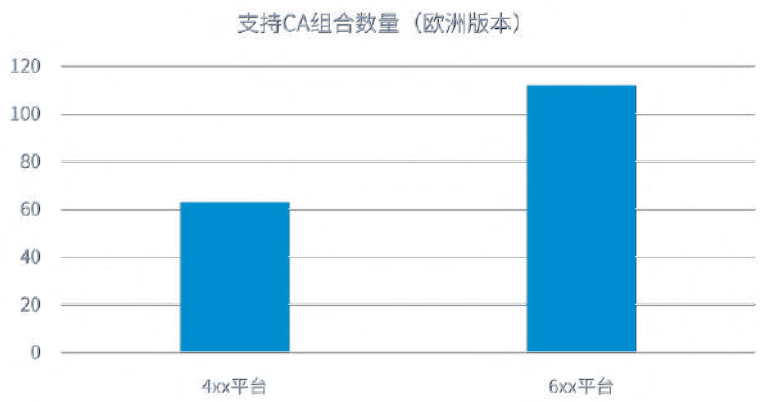


图 5.115: 不同价位平台支持的 CA 组合数量

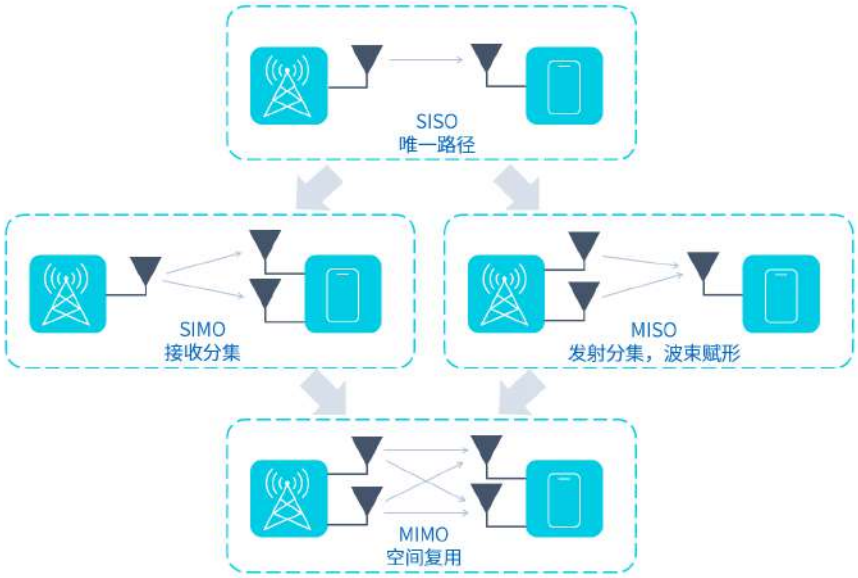


图 5.116: 从 SISO，到 MIMO

5.6.3.3 多入多出（MIMO）能力

MIMO 技术的全称是 Multi-input Multi-output，多输入多输出，这是利用射频提升通信能力的又一重要技术。MIMO 技术提升通信能力的原理主要有两个，分别是：分集提升可靠性；复用提升容量。

其实在从单入单出（SISO，Single-input Single-output）到 MIMO 的演变过程中，还有 SIMO 和 MISO 两个演变，这四种结构的关系如图5.116所示。

在 SISO 系统中，收发均为单一路径，传输 1 路信号。在这种系统中，信号传输不可靠，在受到干扰时很容易影响传输速率。

为了改变这一局面，可以在手机处再增加一个天线，这就是分集天线，这样可以使接收端同时接收到两路信号，实现单发多收 SIMO。单发多收系统可以提升信号的可靠性和稳定性。但由于接收到的信号是从同一个发射天线发射出来的，数据传输相同，所以分集接收并不能增加传输通道数目，也就不能使通信速率翻倍。

同样的道理，可以实现多发单收 MISO。由于接收天线只有一个，MISO 系统同样不能增加传输通道数目，但 MISO 系统可以用于设计发射端的波束赋形系统，利用发射信号间的相位相加，实现基站发射信号向特定方向的叠加增强。

如果将发射与接收信号同时增加，实现 MIMO 系统，就可以实现 2 路信号的同时收发，增加通道数目，使通信速率翻倍。这也就是 4G/5G 中的空间复用技术。对于收发同为 2 根天线的系统，也称为 2x2 MIMO 系统。

频段	发射	接收
n41/n79	必选1T，推荐2T	必选4R
n28	必选1T	必选2R，推荐4R

图 5.117: 中国移动 5G 手机要求

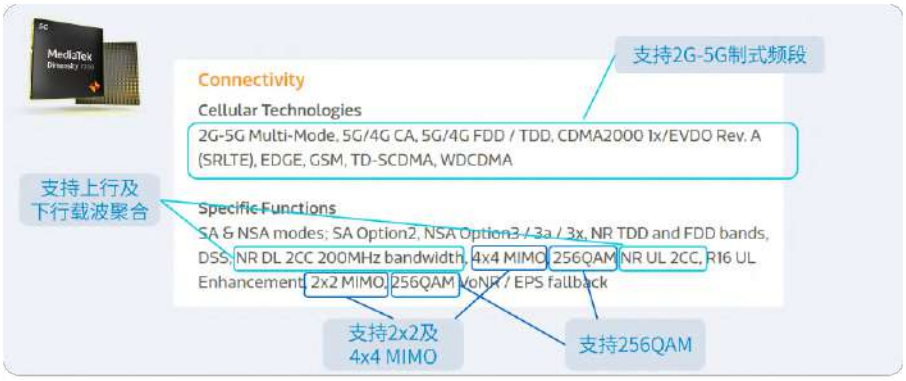


图 5.118: 根据联发科天玑 7200 规格看到的通信能力支持

为了支持 MIMO 系统，标准化组织或运营商会要求终端发射及接收通道数目。在规定中，发射一般用 T 表示（缩写自 Transmit），接收一般用 R 表示（缩写自 Receive）。比如对于中国移动白皮书，对于 5G 手机收发通路数目有如图5.117要求。

支持的 MIMO 的能力也是不同价位手机差异所在。比如对于 iPhone 14 Pro 手机，n41/79 频段均支持 2T4R MIMO。而对于入门级 5G 手机，对于 n41 频段一般只支持 1T4R MIMO，在发射 MIMO 能力上做了简化。

MIMO 功能的支持，需要射频前端与 SoC、Transceiver 配合支持。

5.6.4 性能优化的能力

通过以上技术分析，可以对一部手机的基本通信能力有一个大致了解。比如观察联发科天玑 7200 SoC，就可以看到如图5.118信息。

不过以上评估仅表示手机有无对应的通信“能力”，但要达到用户真正有感知的通信能力提升，还需要在手机设计上多加考虑。例如需要考虑以下几点的性能优化。

5.6.4.1 天线的优化设计能力

与笔记本电脑等其他移动设备不同，手机极其强调便携性，这给手机的天线设计带来极大挑战。并且手机应用环境有着强的不确定性，手机需要在不同持握状态下，均有好的通信信号。

不好的天线设计会让前述通信能力的引入前功尽弃，如图5.119, 苹果 iPhone 4 就经历过所谓的“天线门”，手机在特定握持状态下，会造成天线的短路，进而影响手机通信，给用户使用造成很大不便。



图 5.119: iPhone 4 时代所经历的“天线门”事件

5.6.4.2 射频性能的“增强”能力

在《手机“射频增强”的技术方案》中，我们讨论了一些可能的增强方法，这些方法包括：

- 发射通路功率的增强
- 接收通路干扰的避免与消除
- 结合场景的匹配优化

这些增强的方法虽然不像第一、二节讨论那样，可以从通信原理的本质上提升通信能力，但却可以明显的改善用户的使用体验。这一点也是拉开不同手机使用差距的技术要点。

虽然不同手机使用的芯片可能相同，但在射频性能调试上、结合场景优化上做的不同，会带来用户使用感受的明显不同。

5.6.4.3 其他考虑

手机通信能力的优化不限于以上提到的部分，任何有助于用户可以更快速的刷视频、更少的掉线的技术，都可以用于手机通信能力的提升。比如还可以利用软件与硬件的结合，更好的做好场景优化，甚至可能通过手机与基站的协同，完成手机体验的提升。这些都是不同厂商可以利用自身优势，提升通信能力的考虑。

另外，手机可能是一天中陪伴用户时间最长的智能设备，通信能力的提升还必须要结合功耗、芯片尺寸、成本等多个因素进行优化，带着镣铐跳舞，在众多限制中找到最优解。

5.6.5 总结

自 1973 年大哥大的诞生以来，手机已经快速发展了 50 年。在这 50 年中，人们对手机的要求也越来越高。

通信能力是手机中的重要能力，但因为射频技术复杂，场景众多，较难量化，所以也是最难评价的能力之一。

有关手机射频通信能力的评价，可能可以以下几点作为参考：

- 硬件能力评价：评价手机理论硬件支持能力
- 硬件性能测试：测试手机硬件实际性能
- 实际场景评价：建立仿用户使用场景范例，评价实际使用优化能力

让射频能力可见不容易，您对手机通信能力评价有什么考虑？欢迎扫码参与线上留言讨论。

参考文献

1. https://www.apple.com.cn/iphone-14-pro/specs/
2. 中国移动 4G 手机产品白皮书（2019 年版），中国移动，2019 年 6 月
3. 中国移动 5G 手机产品白皮书（2023 年版），中国移动，2022 年 12 月
4. MediaTek Dimensity 7200，https : //i. mediatek .com/ dimensity -7200
5. C. E. Shannon， A mathematical theory of communication, The Bell System Technical Journal, vol. 27, 1948
6. R1-167414, Enhanced Turbo Codes for NR: Performance Evaluation
7. Global mobile data traffic forecast by ITU，https:/ /www. ericsson .com/en/ reports-and-papers /mobility-report /dataforecasts/ mobile- traf fic-forecast
8. 5G 移动终端天线设计，林辉等，中国工信出版社，2021，

在线阅读

本文摘自慧智微微信公众号，扫描以下二维码，在线阅读

5.7 今年，是时候了解“毫米波相控阵”了

作者：彭洋洋 博士

2023 年伊始，我们做了 2023 年大家最想了解的射频技术话题征集（见文章 2023 年，射频人关注这些技术问题），出乎意料的是，毫米波相关技术获得了 21 个关注话题，排在最受关注话题的第四位。看似高高在上的毫米波技术，已经成为射频人重点关注的方向。

毫米波技术在军用、雷达等领域已经有多年的应用。在民用领域，也随着最近的 5G 移动通信、民用卫星通信，以及车载毫米波雷达等应用的普及，逐渐走进了大众的视野。

我国工信部近日在 2023 年 1 月发文，将 21.2-23.6GHz 和 71-76GHz/81-86GHz 的毫米波频段，列为我国可用于无线通信的频段 [1]。根据统计显示，5G 毫米波手机 2023 年出货将突破 1 亿部，并且在 2025 年有望实现第二波的快速增长 [2]，5G 毫米波手机年出货量如图5.121所示。

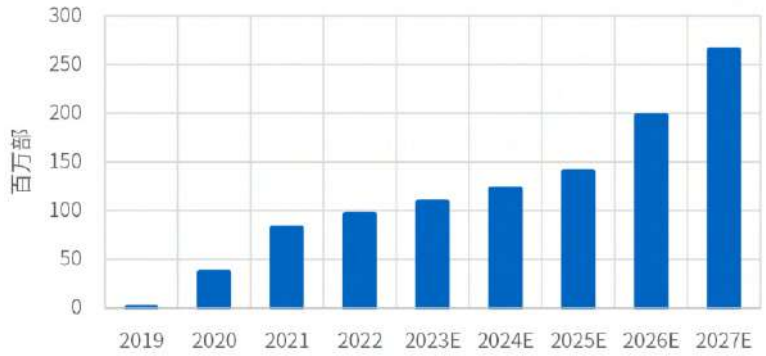


图 5.121: 5G 毫米波手机年出货量

除手机外，其他领域的毫米波应用数量也在快速提升。图5.122分别为车载毫米波雷达市场数据，以及全球卫星发射数量 [3][4]。可以看到二者在近几年都在快速增长。

本文就尝试对毫米波系统中最常用的系统结构：毫米波相控阵结构（Millimeter-wave Phased-Array），做一个讨论。探讨略显神秘的毫米波系统。

5.7.1 什么是毫米波？

无线通信是基于电磁波所进行的通信技术。为了使不同的通信设备传输互不干扰，国际电信联盟等无线电管理机构对无线频谱的使用做了划分，将不同频率的频谱资源，定义到不同的应用中。



图 5.122: （a）车载毫米波雷达市场，（b）全球卫星发射数量数据

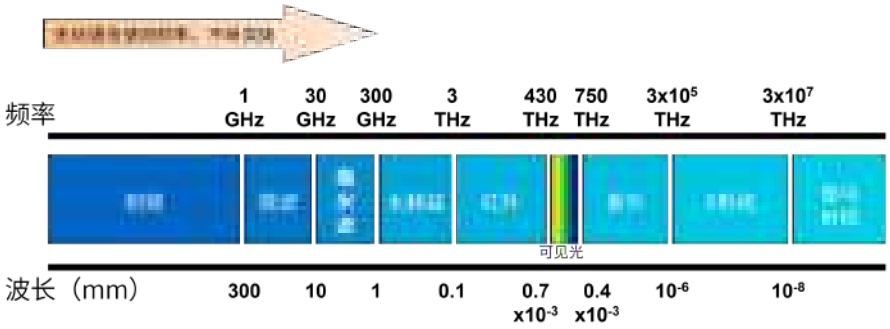


图 5.123: 毫米波在电磁波频谱中所处位置



图 5.124: 4G、5G Sub-6G 以及 5G 毫米波下载速率对比

毫米波一般是指电磁波频率近似在 30GHz 到 300GHz 频段范围内的电磁波，由于此频段电磁波在真空中的波长大约在 10mm ~1mm 之间，波长处于“毫米”量级，所以这个频段的电磁波被称为毫米波。如图5.123所示。

受益于半导体集成电路工艺、通信设备技术的突破，人类对电磁波频谱资源的征服是不断向上延伸的。比如在民用通信领域：

1. 在 20 世纪初年代，主要的无线通信制式是电视和电台广播，所使用的是频率范围在 100MHz 左右的射频频率；
2. 进入 20 世纪 80 年代，人类开始使用约在 1GHz ~3GHz 范围的微波频段，实现手机移动通信；
3. 2020 年，5G 移动通信除了定义 6GHz 以下频段外，还将频率扩展至 24GHz ~40GHz 的毫米波频段。

5.7.1.1 毫米波通信的特点

特点一：大带宽

人类将应用频谱不断向上扩展的源动力，是寻找更丰富的频谱资源，以满足更高通信速率的需求。

无线通信进入毫米波也不例外。相比于 6GHz 以下通信频段，30GHz ~300GHz 的毫米波有着近 50 倍的频谱资源。这就相当于在拥挤的车道旁边，又开辟了一个几十车道的高速公路，大大提升了通信速度。所以毫米波通信的第一个特点就是：大带宽。

大带宽可以完成更高的通信速率。如图5.124根据 Ookla SPEEDTEST 提供的通信速率显示 [5]，相比于 4G LTE，5G Sub-6GHz 网络可提供 5 倍的速率提升，而 5G 毫米波网络，可实现 20 倍速率的明显提升。

特点二：高分辨率

电磁波还可以用来作为雷达探测使用，通过发出电磁波信号，并且监测电磁波遇到物体之后的反射情况，就可以检测出物体的尺寸、距离等信息。这就是雷达探测的原理。

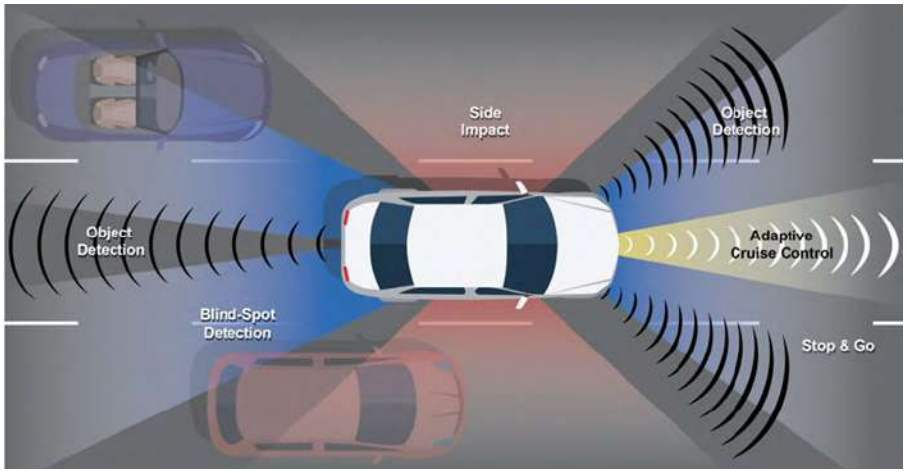


图 5.125: 毫米波雷达在智能汽车中的应用 [6]

作为雷达探测使用时，由于电磁波的衍射效应，电磁波对探测物体的分辨率和电磁波的波长呈正比：波长越短的电磁波，越能分辨出更精细的物体。于是，毫米波就被应用到雷达检测中来。

相比于 1GHz 左右，波长在 0.3 米左右的射频电磁波来说，位于 30GHz 以上的毫米波分辨率更高。车载毫米波雷达是毫米波在雷达领域的典型应用，车载毫米波雷达一般采用 24GHz、77GHz 以及 79GHz 频段，实现最高厘米级的高精度探测。毫米波雷达在智能汽车中的应用如图5.125所示。

特点三：损耗大，易受干扰

毫米波通信也有缺点，就是路径损耗大，易收到干扰。

根据 Friis 信号传输公式，在传输距离一定时，电磁波的损耗与波长尺寸呈正比：波长越短的电磁波，路径损耗越大。

路径损耗过大就使得毫米波通信无法传输足够远的距离。例如，对于 1GHz 移动通信，通信基站的覆盖范围可达到数公里范围。而对于毫米波，覆盖范围就快速缩小至数百米。这就对基站的部署提出了更高的要求。

除了路径损耗外，毫米波还容易受到物体遮挡的干扰。毫米波由于波长短，厘米尺寸的物体就会对信号形成遮挡和反射，这个特点在雷达检测中是优点，但在移动通信中却是致命缺点。造成毫米波只能用做“视距传输”，而无法进行绕射传输。如图5.126所示。

特点四：电路尺寸小

在射频微波电路的实现中，所用到的元器件值通常与电路工作的波长呈正比、频率呈反比。于是，工作在更高频率的毫米波电路通常可以做到更小的尺寸，这在一定程度上降低了电路成本，同时也为后续的相控阵技术提供了基础。

如图5.127所示，文献 [7] 中展示了工作于 24GHz 的 4 通道毫米波相控阵完整发射机系统，整个系统包含本振、上变频器、功率放大器等各个模块，并且包含 4 个通道数。如此复杂的通信系统在 2.1mm x 6.8 mm 的芯片下即可实现，只有一粒大米大小。

5.7.2 什么是相控阵？

相控阵（Phased Array）技术是控制阵列天线各单元的相位、幅度，来形成对信号空间波束控制的技术。

相控阵技术起源于 20 世纪初发明的相控阵天线技术，并最早在军用雷达技术中得到了广泛应用和迅速发展。进入 21 世纪后，随着民用电磁波频率的不断提高，相控阵技术在民用技术中也开始崭露头角。

在相控阵技术中，有两个重要的技术概念，分别是“相控”和“阵”。以下分别就这两个概念进行讨论。

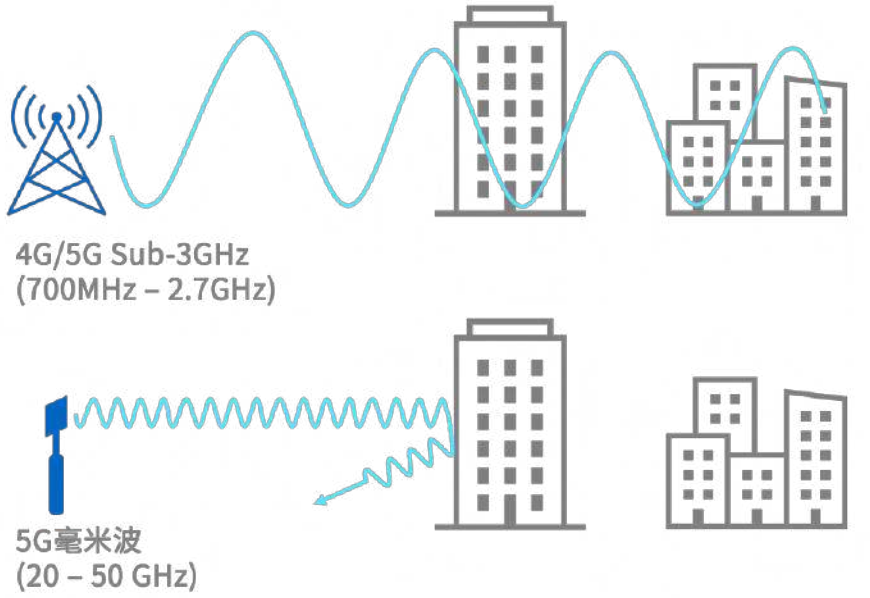


图 5.126: 毫米波传输，容易受到物体干扰

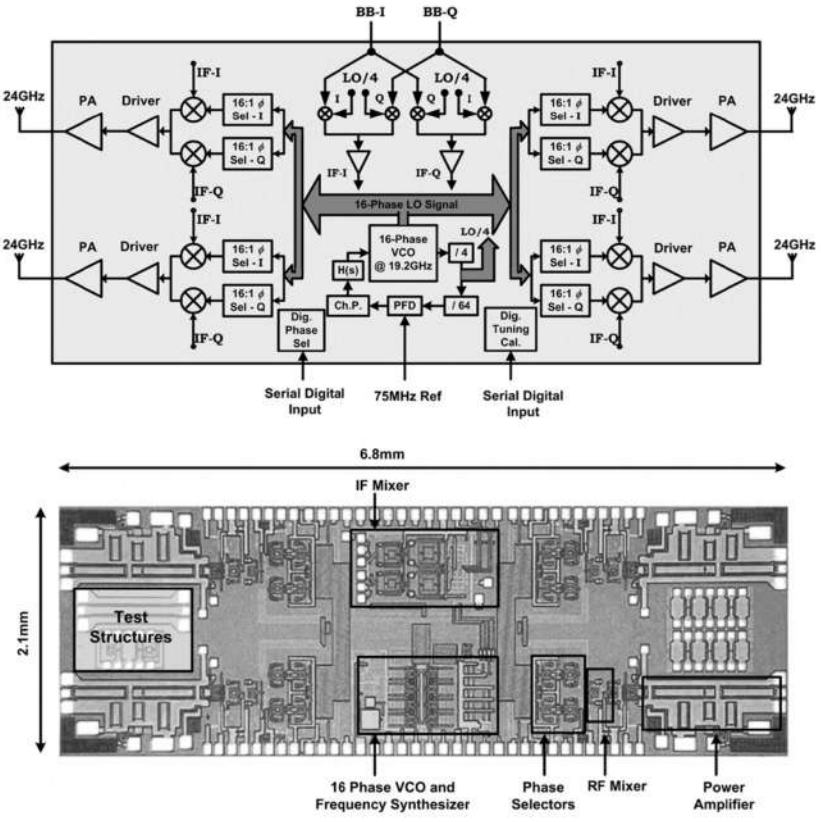


图 5.127: 4 通道 24GHz 毫米波系统

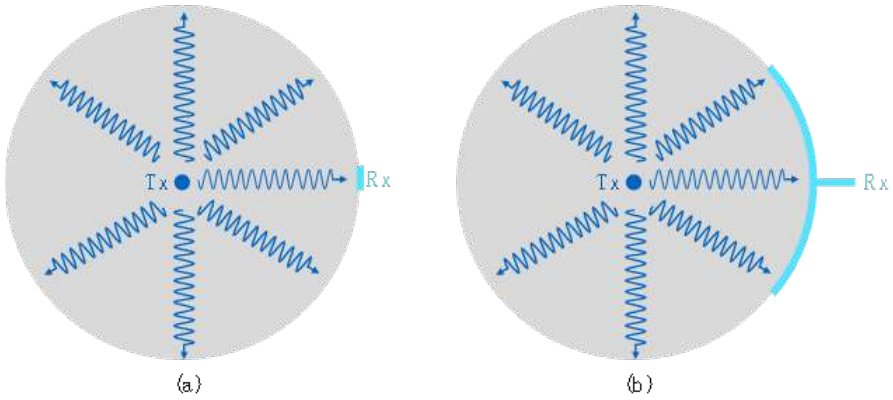


图 5.128: （a）全向辐射的电磁波，（b）增大天线面积来接收更多信号

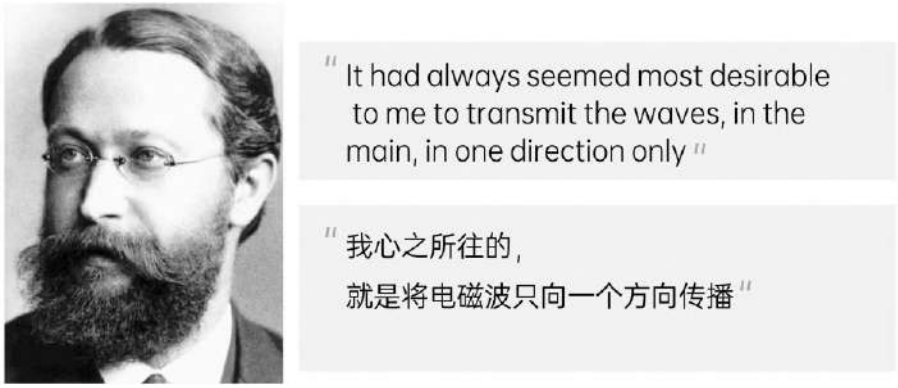


图 5.129: 卡尔 • 费迪南德 • 布劳恩，1905 年天线阵系统发明者

5.7.2.1 “阵”的引入：实现定向收发

在天线“阵”被发明之前，电磁波的辐射通常被认为是向外接近全向辐射的：发射信号能量以接近球面的方式向外扩散。根据能量守恒，发射距离越远，球面半径越大，单位面积得到的能量越小。当能量小到一定程度，接收机将无法接收到有用信号。这就是空间传输中“路径损耗”的主要来源之一。

当然可以用增大接收面积来接收更多的能量，很多球面天线就是采用这样的原理，但这样做的结果是球面面积大，并且球面始终需要对准发射源，不适用于发射、接收快速运用的场景。如图5.128。

于是，天线“阵”就被发明了出来。

天线阵是诺贝尔物理学奖获得者，著名物理学家卡尔 • 费迪南德 • 布劳恩（Karl Ferdinand Braun，1850 年 6 月 6 日－ 1918 年 4 月 20 日）于 1905 年所发明的。布劳恩是阴极射线管的发明者，同时也是无线通信技术的先驱者。1909 年，因为在无线电报技术中的贡献，布劳恩与马可尼分享了当年的诺贝尔物理学奖。

在获得诺贝尔奖时，布劳恩表示：“我心之所往的，就是将电磁波只向一个方向传播”[8]。只向一个方向传输的电磁波可以避免无谓的损耗，并且单方向的传输能量更强，传播距离也更远。如图5.129所示。

布劳恩设计的天线阵系统包含 3 根垂直单极天线，分别放置于等边三角形的三个顶点处，两两相距 1/4 波长。通过控制输入信号的相位，就可以实现三根天线发出的信号在三个方向上的叠加情况，从而实现天线向三个方向的分别定向发射。1905 年布劳恩发布的天线阵系统，及其远场辐射图如图5.130。

天线阵技术被发明后，受到了军方极大的关注。其定向发射接收、不需要物理转向调节、传播距离远等特性非常适用于军用雷达领域。于是在 1920 年左右，美国、德国等国家开始研究将天线阵应用于军事雷达中。在 1941 年，美方将天线阵雷达 SCR-270 系统部署于珍珠港 [9]，该系统包含由 32 根天线构成的天线阵列。虽然这个雷达系统并没有阻止住日本的攻击，但天线阵雷达的可行性得到了完整验证。在现代军用系统中，相控阵系统已经得到的广泛应用。如图5.131。

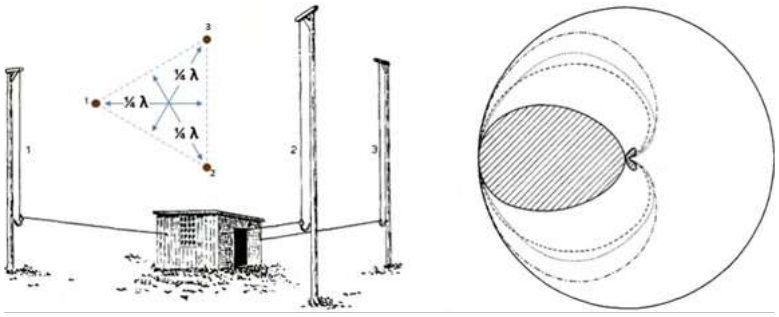


图 5.130: 1905 年布劳恩发布的天线阵系统，及其远场辐射图

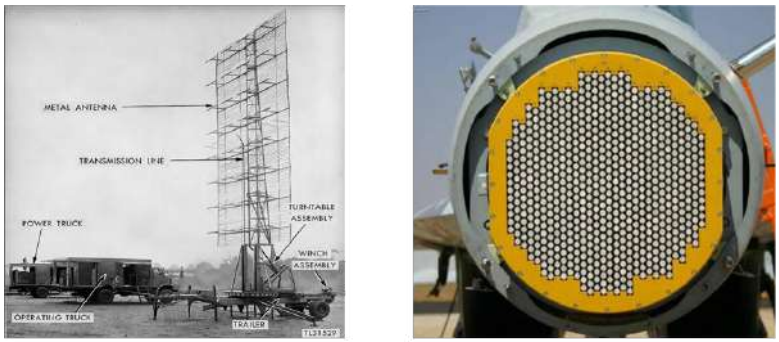


图 5.131: （a）美军 1941 年在珍珠港部署的 SCR-270 天线阵雷达系统（b）俄米格-35 战斗机装备的甲虫-AE 相控阵雷达系统

5.7.2.2 “相控”技术：控制瞄准方向

天线阵的引入为电磁波的定向收发提供基础，但实现方向的控制与扫描，还需要引入“相位控制”技术，也就是“相控”。

以接收信号时举例，但天线阵系统进行信号接收时，由于进入各天线的信号经过的传输路径不同，如果直接相加，并不能实现信号的完美加和。这个时候，就需要将各路信号进行移相对齐后，再叠加起来。这个移相对齐的过程，就称为“相控”。通过控制不同通路间的相位关系，就可以接收不同位置发出来的电磁波信号。接收通路中的相位控制示意如图5.132。

在发射信号时也是一样，通过对输入信号的相位设计，可以控制输出信号在哪个方向进行叠加。如此，如果需要变换发射角度时，只需要改变各信号的相位差。这样就建立起信号发射角度与相位之间的联系。

为简单描述，以两天线组织的阵列分析如图5.133所示，当两天线发出的信号之间相位无偏移时，两天线发出的信号在中间对称处叠加，而在其他位置抵消，信号集中于垂直方向发射；当两天线信号有相位差时，以天线 1 的相位延迟大于天线 2 为例，天线 2 发出的信号超前于天线 1，此时叠加方向向左倾斜。通过控制天线 1 与天线 2 之间的相位差，即对发射信号的波束方向进行控制。因为这种技术像是在对波束形状进行赋形，所以也被称为“波束赋形（Beam forming）”技术。

5.7.2.3 “移相”的实现

由于各信号的“相位”与信号的发射方向、叠加强度直接相关，所以“移相”功能是相控阵系统中非常重要的功能模块。在现代相控阵系统中，移相功能通常由移相器电路实现。

顾名思义，移相器就是实现信号相位变化的电路，通过信号延迟、信号叠加等方式，使输入信号产生相移，从而改变输入信号的相位。

一般在电路实现上，分为无源移相和有源移相两种。两种移相方式常见的电路结构与特点如图5.134。

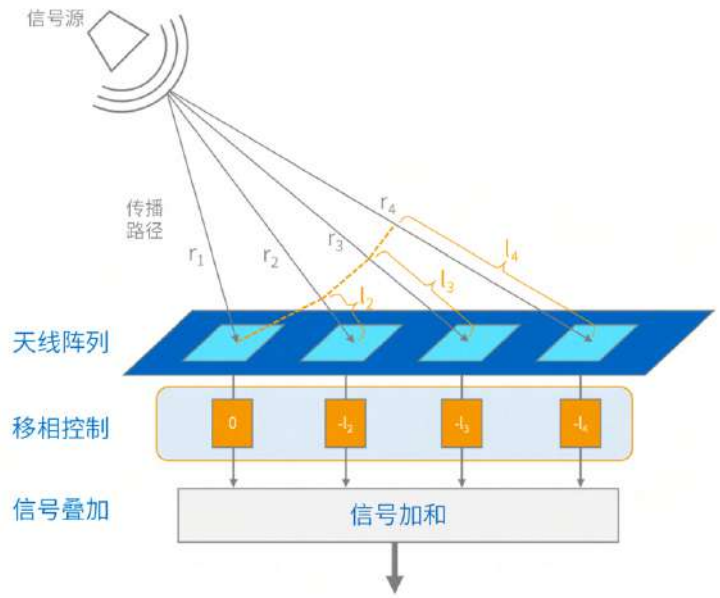


图 5.132: 接收通路中的相位控制

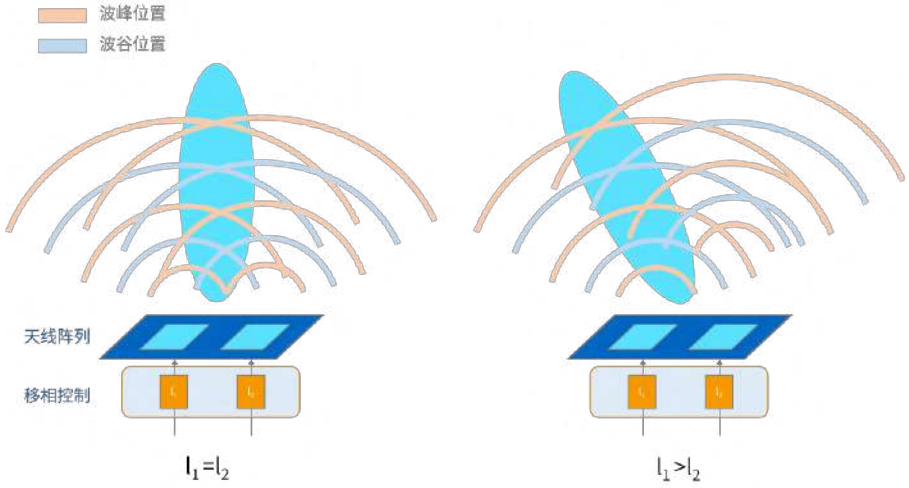


图 5.133: 通过两天线间信号相位差，对发射波束形状进行控制

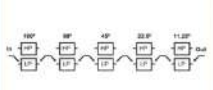
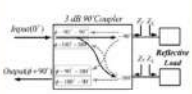
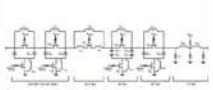
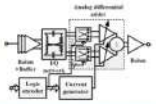
无源移相器			有源移相器 APS
高低通结构移相器 HP/LP-PS	反射式结构移相器 RTPS	开关电感电容移相器 STPS	
			
切换高通与低通滤波器实现相位差	改变耦合器的负载来改变反射信号的相位	由几个基本的开关与电感/电容移相单元构成	采用矢量合成的方式
• 设计较为简单 • 面积较大 • 插损较大 • 较难达到毫米波频率	• 适合于较高工作频率 • 移相范围有限，需要和其他结构移相器配合完成360度移相	• 精度较高、尺寸较小 • 插损较大 • 依赖精确的电磁仿真	• 占用面积小 • 精度高、误差小 • 电路结构较复杂

图 5.134: 不同移相器的架构及特点

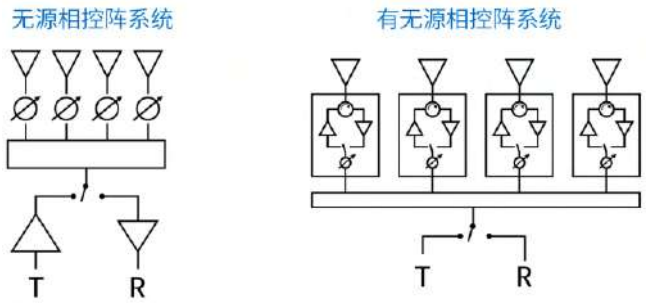


图 5.135: 无源相控阵系统，及有源相控阵系统架构

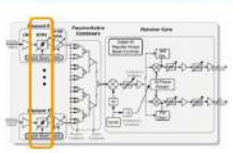
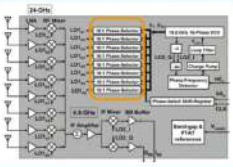
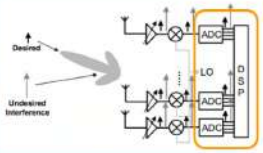
	射频移相架构	本振移相架构	数字移相架构
架构	移相器位于射频通路	移相器位于本振通路	在数字基带信号完成合成
			
优点	• 对强干扰信号有抑制 • 功耗及干扰较小	• 具有强干扰抑制能力 • 射频通路简化	• 模拟通路简化 • 可以完成复杂算法
难点	• 对移相器、可变增益放大器性能要求高（位于射频通路）	• 混频器需要抵抗强干扰 • 本振数量多，功耗及干扰增加	• 对整个系统动态范围要求提高 • 对ADC能力要求极高

图 5.136: 有源相控阵的系统架构

5.7.2.4 相控阵系统的分类

在相控阵系统分类中，主要分为无源相控阵和有源相控阵两种。无源及有源相控阵系统架构如图5.135。

两种系统都可以实现定向收发的天线阵，在实现上，无源相控阵系统的阵列由无源天线 + 移相器部分实现，信号的接收和发射均由中央接收机和发射机来实现。在有源相控阵雷达中，每个辐射器均配置有独立的有源接收/发射组件。

有源相控阵系统中，由于功率源前置至天线阵元，雷达系统更为稳定。并且因为每个通道上均有 T/R 组件，即使有少量的 T/R 组件损坏，整体性能也不会受到明显影响。由于每个通道可以独立工作，还可以对有源相控阵系统的单元组件进行分组，实现多目标同时跟踪等特性。

虽然无源相控系统只有一个发射接收组件，实现相对简单，成本也相对更低，但有源相控阵系统应用灵活、可靠性高，在雷达、无线通信中的应用更为广泛。

5.7.2.5 有源相控阵系统架构

相控阵系统实现中，最主要的功能就是实现移相。根据移相器在系统中所处的位置，有源相控阵系统可以分为如下三种架构 [10]。分别为：

- 射频移相架构
- 本振移相架构
- 数字移相架构

三种架构的实现方式和优缺点对比如图5.136。

在以上架构中，射频移相架构是当前应用较为广泛的实现架构。

5.7.3 毫米波 + 相控阵：优劣互补，相得益彰

以上分别讨论了毫米波、相控阵两大技术。虽然二者是独立的两大技术，但在使用中，经常将二者结合使用，两种技术相得益彰，实现优势互补：

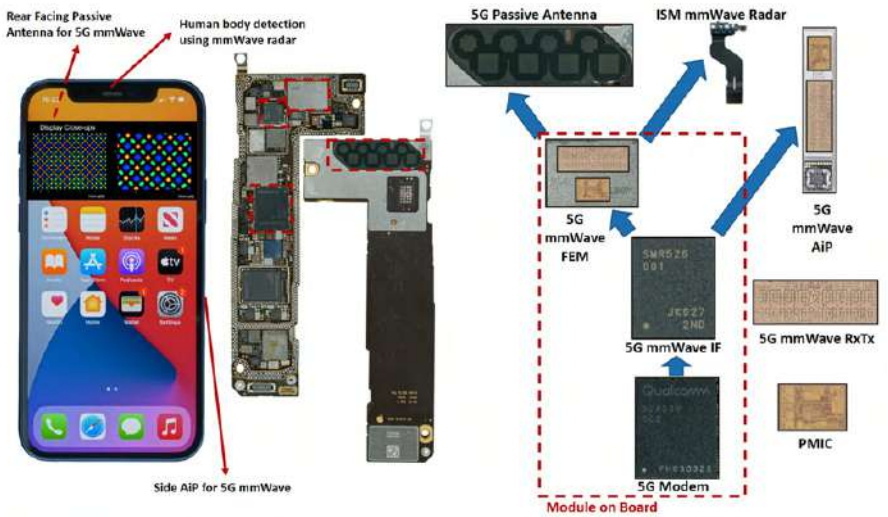


图 5.137: 搭载高通毫米波相控阵方案的 iPhone 12 手机（美版）

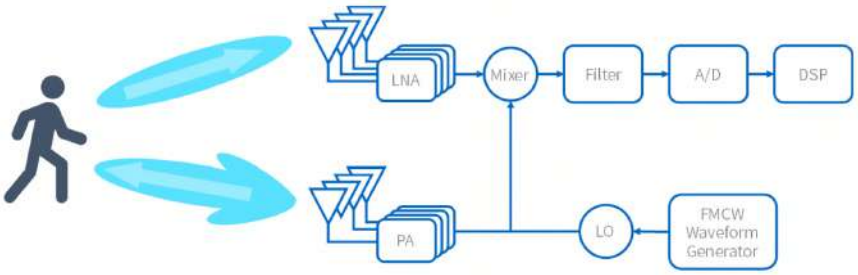


图 5.138: 典型车载雷达工作原理 [12]

1. 毫米波技术的特点是带宽大，但其路径损耗大、传播距离短，利用相控阵技术的波束聚焦功能，刚好可以将毫米波实现定向发射，增大传输距离。
2. 相控阵系统优点是可实现信号的定向发射，但由于需要几十甚至成百上千个阵列，造成电路面积增大。而毫米波电路面积小这个优势，刚好可以用于实现大规模阵列。

于是，“毫米波相控阵”这一组合相辅相成，在一些特定应用领域所向披靡。

5.7.3.1 毫米波相控阵系统应用

5G 手机

毫米波相控阵技术离我们并不遥远，不少 5G 手机中已经装备了此项技术。

在 2020 年 10 月份，苹果公司发布的 iPhone 12 中，北美版本中就加入了毫米波支持。iPhone 12 采用高通的毫米波方案，在手机顶部及侧面分别部署 4 天线毫米波阵列，实现毫米波信号的收发功能 [11], 如图5.137。

根据苹果公司提供的数据显示，搭载毫米波技术的 iPhone 12，最高可实现 4Gbps 的峰值下行速率。

车载毫米波雷达

车载毫米波雷达的工作原理是向被探测物体发射毫米波电磁波信号，并接收从目标反射回来的反射波，通过计算发射和接收信号的时间差，就可以对被测物体进行探测，如图5.138。

在实现方式上，车载毫米波雷达也需要借助毫米波相控阵技术，利用多天线阵列的方向，实现毫米波信号的精准赋形，实现对物体的精准探测。

图5.139为 24GHz 车载毫米波雷达的实现方案之一，在接收通路中，采用了 4 通道相控阵列的方式进行设计

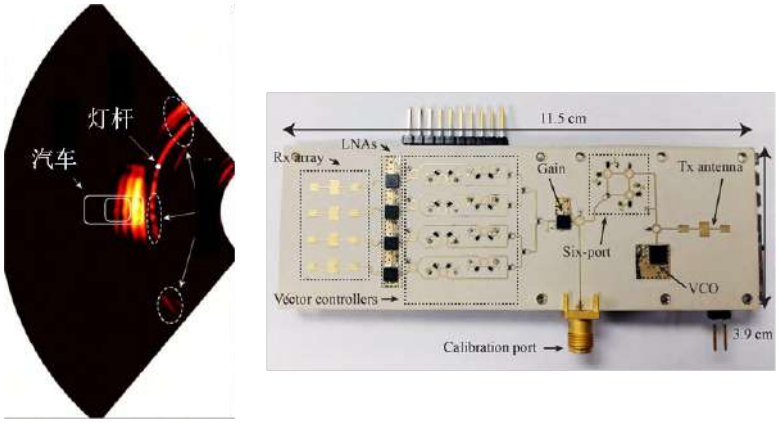


图 5.139: 24GHz 车载毫米波相控阵雷达系统

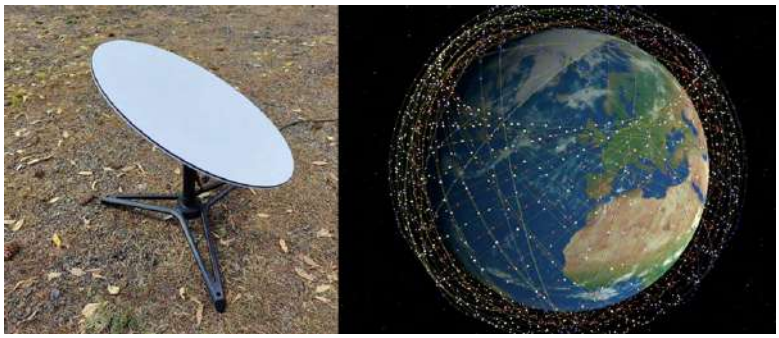


图 5.140: 星链系统所使用的地面站以及低轨卫星系统

[12]。

卫星通信

卫星通信是现在无线通信研究的一大热点，尤其是低轨卫星领域，由于其低延时、大带宽的特性，可以作为蜂窝通信很好的补盲使用。

虽然卫星通信有不受地理位置限制的优点，但实现起来并不容易。即使对于低轨卫星，其距离地球的距离也在 1,000 公里量级，基本相当于北京到上海的距离。而普通的地面蜂窝基站的传输距离只有数公里。想要在地面到卫星这种距离范围内直接建立信号连接并不容易，需要有高的发射功率，或者采用定向性强的发射系统。

另外，卫星的快速运转也给地空连接提出挑战。低轨卫星绕地球一圈的时间大约只有 100 分钟左右。如果以 60 度的可视角度计算，每一颗卫星在视角范围内的时间只有 17 分钟。并且卫星还在以每小时 3 万公里的速度快速飞行。这就需要地面站必须要有信号波束的快速扫描特性。

毫米波相控阵系统的波束定向性，以及电子相位控制的快速扫描特性刚好可以在卫星通信中一显身手。在 SpaceX 公司星链系统中，就使用了工作于毫米波的相控阵系统。星链系统所使用的地面站以及低轨卫星系统如图5.140。

星链系统将其地面站称为 Starlink Dish（星链盘），其直径为 58.9 厘米，外观类似于一个圆盘。在圆盘中，密集排列着 1,280 个天线阵列单元 [13]。通过下层连接的移相控制以及射频收发电路，实现高指向和快速扫描的毫米波相控阵系统，完成以 550 公里以外，3 万公里/小时快速移动的卫星连接。星链系统地面收发装置构成如图5.141。

5.7.4 总结

自 19 世纪末电磁波被发现以来，无线通信技术迅速发展。经过 100 多年的发展，无线通信技术已经不再是单纯的“收”、“发”这么简单，而是借助于不同频率、不同信号，甚至不同的天线技术完成强大的无线通信功能。

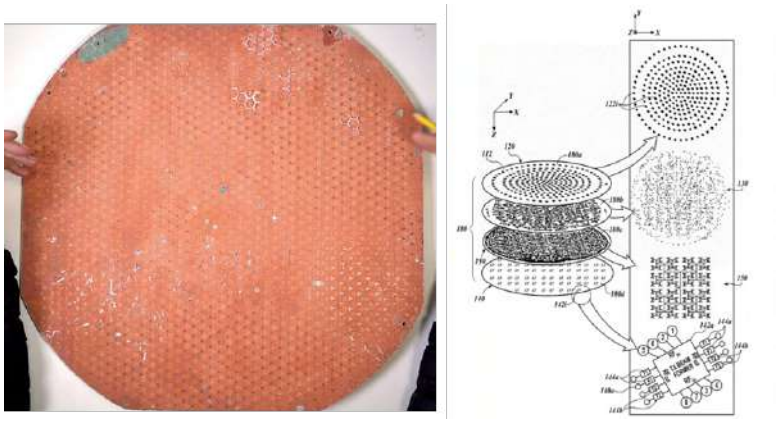


图 5.141: 星链系统地面收发装置构成

毫米波相控阵系统是无线通信技术发展中有代表性的技术突破，通过对大规模天线阵中输入信号的相位控制，实现了大带宽毫米波信号的定向传输，解决了毫米波信号路径损耗大的难题。

在 2020 年之前，对于毫米波相控阵系统的研究主要集中于军用、学术领域。在 2020 年之后，随着民用 5G 通信、智能汽车用毫米波雷达、民用卫星通信的发展，毫米波相控阵系统开始在民用领域逐渐普及。

2023 年，是时候了解“毫米波相控阵”技术了。慧智微期待和您一起，把握最新的无线互联技术。努力“化繁为简，使一切智慧互联”。

参考文献

1. 工业和信息化部关于微波通信系统频率使用规划调整及无线电管理有关事项的通知，2023/1/4
2. Cellular Broadband Device & Module Market, 2022, TSR
3. ICV: Automotive MMW Radar Market Research Report for 2022 H1
4. 卫星互联网开始加速，关注低轨卫星产业链机会，国信通信，2022/10/30
5. https://www.speedtest.net/
6. New Radar Applications Driven by Technology, Pasternack, Microwave Journal, 2017。
7. A. Natarajan, A. Komijani and A. Hajimiri, ”A fully integrated 24-GHz phased-array transmitter in CMOS,” in IEEE Journal of Solid-State Circuits, vol. 40, no. 12, pp. 2502-2514, Dec. 2005
8. Phased array on a fingertip, Thomas F. Krauss, Nature, 2013
9. SCR-270/SCR-271 Radar, https://pacificeagles.net/scr-270-scr-271-radar/
10. 面向射频移相架构相控阵，权星，2018
11. Elaboration on iPhone 12 5G antennas, https://www.reverse-costing.com/teardown-notes/elaboration-iphone-12-5g-antennas/
12. A K-Band Portable FMCW Radar with Beamforming Array for Short-Range Localization and Vital-Doppler Targets Discrimination, IEEE Transactions on Microwave Theory and Techniques, 2017
13. Starlink dish teardown, https://mobillegends.net/starlink-dish-teardown-part-1-spacex-bugbounty-is-open-during
14. https://www.nonstopsystems.com/radio/frank_radio_antenna_yagi.htm

在线阅读

本文摘自慧智微微信公众号，扫描以下二维码，在线阅读



5.8 “iPhone 13 支持卫星通信”，支持的究竟是什么？

作者：彭洋洋 博士

2021 年 9 月 7 日凌晨，苹果公司发布公告，2021 年的新 iPhone 将于 9 月 15 日发布。

前几日，一则“iPhone 13 手机将支持卫星通信”的消息传遍朋友圈，“可以卫星上网的新 iPhone”听起来着实酷炫。

有关苹果 iPhone 13 将支持卫星通信的消息，源自于著名的苹果分析师 Ming-Chi Kuo，Kuo 给出他的分析“基于定制的高通调制解调器，如果苹果手机启用了相关的软件功能，iPhone 13 的用户可以在没有 4G/5G 网站覆盖范围的情况下通过卫星打电话和发送消息”。Kuo 指出，与苹果合作的是 Globalstar 公司。

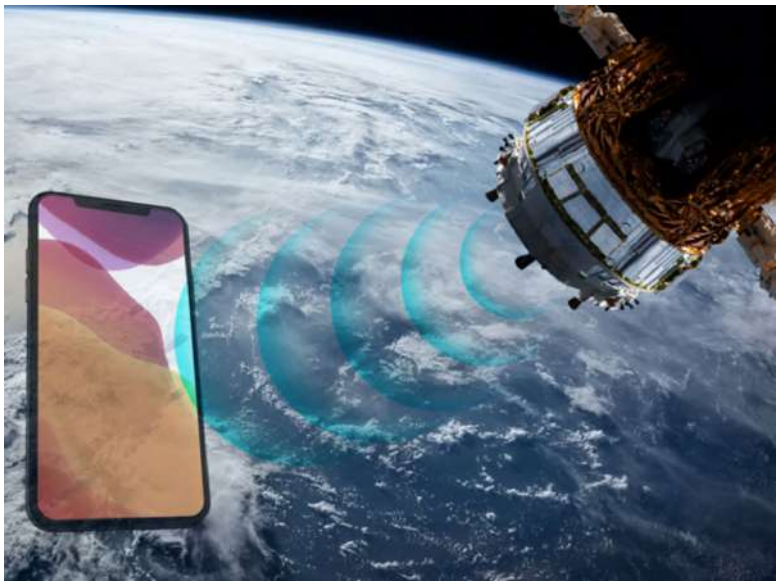


图 5.143: 传言 iPhone 13 手机将支持卫星通信

Globalstar 公司由来已久，是由美国劳拉公司（LoralCorporation，主要业务为太空通信）与高通公司于 1991 年共同出资成立的卫星通信公司，专注于低轨道卫星通信通信，与被广泛熟悉的“铱星计划”提供相类似的低速率数据传输服务。2020 年 3 月，Globalstar 宣布其所掌握的用于卫星通信的 n53/B53 频段，被 3GPP 批准为正式的 5G 频段 [1,2]。

卫星通信究竟有什么功能，需要什么条件？传言苹果 iPhone 13 可支持的 n53 频段究竟能做什么？什么样的射频前端可以支持到 n53 频段？

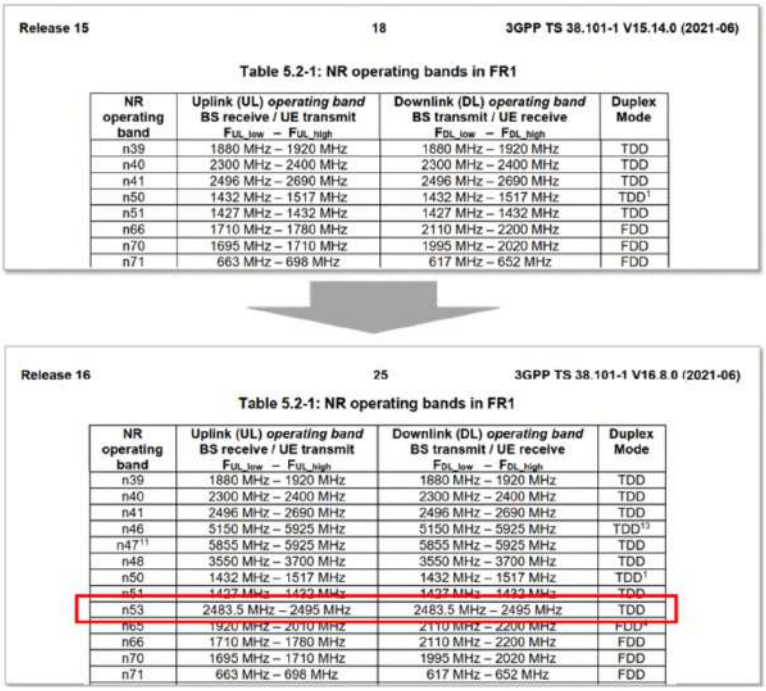
5.8.1 卫星通信:LEO(近轨) 卫星最被关注

卫星通信是利用卫星做为通信“基站”，移动用户与卫星直接相连后，通过卫星的中继，实现移动用户之间、用户与地面固网之间的互联。如图5.145，按照在轨高度，卫星移动通信分为:

1. GEO（Geosynchronous Earth orbit，静止轨道）卫星系统;
2. MEO（Medium EarthOrbits，中轨道）卫星系统;
3. LEO（Low Earth Orbit，低轨道）卫星系统。

其中，GEO（静止轨道）卫星系统运行在赤道上方高度约 35,800 km 的地球同步轨道，运行周期同地球相同，一般较少卫星即可实现全球除极低以外的覆盖。不过由于高度高，传播延时大。

目前有代表性的 GEO 卫星系统主要有由国际移动卫星公司经营的 Inmarsat（国际海事卫星）系统。该系统由 13 颗地球静止轨道卫星构成，可以向南北纬 83 度以内的地区提供电话、传真和数据业务，主要承担国际船舶、飞机等的航行、遇险安全通信任务，目前终端用户约 30 万。目前 Inmarsat 已经历经五代系统更替，前四代



Release 15 18 3GPP TS 38.101-1 V15.14.0 (2021-06)

Table 5.2-1: NR operating bands in FR1

NR operating band	Uplink (UL) *operating band* BS receive / UE transmit F_{UL_low} – F_{UL_high}	Downlink (DL) *operating band* BS transmit / UE receive F_{DL_low} – F_{DL_high}	Duplex Mode
n39	1880 MHz – 1920 MHz	1880 MHz – 1920 MHz	TDD
n40	2300 MHz – 2400 MHz	2300 MHz – 2400 MHz	TDD
n41	2496 MHz – 2690 MHz	2496 MHz – 2690 MHz	TDD
n50	1432 MHz – 1517 MHz	1432 MHz – 1517 MHz	TDD[1]
n51	1427 MHz – 1432 MHz	1427 MHz – 1432 MHz	TDD
n66	1710 MHz – 1780 MHz	2110 MHz – 2200 MHz	FDD
n70	1695 MHz – 1710 MHz	1995 MHz – 2020 MHz	FDD
n71	663 MHz – 698 MHz	617 MHz – 652 MHz	FDD

Release 16 25 3GPP TS 38.101-1 V16.8.0 (2021-06)

Table 5.2-1: NR operating bands in FR1

NR operating band	Uplink (UL) *operating band* BS receive / UE transmit F_{UL_low} – F_{UL_high}	Downlink (DL) *operating band* BS transmit / UE receive F_{DL_low} – F_{DL_high}	Duplex Mode
n39	1880 MHz – 1920 MHz	1880 MHz – 1920 MHz	TDD
n40	2300 MHz – 2400 MHz	2300 MHz – 2400 MHz	TDD
n41	2496 MHz – 2690 MHz	2496 MHz – 2690 MHz	TDD
n46	5150 MHz – 5925 MHz	5150 MHz – 5925 MHz	TDD[13]
n47[11]	5855 MHz – 5925 MHz	5855 MHz – 5925 MHz	TDD
n48	3550 MHz – 3700 MHz	3550 MHz – 3700 MHz	TDD
n50	1432 MHz – 1517 MHz	1432 MHz – 1517 MHz	TDD[1]
n51	1427 MHz – 1432 MHz	1427 MHz – 1432 MHz	TDD
n53	2483.5 MHz – 2495 MHz	2483.5 MHz – 2495 MHz	TDD
n65	1920 MHz – 2010 MHz	2110 MHz – 2200 MHz	FDD[4]
n66	1710 MHz – 1780 MHz	2110 MHz – 2200 MHz	FDD
n70	1695 MHz – 1710 MHz	1995 MHz – 2020 MHz	FDD
n71	663 MHz – 698 MHz	617 MHz – 652 MHz	FDD

图 5.144: 3GPP 批准 n53 做为正式的 5G NR 通信频段

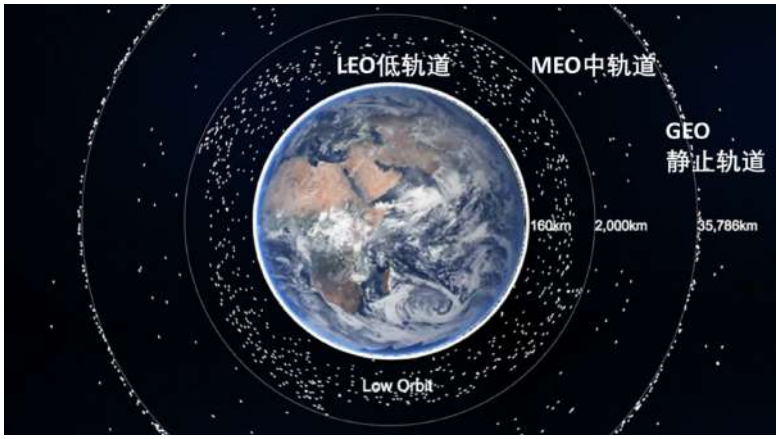


图 5.145: 卫星通信系统分类



图 5.146: 第一代铱星系统的宣传广告，及第一代铱星卫星

采用 L 波段频率（1 ~2GHz 频率范围内），第五代采用 Ka 波段（27 ~40GHz 频率范围内），可提供约 50Mbit/s 的下行及 5Mbit/s 的上行传输。

MEO（中轨）卫星系统轨道高度在 2,000 ~30,000 km 左右，因为其高度高，且无法实现与地球周期同步，目前发展不如 GEO 和 LEO 卫星系统。目前代表型的为 Odyssey 系统。Odyssey 为 TRW 空间技术公司推出，由 12 颗高度为 10,000 km 的卫星构成，可在全球范围为 280 万个用户提供服务，通信速率约为 4.8 ~19.2kb/s。

LEO（近轨）卫星系统是近年来广受关注的卫星通信系统，LEO 卫星系统轨道高度在 1,500 km 以下。由于处于近地轨道，若实现全球覆盖需要较多的卫星数量。但 LEO 卫星系统有卫星链路传播损耗小、传播时延短的优点，适合于多用户的全球卫星互联，是当前卫星通信关注的重点。以下将对 LEO 近轨卫星展开详细讨论。

5.8.1.1 第一代近轨卫星系统: 人类的雄心勃勃

LEO（近轨）卫星系统在 20 世纪末期经历了第一个发展高潮。当时全球移动通信的概念被全球关注及推行，在 90 年代所推行的第二代蜂窝移动通信系统，就取名为 GSM（Global System for Mobile Communications，全球移动通信系统），但蜂窝通信由于受限于基站建设，只能覆盖地面基站所覆盖的数公里范围，并不能实现真正意义的全球覆盖。

传言摩托罗拉一位工程师的妻子在加勒比海度假时，向丈夫抱怨没有信号无法接到客户电话。这位工程师突发奇想，如果用环绕地球的卫星组成一个星座，不就可以让通信覆盖地球上的每一个角落了吗？图5.146为第一代铱星卫星。

5.8.1.2 第二代近轨系统: 星向左，星链向右

第一代近轨卫星通信系统以商业的失败而告终。面对如何实现商业成功上，第二代铱星系统与 SpaceX 公司的星链（Starlink）系统选择了完全不同的路线。

铱星选择了寻找“土豪”用户。

铱星 2000 年申请破产保护和债务重组，并重新组建了新的铱星公司。新的铱星公司将营销对象转向了美国军方。2001 年，”911 事件”爆发，美国开始在全球进行反恐战争。新铱星公司等待的“土豪”用户终于出现，铱星公司与美国军方的合同进一步巩固。同时，不少美军官兵个人也购买了铱星电话和家人联系。铱星公司拥有了全球约 85 万的用户，营收也在稳步提高，2019 年，铱星达到了 5.6 亿美元的营收。

星链（Starlink）选择了降低卫星发射成本，服务广泛市场。

如果说当年的铱星计划是雄心勃勃，那么 2020 年的星链计划则是近乎疯狂。星链卫星系统准备在 2024 年前发射 4,425 颗卫星，这个数字甚至大于当前全球在轨的合计 3,300 颗卫星（截至 2020 年 11 月）。星链卫星系统计划长期共计发射 4.2 万颗 LEO（近轨）卫星，实现全球的覆盖。星链的目标用户为数亿需要卫星进行互联网



图 5.147: 一次性搭载 60 颗星链卫星的运载火箭

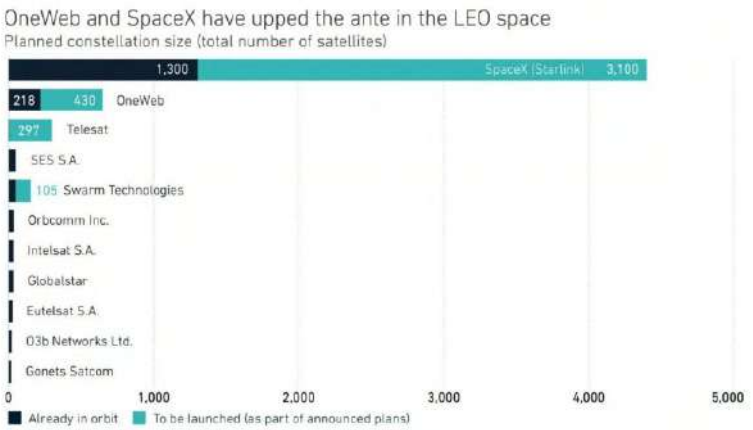


图 5.148: LEO（近轨）卫星当前发射数量及计划（来源：GSMA）

链接的用户。

为了满足广泛用户对于上网成本及上网质量的需求，SpaceX 开发了一系列降低卫星发射成本的技术。比如 SpaceX 的星链卫星仅 227 千克左右，一次火箭发射可以发射 60 颗星链卫星，如5.147；同时 SpaceX 的火箭还可以进行火箭回收，重复利用，用以进一步降低发射成本。这和当年铱星时代的一次发射只承载几颗卫星并且火箭不可被回收相比，卫星发射成本有了数量级的降低。

除了 SpaceX 公司的星链计划，英国卫星通信公司的 OneWeb 计划提出将发射 720 颗低轨卫星（第一期 648 颗）。我国也不甘落后，我国中国航天科工集成的“虹云”卫星通信和中国航天集团的“鸿雁”卫星通信系统，均是为低轨通信设计的卫星通信系统。LEO（近轨）卫星当前发射数量及计划如图5.148。

5.8.2 卫星通信: 是补充, 非替代

大家看到卫星通信，就会想当然的认为可以绕过运营商进行通话、免费上网打电话了，就可以取代现在在用的 4G/5G 甚至是宽带了。其实不然，**卫星通信只能做为陆地蜂窝通信的补充，无法做到对陆地通信系统的取代。**

5.8.2.1 卫星通信速度受限

随着星链计划的推出，以及铱星计划的升级，近地卫星通信开始进入 Mbit/s 量级，星链计划所推出的 150Mbit/s 为最快标称速率。而陆地移动通信的 5G 系统愿景为达到 10Gbit/s 的通信速率，相比近地卫星通信高近 100 倍。而对于类似全球星（Globalstar）系统，通信速率只有数 kbit/s，几乎只能完成语音及短信业务，无法完成互联网数据业务。全球主要卫星系统的数量及速度如图5.149所示。

	卫星	频率	通信速率	覆盖范围
Inmarsat （国际海事卫星组织）	11颗GEO卫星	上行1626.5MHz-1660.5MHz 下行1525-1559MHz	最高5Mbit/s，下行50Mbit/s	
铱星 (Iridium)	66颗LEO	1621.32-1626.5MHz L band; 第二代采用Ka band	4.8kbit/s (L band); 1.5Mbit/s; 10Mbit/s; 30Mbit/s (Ka Band)	
全球星 (Globalstar)	48颗LEO	上行1610-1626.5MHz 下行2483.5-2500MHz	2.4/4.8/9.6kbit/s的语音业务 7.2/9.6kbit/s的数据业务	
星链 (Starlink)	LEO卫星： 2024年4425颗 2027年7518颗 预备占位30000颗 总计4.2万	Ka、Ku波段	50~150Mbit/s	目标全球，当前建设中

慧智微整理

图 5.149: 主要卫星通信系统特点

5.8.2.2 卫星通信受环境制约

地面站与卫星通信，星间链路通信通常采用微波毫米波波段（如 Ka、Ku 波段），容量大，并且高频率有助于做高天线增益。

移动终端与卫星通信的频段主要有两种，分别是 1 ~4GHz 频率范围内的 L/S 波段，以及微波毫米波的 Ka/Ku 波段。

1. L/S 波段的优势是技术简单，不需要毫米波的波束成形技术，缺点是无法得到天线增益，空间损耗大，速率受限；
2. 微波毫米波频段通信可以利用波束成形技术提高天线增益，但终端复杂度提高，通常需要专用的发射接收设备。

另外，卫星通信距离长，卫星通信的长度约为 1,000 公里。根据自由空间损耗计算公式：

$$Loss = 32.45 + 20\log_{10} F + 20\log_{10} D$$

其中，F 为传输频率（MHz），D 为传输距离（km）。对于使用 1.65GHz 进行上行传输的全球星系统，1,000 公里传输的路径损耗为 156.8dB，相较于距离通信基站 1km 左右距离，损耗增大约 60dB。

5.8.2.3 卫星通信连接数受限

目前，陆地蜂窝通信系统已在全球建成约 700 万个物理站点，覆盖 52 亿全球用户使用 [4]。即使星链卫星系统宏大的 4.2 万颗卫星系统建立成功，其数目也只是全球基站数目的 0.6%。每颗卫星需要覆盖相当于两个上海的面积。单个卫星需要覆盖范围面积大，用户容量大大受限。

根据测算，铱星系统全球最大承载的用户数目为 1,470 万 [5]，这只相当于上海人口的一半左右，远低于陆地蜂窝通信系统。

5.8.2.4 卫星通信成本高

在目前提供卫星通信的服务商中，星链公司的服务费用为每月 99 美金（约 650 元人民币），可提供网络服务；全球星服务费约 20 美金（约 130 人民币），提供紧急呼叫服务。对比陆地移动通信系统部署发达地区的单月 10 美金（约 65 人民币）的上网套餐相比，通信成本高。

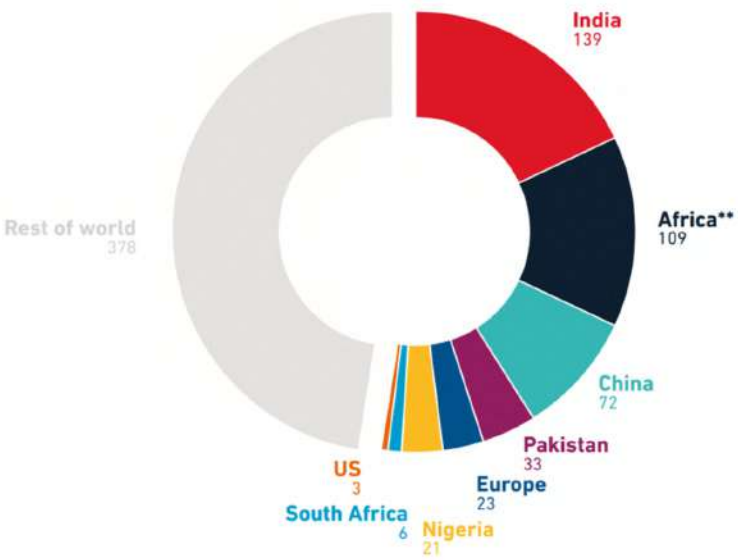


图 5.150: 未被陆地蜂窝网络所覆盖人员数量（百万，来源 GSMA）

虽然有以上不足，但相比于依赖于陆地基站建设的陆地蜂窝系统，卫星通信系统有以下优势。

5.8.2.5 优势 1: 卫星通信覆盖范围广阔

在陆地蜂窝系统建设中，由于基站建设受限于物理环境，无法建设和覆盖沙漠、海洋等无人区域。而卫星通信的广泛覆盖，可以提供“无人区”的覆盖。

同时，由于全球经济发展的不均衡，如图5.150，全球大约有 7 亿用户无法使用互联网 [4]，占全球总人数的的 10%，其中印度及非洲分别有 1.39 亿及 1.09 亿人口。卫星通信的的广泛覆盖特性，可帮助全球还没有联网能力的地区实现联网覆盖。

随着物联网的发展，大量处于非基站建设范围的物联网终端需要联网，这些物联网终端可能无法采用陆地基站连接，如自然灾害监测系统、海洋轮船、远程无人机等，卫星通信可以解决这些物联网终端的互联需求。

5.8.2.6 优势 2: 低密度地区更适合卫星通信

对于人员低密度地区，卫星通信可以避免光纤基建的过度投入，在一些低密度地区，如美国、加拿大、澳大利亚、俄罗斯等地区，由于人口密度低，光纤全面建设的成本和人口的全面联网的需求之间构成一对矛盾，卫星通信可以很好的解决这一矛盾。

综上所述，卫星通信有以下特点：

- 是陆地蜂窝移动通信的补充，不会对陆地蜂窝移动通信构成替代。
- 主要以提供广泛覆盖下的低速连接服务为主。
- 通信速率和连接数目受限，无法为全球用户提供高速率的连接。

对于文章开头提到的 iPhone13 对卫星频段的支持，传言与苹果合作的是全球星（Globalstar）公司。全球星系统仅可支持低速率的语音和短信发送功能。**即使 iPhone 13 可对卫星通信支持，还是无法利用卫星通信享受到高速上网功能，但可享受到全球星覆盖范围内的 SOS 紧急定位及呼叫功能。**

5.8.3 射频前端方案: 可重构支持链路复用, 场景优化

过去 20 多年移动通信的发展，通信协议从 2G 快速演进到 5G，但演进还远没有结束。随着近年来卫星通信的发展，移动设备有可能加入卫星通信功能，作为陆地蜂窝通信的补充，实现真正意义的“全球通”终端。

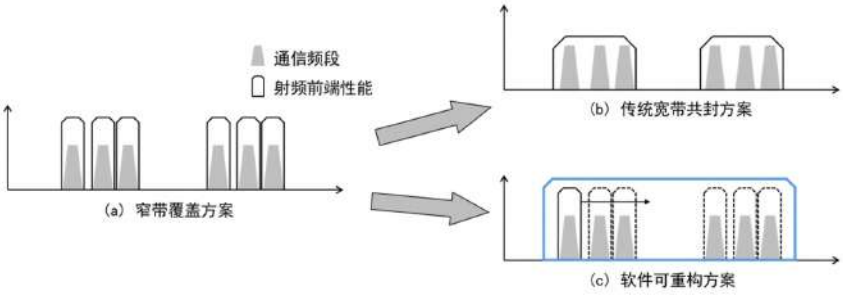


图 5.151: 针对单个模式场景优化的软件可重构方案

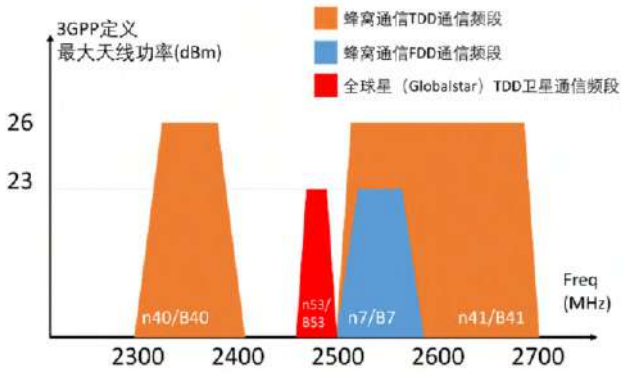


图 5.152: 全球星（Globalstar）卫星通信频段与蜂窝移动通信频段关系

但移动终端目前已经同时共存了大量的蜂窝通信制式与模式，统计显示，目前手机中已经支持近 40 个通信频段，覆盖 2G (GSM、CDMA)、3G（WCDMA、CDMA2000、TD-SCDMA）、4G（TDD-LTE、FDD-LET）、5G（NR）等 [6]。在对最新的 iPhone 手机拆解中，有近一半的芯片与无线通信相关。卫星的连接的加入，加进一步增加射频前端方案的复杂度。

与当前无线终端现有的射频前端复用是一种解决方案，但卫星通信的功率、制式、波形等与当前移动终端不同，射频前端需要有能力针对卫星连接场景进行优化。

可重构射频前端可以实现链路复用，以及针对不同场景的优化。

5.8.3.1 可重构射频前端: 场景针对性优化, 支持后续升级

为了支持多频段多模式，传统宽带射频前端方案一般将相近的频段进行宽带覆盖，虽然损失了性能，但可以在一定程度上减少芯片的射频通路，在过去 10 年中，被业界广泛采用。

采用传统方案对性能的折损在 LTE/NR Sub-3GHz 高频段表现的尤为明显，在这个频率范围，对功率有高要求的 n41 与对效率有高要求的 B7 在同一组频率范围中，构成一对矛盾。传统宽带方案用一颗固定的宽带 PA 将两个频段同时覆盖，造成 n41 的高线性、B7 的低电流均无法达成。

近年来，不断有公司投入到移动终端多频段多模式可重构射频前端方案中，慧智微电子是此技术方案的开创和代表公司。可重构方案是利用软件和硬件的协同，使射频前端在感知当前应用场景后，通过软件调谐的方式，对此应用场景进行有针对性的性能优化，使各个频段和模式应用时，性能均达到最优。如图5.151。

5.8.3.2 可重构射频前端对 n53 的场景优化支持

在卫星通信连接中，由于“iPhone13 支持 n53”的传言，n53 成为备受关注的卫星通信频段。n53 频段位于 LTE/NR Sub-3GHz 通信频段的高频段频率范围内（2.3-2.7GHz），属于高频 TDD 频段。此频段处于 HPUE TDD 关键频段 n40/n41 之间，处于 FDD 关键频段 B7 旁边。如图5.152。

频段名称	n53
频率范围	2483.5 MHz – 2495 MHz （上行） 2483.5 MHz – 2495 MHz （下行）
双工类型	TDD 时分
最大信号带宽	10MHz
功率等级	23dBm +/-2dB （PC3）

图 5.153: n53 的主要性能参数

	可重构 5G NR Phase5N 射频前端 S55643-11		竞品 A	
	线性度 ACLR (dBc)	电流 (mA)	线性度 ACLR	电流
n53	-42 dBc	689 mA	-36 dBc	861 mA
B7 (FDD 关键频段，电流有高要求)	-40 dBc	653 mA	-38 dBc	857 mA
n41 (PC2) (TDD 关键频段，线性度有高要求)	-40 dBc	952 mA	-37 dBc	1150 mA

测试条件：Room temp；EVB测试；PC3=29dBm/PC2=31dBm

图 5.154: S55643-11 产品对 n53 频段支持性能参数

在最新的 3GPP 定义中，n53 的主要性能参数要求如图5.153[1]：

S55643-11 是慧智微基于慧智微第三代可重构射频前端平台 AgiPAM® 3.0 开发的 Phase5N 射频前端发射模组。该模组尺寸为 4x6.8mm^2，利用软件调谐方式，支持 Sub-3GHz 低中高 4G 及 5G 频段。针对 n53 作为卫星通信频段，窄带、高线性的需求，S55643-11 可做针对性的软件调谐优化，并且不需要对其他频段性能进行折中。图5.154为 S55643-11 产品对 n53 频段及高频关键频段 B7/n41 进行性能测试。

通过以上测试数据，可以看到，慧智微可重构射频前端 S55643-11，可实现针对不同场景特殊优化。在新引入的 n53 卫星通信频段，可实现线性度优于竞品 6dBc，电流优于竞品 172mA。同时保持关键 FDD 频段 B7 实现电流优于竞品 204mA，关键 TDD 频段 n41 线性度优于竞品 3dBc。

目前，3GPP 协议对 n53 的定义还不完整，一些如带外抑制、共存干扰等指标还没有明确定义，并且可能还有其他协议对 n53 频段的使用进行规范，n53 的性能需求无法短期收敛。而传统宽带射频前端只能等待协议固化后，再重新做硬件设计来达成对 n53 的优化支持。慧智微软件定义的 Phase5N 方案射频前端 S55643-11，可支持协议后续升级，确保客户 n53 终端方案以用最短的时间推向市场。

5.8.3.3 可重构射频前端对 1.6GHz 卫星频段的扩展支持

对于蜂窝通信系统，中频段的最低频率为 Band3，1.71GHz-1.785GHz，这也是传统宽带射频前端的中频段功率放大器支持的最低频率。而卫星频段通常用到的 1.6GHz 频段无法支持。

可重构射频前端可以利用软件调谐的特性，将最低支持频率调谐至 1.6GHz，在不增加射频通路的情况下，支持卫星通信频段。如果有这方面的需要，请与慧智微销售及技术支持团队联系，进行具体规格讨论。

5.8.4 总结

卫星通信由于其广覆盖的特性备受关注，可以作为现在移动通信的有效补充，有望是移动终端通信的下一个功能引爆点。当前射频前端芯片及方案需要兼容考虑对未来潜在通信制式的支持。

卫星通信由于建设成本高，运行周期长。在卫星通信设计中已经考虑支持在轨软件定义和功能重构 [7]，确保在卫星使用的 7.5 到 10 年生命周期内，实现星上通信资源的高效按需调度和配置，能够针对不断变化的应用

需求，及时做出调整。

对于移动终端而言，需要有能力承接快速变化的协议需求，灵活的支持新协议的加入及未来演进。可重构射频前端利用软硬件协同特性，可以支持在协议演进中快速收敛产品，帮助终端快速面市。

参考文献

1. 3GPP TS 38.101-1 V17.2.0, 3GPP.
2. 3GPP TS 38.101-1 V15.14.0, 3GPP.
3. GSMA: Radar-Connectivity from the sky, GSMA.
4. GSMA: The Mobile Economy 2020, GSMA.
5. 吴寅, 徐展琦, 周诠. 低轨道卫星通信系统的容量与星间链路数计算 [C]//2006 军事电子信息学术会议. 中国电子学会, 2006.
6. RF Front Ends (RFFE) for Mobile Devices, 2021,Mobile Expert.
7. 肖永伟, 孙晨华, 赵伟松. 低轨通信星座发展的思考 [J]. 国际太空, 2018, 000(011):24-32.
8. https://en.wikipedia.org/wiki/Globalstar
9. https://en.wikipedia.org/wiki/Starlink

在线阅读

本文摘自慧智微微信公众号，扫描以下二维码，在线阅读

5.9 为什么有的手机能连卫星？

作者：彭洋洋 博士

2024 年 11 月 26 日，华为 Mate70 发布。这部手机的到来，再次把手机卫星通信技术推向了新高度。

在这部手机中不仅在体验上加入了“卫星寻呼”功能，让用户可以一次连星之后持续在线，解决了连接卫星的焦虑。并且，还带来了另一项重磅消息：MateX6 手机中，还会连接低轨卫星互联网，成为全球首款支持三网卫星通信（北斗卫星消息、天通卫星通信、低轨卫星互联网）的大众手机。



图 5.156: 华为消费者业务 CEO 余承东发布华为三网卫星通信手机 Mate X6

手机究竟是如何连接上卫星的？为什么有的手机连接几公里的基站都费劲，有些手机却可以连接上几万公里外的卫星？手机连接卫星是未来的趋势吗？

带着以上问题，我们对手机连接卫星的技术方案做一个梳理，尝试去理解一下未来手机直连卫星的技术发展。

5.9.1 手机连接卫星

是噱头，还是真有需求？

自 2022 年 9 月 6 日华为发布首款消费类大众卫星手机 Mate50 以来，苹果、vivo、OPPO、小米、荣耀等厂商陆续发布可以直接连接卫星的手机。

但即使得到众多厂商的推崇，直连卫星手机的出货量并不高。除华为和苹果外，其他品牌厂商支持卫星通信的手机年发货量普遍在百万台甚至以下，这和每个品牌手机 1 亿左右的年出货量相比，占比只有约 1%。即使对于直连卫星手机的引领者华为，也并没有把直连卫星定义为日常应用，而是归类到了“一生用一次，一次续一生”的应急场景。这不仅让人有了疑问：**手机连接卫星，是不是只是个噱头？**

5.9.1.1 在少部分场景下，确实有需求

在过去 20 年智能手机的发展中，智能手机已经变的非常强大。一部小小的手机，可以覆盖我们生活中的大部分场景。

作为消费类电子产品中，每年出货量最大的品类，手机在过去的设计思路，也是尽量的去满足最广大用户的最频繁需求，然后把这些需求不断的去做优化。

比如，过去 10 年期间，手机先是卷流畅顺滑，再卷摄像头，然后到屏幕、续航等等，这些需求作为普通用户都是可以被明显感知的，所以也成为不同手机厂商竞争的重点。

相比这些显性功能而言，**“通信”这一手机的基础功能，反而在过去并不是宣传的重点。**

通信未能成为手机宣传重点的原因，是因为通信的差异太难被用户感知了。

对于大部分用户来说，通信略显专业，让人望而却步；对于广大通信从业者来说，也很难讲清楚自己花大量努力做的优化，可以给用户体验做什么提升。所以在很长的时间里，手机的通信技术就像是一个黑盒子，被大家避而不谈。

但这一切，在手机各个功能做到极致，用户追求更差异化的需求的时候，会有改变。

近两年来，各个手机厂商开始不断的在通信能力上加大投入，开始“卷”通信这个基础能力。

- 2023 年，荣耀 Magic5 Pro 搭载自研射频增强芯片 C1，它能在弱信号环境下，帮助用户快速回网，确保直播流畅不卡顿。
- 2024 年，vivo X200、小米 15 Pro 发布手机无网通信技术，可支持在无网空旷环境下，支持最远超过 1500 米点对点语音对讲及文字传输，确保用户在徒步、露营等无网环境下仍能保持通信不断联。
- 2024 年，OPPO Find X8 中发布 VIP 专属连接的山海通信技术，确保用户在各种复杂的环境，尤其是在演唱会等网络拥堵场合，都能享受到快速且稳定的网络连接。



图 5.157: 荣耀、vivo、小米、OPPO 等各手机厂商，均在通信上展开投入

无论是弱网环境、无网郊区、演唱会，这些场景有一个共同特点是，都不是用户常用场景。但这些特殊场景的优化，却可以明显拉开不同厂商通信能力的区别。成为体现手机差异化的“明星场景”，也成为厂商通信能力差别的重要体现。

手机连接卫星，需要解决的，就是这样一个场景。在广阔无人的荒漠中，当所有的通信手段都失效时，手机直连卫星就成为与外部连接的唯一手段，甚至可以拯救人的一生。

这些场景对于 99% 的用户来说，可能都不会考虑。但对于那 1% 的用户来说，关键的时候确实可以用来“续一生”。

5.9.1.2 技术仍受限，体验待提升

手机直连卫星还没有得到广泛应用的另外一个原因，是还没有办法做到丝滑的使用体验。

和我们日常使用手机，拿起手机就可以电话和视频不同，手机直连卫星的体验仍然没有办法做到如此丝滑。

首先，在手机和卫星通信之前，手机需要连接卫星。连接卫星需要把手机不断的尝试切换角度，尽量对准天空那些看不到的卫星。连接时间很有可能需要一分钟甚至更久的时间。



图 5.158: 使用前需要对星

其次，一些卫星方案只支持短消息，甚至还是单向短消息。这就是说，如果只是单向消息，你只能发送一条消息，但这条消息对方是否接收，你是完全不知道的，对方也没有办法向你回复消息。



图 5.159: 苹果手机的卫星短消息功能

虽然有一些手机开始支持卫星语音，但语音数据必须要进行极大的压缩，来满足当前低速率的传输需求。在压缩之后，语音的清晰度和可辨别性都大量降低。

网络上有手机卫星语音通话测试，大家可以尝试听听看，看看是否可以清楚识别语音内容。

虽然以上这种体验的时间消耗、精力消耗在“一次续一生”时都是值得去等待的，但在日常使用中，这种体验目前确实很难打动消费者。也成为卫星直连手机现在较难被广大民众所接受的原因。

值得一提的是，以上所有的讨论是手机直连卫星这一应用。虽然有一些低轨宽带卫星方案（如 SpaceX 的星链，我国垣信卫星的千帆星座等），可以实现数十 M 每秒的稳定数据传输，但这些方案是利用毫米波相控阵方案实现，需要在地面端使用较大的收发装置，并不是我们讨论的便携式的“手机”直连场景。



图 5.160: 宽带低轨卫星需要部署的地面收发装置

5.9.2 手机直连卫星的技术方案

直连卫星的手机在使用体验上和用户需求仍有差距，正是这些差距让手机直连卫星技术有了发展方向。

在技术实现了，目前手机直连卫星主要有三种实现形式。分别为：

1. 卫星不变，改手机
2. 手机不变，改卫星
3. 卫星、手机一起升级

5.9.2.1 卫星不变，改手机

“卫星不变，改手机”，是指利用现有的卫星，利用现有的卫星通信协议，改变手机方案，在手机中增加相关卫星制式的支持。

这些方案一般由卫星运营商、手机厂商、芯片厂商共同开发，通过定制手机专用的卫星通信芯片，完成卫星通信的支持。近年来所发布的华为 Mate50（北斗短报文）、Mate60（天通语音 + 短信），苹果 iPhone14（Global Star 短信）等手机终端均是采用此方案。

在这些方案下，手机不再只是一部支持地面蜂窝连接的手机，而是变成一部手机卫星终端。比如在华为 Mate60 Pro 的手机包装上，此款手机就被的名称就改为：卫星移动终端。



图 5.161: 华为手机，变成了卫星终端

采用此方案的优点是只需要改变手机方案，而不需要重新发射和运营卫星。这在卫星建设不成熟的阶段，是最为有效的部署方式。

但采用这种方案，可采用的卫星资源并不多。原来的卫星方案也并不是为广泛的手机用户所设计的，所以在连接数目、连接速率上，都会存在限制。目前，采用这种方案的卫星直连手机通信速率在 0.8 ~2.4kb/s 左右 [1]。

5.9.2.2 手机不变，改变卫星

利用现有卫星来实现手机直连虽然实现起来更为快捷，但受限于可服务用户数量，这种技术方案在大规模应用时会遇到限制。在卫星技术逐渐发展后，有厂商开始考虑将卫星技术做升级，从而可以在不改变用户手机的情况下，让卫星可以连接上数以亿计的存量手机。

美国的 SpaceX 公司、AST&Science 公司、Lynk Global 公司等，发布了一些技术原型。这些公司基于自己不同的技术优势来做这方面的尝试，但思路均是“把地面的基站搬到天上”，同时，用大规模的相控阵系统，增大天线增益，让卫星可以接收到手机发出的微弱信号；同时有足够大的发射功率，让地面的手机可以接收到。

采用这种技术的最大优点是地面用户几乎是无感的：手机用户本来就是要在不同的基站之间做切换，而增加一个卫星基站，只不过是让用户在必要的时候，切换到天空的基站而已。如此一来，用户的切换成本大大降低。

但这么做并不是没有成本，现在地面通信协议是为地面几公里范围内的手机和基站通信所设计，如果把基站搬到天上，保证同样的传输下，就需要基站的能力更强，所以就需要卫星上基站有更高的输出功率、更大的天线增益。卫星的体积也会更大，发射成本会极大提高。以 AST 公司的 BlueWalker3 举例，其卫星重量达到了 1.5 吨，天线总面积达到 64 平方米。而相比起来，用于低轨宽带通信的卫星重量一般只有 200 千克。



图 5.162: AST BlueWalker3 卫星相控阵天线图

目前，存量手机直连技术方案也是创业热点，AST 公司计划在 2025 年前后部署 168 颗卫星，Lynk 公里也宣称在 2022 年已可以提供最高 200kb/s 速率的 LTE 通信，SpaceX 的 V2.0 与传统运行商 T-Mobile 合作，完成了现有手机直连卫星的测试，速率达到 17Mb/s。

5.9.2.3 卫星、手机一起升级

虽然基于现有技术，已经可以实现手机和卫星的连接，但未来更优的方案，无疑是根据未来需求，将卫星与手机一起升级。

虽然手机直连卫星是行业普遍认可的技术发展方向，但标准如何统一成为行业的一大难点。卫星企业有一套自有的标准，并且长时间以来不在手机通信产业中；手机通信产业的标准又没办法适应卫星通信需求。以上现状对标准的统一和升级提出了极高的挑战。

手机通信产业一直对卫星通信技术保持高度关注。2022 年 6 月，手机通信行业的标准化组织 3GPP 宣布 5G R17 版本标准协议冻结，在 R17 协议中，一项重要的特性就是 NTN（NTN：Non Terrestrial Network，非地面网络），这意味着未来行业终于有了专门针对手机直连卫星的通信协议。

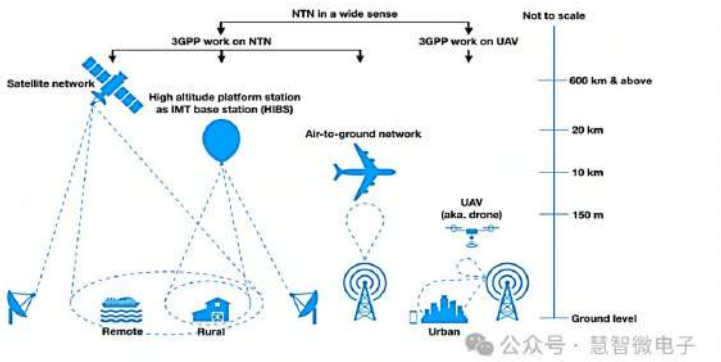


图 5.163: 3GPP 对于 NTN 的规划

虽然协议的首个版本基于原来的 NB-IoT 协议演进而来，R17 协议也先研究了用于窄带通信的卫星通信通信协议，但这丝毫影响不了行业对 NTN 协议的期待。2023 年 1 月，欧盟 6G-NTN 研究项目正式启动，2023 年 3 月，美国 FCC 通过了太空的补充覆盖新规则，来支持卫星企业和运营商的合作。

我国也快速展开了对于 NTN 的研究工作，包括中国移动、中国电信在内的运营商，华为、vivo、小米等手机企业，以及展锐、慧智微等芯片厂商，共同讨论并起草 NTN 行业标准，推动我们 NTN 的行业发展。

在 2024 年 MWC 上，展锐首次发布了 NR NTN 卫星通信系统验证，联发科展示了 Pre-6G NTN 卫星宽带技术方案。华为、中兴等厂商完成了 NTN 基站研发，vivo 等公司也推出支持 NTN 的试验终端。



图 5.164: 展锐展示的 NTN 芯片，及 vivo 的 NTN 手机

目前手机直连卫星主要集中于以上三种方案。现有手机直连卫星系统中，所采用的卫星及协议总结如下表5.165所示。

卫星系统	卫星			协议		
	现有高轨	现有低轨	新建低轨	原卫星协议	原蜂窝协议	新NTN协议
北斗	✓			✓		
天通	✓			✓		
Global Star（苹果）		✓		✓		
铱星（高通）		✓		✓		
AST			✓		✓	
星链V2			✓		✓	
Lynk			✓		✓	
Skyterra（Ligado）	✓					✓
Immarast（联发科）	✓					✓

公众号 · 慧智微电子

图 5.165: 现有手机直连卫星系统中所采用的卫星及协议

5.9.3 手机直连卫星的技术趋势

随着技术发展，在手机直接卫星技术上，未来可能有以下趋势。

- **需求场景将更加广泛且具体化**

  在当前阶段，虽然手机直连卫星的技术已经初露锋芒，但其应用需求尚显得较为笼统与不确定。展望未来，这一技术的需求场景将逐渐明确并呈现多样化趋势。

  随着技术的不断成熟与普及，手机直连卫星将不再仅仅局限于紧急通讯，而是渗透到日常生活的方方面面。无论是户外探险者需要的实时定位与导航，还是海上航行者对于稳定通信的渴望，抑或是自然灾害发生时对于应急通信的迫切需求，手机直连卫星都将成为不可或缺的技术支撑。

  这些具体而实际的需求场景，将推动手机直连卫星技术不断向前发展，实现更广泛的应用与更深入人心的服务。



图 5.166: 卫星实现未来通信的全球覆盖

- **用户体验将实现质的飞跃**

  当前，手机直连卫星在速率、连接稳定性等方面仍存在诸多限制，影响了用户的实际使用体验。然而，随着技术的快速发展，这一现状并不会持续太久。

  以华为 Mate 系列的手机为例，从 2022 年 Mate50 只支持短消息，到 Mate60 支持通话，再到 Mate70 的卫星寻呼，卫星直连手机的体验在不断优化，展示了手机直连卫星在提升用户体验方面的巨大潜力。

  随着技术的不断进步与迭代，预计未来手机直连卫星的速率将大幅提升，连接稳定性也将得到显著增强。用户将能够享受到更加流畅、稳定的卫星通信服务，无论是在偏远山区还是广阔海域，都能实现无缝连接与高效通信。这种用户体验的质的飞跃，将使得手机直连卫星技术成为更多用户的选择。



图 5.167: 华为手机直连卫星技术发展

- **产业发展助力成本降低，推动广泛应用**

  手机直连卫星技术的产业发展，不仅将推动技术本身的不断进步，还将带来成本的大幅降低。随着产业链

的不断完善与规模化生产，手机直连卫星所需的硬件设备成本将逐渐下降。

这意味着，未来不仅手机能够实现与卫星的直连，无人机、智能穿戴设备、远程监控设备等多种终端设备也会接入卫星网络。这种成本的降低与应用的拓展，将形成良性循环，进一步推动手机直连卫星技术的普及与应用。

随着越来越多的用户享受到这一技术带来的便利与高效，手机直连卫星将成为未来通信领域不可或缺的重要组成部分，引领着通信技术的革新与发展。

5.9.4 总结

手机直连卫星技术，作为通信技术领域的一项重大突破，正逐渐从概念走向现实，并展现出广阔的应用前景。

在未来 6G 到来之时，手机直连卫星技术作为未来通信领域的重要组成部分，将和 6G 一起，引领着通信技术的革新与发展。

随着技术的不断进步与应用的不断拓展，我们有理由相信，手机直连卫星将成为未来通信领域的璀璨明星，为人们的生活带来更加便捷、高效、安全的通信体验。

参考文献：

1. 5G from Space: An Overview of 3GPP Non-Terrestrial Networks，Lin, Xingqin , et al.
2. 手机直连卫星系统发展及关键技术综述，鲍莉娜
3. 传遍朋友圈的“iPhone 13 支持卫星通信”，支持的究竟是什么？慧智微
4. AST SpaceMobile Announces Successful Orbital Launch of Its First Five Commercial Satellites
5. SpaceX unveils “V2 Mini”Starlink satellites with quadruple the capacity
6. Lynk Begins Operation of Next Generation Fifth “Cell Tower in Space”Satellite

在线阅读

本文摘自慧智微微信公众号，扫描以下二维码，在线阅读

5.10 宽带卫星通信市场综述

作者：Microwave Journal

近年来，宽带卫星通信市场快速发展。2022 年 SpaceX 公司宣布，其全球用户数已经超过 100 万。

手机市场也纷纷投入到卫星连接技术中来，在 2021 年，传言 iPhone 13 即支持卫星通信（见文章《传遍朋友圈的“iPhone 13 支持卫星通信”，支持的究竟是什么?》），2022 年 9 月 6 日，华为发布 Mate50 手机，手机即可连接北斗卫星网络发送短报文，小小的智能手机也拥有了“捅破天”技术。

过去 10 年，宽带卫星技术经历了什么样的发展？未来 10 年，卫星通信技术又将走向何处？

在 2023 年 2 月最新的 Microwave Journal 杂志中，欧林工程学院 Whitney 等人发表名为《当前宽带卫星通信市场综述》文章，对卫星市场做了全面综述。原文见文章底部阅读全文链接，以下为此文章的中文翻译（有修改），欢迎留言讨论。

5.10.1 当前宽带卫星通信市场综述

Whitney Q. Lohmeyer, Phillip Post, Katie Fleming, Celvi Lisy, Argyris Kriezis and Abby Omer, Olin College of Engineering, Needham, Mass.

自 2012 年到 2023 年以来，宽带卫星产业的发展可以用前所未有、神秘莫测来形容。

本文对过去 10 年宽带卫星的发展的历史里程碑、系统发展、面临的挑战做了整体梳理。主要关注以下三种卫星系统的发展：

1. 非地球静止轨道（NGSO，Non-geostationary）
2. 低地球轨道星座（LEO，Low Earth Orbit）
3. 地球静止轨道 (GSO，Geostationary)

根据 OSSTP 在 2022 年 4 月的统计，已有超过 20 个独立实体向 FCC 提交了美国卫星通信市场准入申请。这些申请通过使用覆盖 Ku、Ka 以及 V 波段的共计 70,000 颗卫星，提供固定宽带服务，部分实体已经开始进行卫星部署。截至 2022 年 12 月 1 日，OneWeb、Starlink 等公司共有 4,000 多颗非地球静止轨道（NGSO）卫星在轨运行。发射数量与在轨运行数量如下图所示。

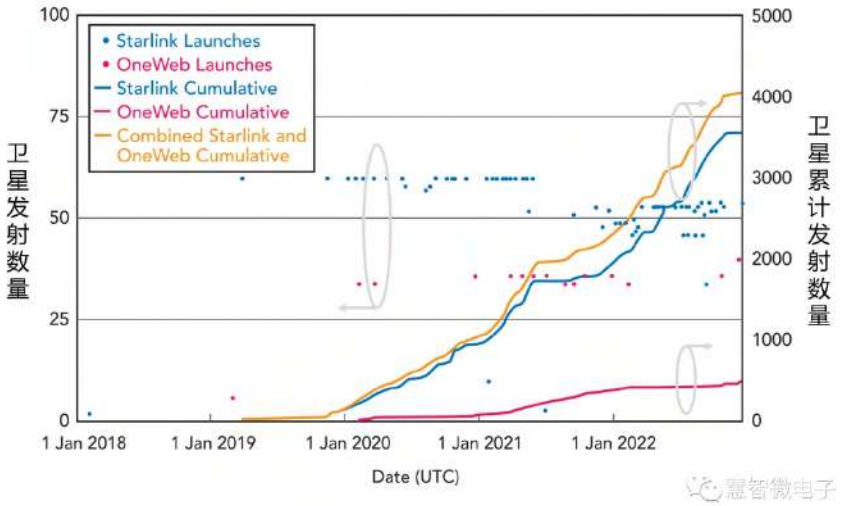


图 5.169: NGSO 固定宽带卫星发射数量（美国）

这些卫星实现技术的技术框架及法律基石主要是在 20 世纪 90 年代末到 21 世纪初建立的，同时在一些早期的系统如 Skybridge、Teledesic 等做了落地。虽然这些早期的卫星网络系统没有在商业上取得成功，但他们却触发了整个卫星行业的发展。这些努力有且于在国际电信联盟（ITU）内推动建立一些技术及规范标准，也有助于

改善 FCC 内处理卫星相关提案的框架。另外，这些网络还推动了终端、网关、传输架构等的低成本方案技术发展，这些都是当前宽带卫星产业蓬勃发展的重要基础。

5.10.2 宽带卫星历史：十年回顾

2012 年 11 月，卫星领域的权威人士、连续创业者雷格 • 惠勒 (Greg Wyler) 在国际电联（ITU）提交了 Ku 波段用于宽带卫星通信的申请，新一代非地球静止轨道卫星（NGSO）系统如 OneWeb、Starlink、Lightspeed、Kuiper 等评估的序幕就此拉开。同一时期，高容量地球静止轨道（GSO）卫星如 Viasat-1 及中轨（MEO）卫星 O3b 也开始提供服务。O3b 通常被描述为最成功的卫星网络，可以提供高容量通信网络，并且运营良好。值得一提的是，O3b 的创始人即为首次向 ITU 提出 Ku 卫星频段申请的 Greg Wyler。



图 5.170: Greg Wyler 的卫星计划时间线

在提出宽带卫星构想后，Wyler 将他的 NGSO 项目带到了 Google，并在 Google 的接入部进行了短期工作。之后，Wyler 与 SpaceX、Tesla 的 CEO Elon Musk 合作完成了一些有关 NGSO 卫星宽带接入的工作，二人在 2015 年分开后，分别造就了 OneWeb 和 Starlink 两大通信卫星星座。

在 2015 年世界通信大会（WRC-15，World Radiocommunication Conference 2015）中，一系列 NGSO 的技术规则被建立起来，成为促成未来 NGSO 革命性发展的技术框架。这些技术框架包括：针对 GSO 系统的保护标准研究（等效通量功率密度限制）、V-Band 演进、NGSO 交付需求、5G 空间/地面频率分配等。

同年，Wyler 与加拿大通信卫星公司 Telesat CEO Dan Goldberg 沟通了 LEO 宽带星座计划。Goldberg 首先在加拿大申请接入 ITU 现有的 COMMStellation 星座，在申请频谱获批后，Goldberg 组建了自己的星座，也就是现在被称为 Lightspeed 的 Telesat NGSO 星座。

自此开始，建设星座开始成为卫星行业的主流，新的初创企业和科技集团开始申请 NGSO 星座。2016 年，SES 通过收购 O3b 在这场比赛中占据一席之地，而亚马逊则宣布计划推出 Kuiper LEO 卫星星座。2017 年，铱星计划也通过由 SpaceX 发射的 NEXT 系列卫星升级了其卫星消息服务系统。如下图所示。

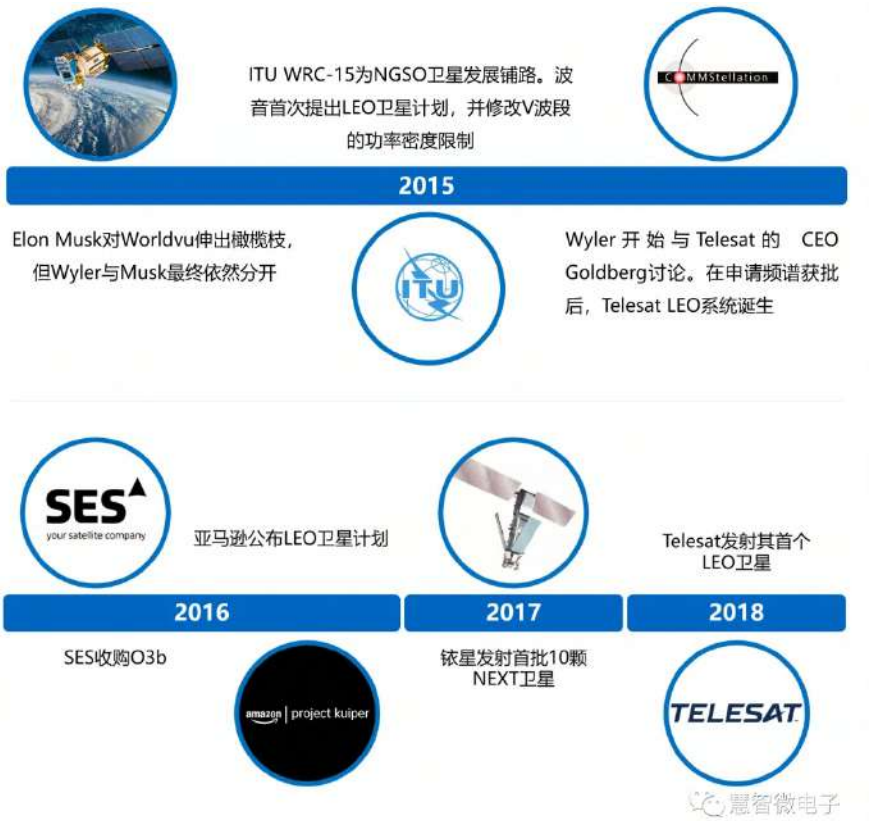


图 5.171: NGSO 卫星主要事件

尽管取得了一些成功，但 LEO 星座卫星的运营并不是没有风险。Greg Wyler 的 OneWeb 是新格局中这个领域最大的参与者之一，但由于资金的不确定性，OneWeb 于 2020 年 3 月宣布破产。在八个月后的 2022 年 11 月，OneWeb 设法找到了新的资金，才从破产边缘走了回来。

在 2020 年 5 月，FCC 进行了 Ka 频段公开申请的受理，共有 10 位 LEO 星座访问的申请人参与。SpaceX 于 2020 年 10 月宣布了其卫星系统的公共版测试，SpaceX 为广大用户推出了快速、低延迟的 LEO 卫星互联网服务。

即使行业依然存在风险，NGSO 卫星产业在快速扩张及发展。

进入 2021 年，快速发展依旧，加拿大为 Telesat 的 Lightspeed NGSO 星座投入 11.5 亿美元。在 2022 年，卫星产业迎来挑战：新型冠状病毒 COVID-19 的大流行、俄-乌战争等几乎影响到了全球人的生活。Telesat 表示他们计划的卫星数目将减少 1/3，OneWeb 也将卫星发射进行了推迟。

尽管挑战存在，但宽带卫星行业还是取得了一些令人瞩目的成就：Starlink 和 OneWeb 达成了频谱协调协议美国 AST SpaceMobile 公司也宣布可以直接用于手机的低速率通信服务（非宽带通信）ViaSat 和 Inmarsat 这两个最大的卫星服务公司之一，也同意以 73 亿美元的协议合并



图 5.172: 卫星市场近期进展

下表总结了主要 LEO，MEO 和 GEO 卫星的任务执行时长，轨道高度，计划卫星数量（以提交申请文件为参考），频率和当前轨道卫星数量。

	Mission Lifetime (Years)	Altitude (km)	Frequency Bands	No. of Planned Satellites	No. of Orbiting Satellites
SpaceX Starlink	5-7	328.3 - 614	Ku Ka E	>30,000	3347[2]
OneWeb	10	1200	Ku Ka	47,844	428[3]
Telesat Lightspeed	10	1000-1325	Ka	198	2
Amazon Kuiper	7	590-630	Ka	3236	0
SES O3b	10	8062	Ka	20	20
SES mPower	12	8062	Ka	11	0
Mangata	10	6400 1215-3800 9000-11,585	Ka V	0	791
ViaSat	15	35,786 (GEO) 8400 1300	Ka V	3	
EchoStar	15	35,786 (GEO)	Ku Ka S	9	
Intelsat	15+	35,786 (GEO)	Ku Ka C	50	
Inmarsat	15	35,786 (GEO)	Ka L S	14	

慧智微电子

图 5.173: 部分 LEO/MEO/GEO 卫星系统统计（至 2022 年 12 月 3 日）

5.10.3 技术和监管挑战

如果没有克服技术和监管方面的挑战 [4]，过去十年 LEO，MEO 和 GEO 宽带卫星行业就无法取得的令人难以置信的成就。这些挑战主要有：

- 低成本方案
- 用户终端
- 高效固态功率放大器

- 发射器

主要的监管挑战集中在：

- 干扰的限制
- 国际电联申请中的部署截止日期
- FCC 对 NGSO 的更新许可规则以及共享框架

5.10.3.1 技术挑战

为了满足 NGSO 系统通信需求，在终端侧，用户终端一般由一对传统抛物面天线，或者一面相控阵构成 [5]。虽然抛物面天线比较便宜，但需要第二天线来支持卫星到卫星之间的切换需求。相控阵天线就可以大大简化这一过程，相控阵天线可以利用多波束扫描来实现多卫星的追踪和切换，但成本要远高于抛物面天线 [6]。Starlink 的地面终端采用的相控阵天线，通过价格补贴，价格还是达到了 599 美元 [7]。不论采用哪种技术，信号的发射与接收必须符合国际电联与 FCC 的规定，比如功率通量密度的限制、地面站增益模板、等效各向同性辐射功率（EIRP）的限制等 [8]。

微波固态功放技术是另一项关键技术。微波固态功放位于卫星有效载荷上。这些功率放大器的半导体材料已从主要基于“硅”基，发展为 GaAs 和 GaN 工艺 [9]。在过去几年中，Akash Systems 等公司正在开发诸如金刚石上 GaN 的解决方案 [10]，可以将功放效率由 30% 提升到 50% 以上，这将有助于显著提升系统整体性能。

业界预计从 2024 年开始，发射需求将激增，这一年有可能出现发射资源紧缺现象。目前可用的四种中型/重型运载火箭分别为：

- SpaceX（美国）：猎鹰 8 号
- ULA（美国）：Atlas V
- Arianespace（法国）/Roscosmos（俄罗斯）：Soyuz-2
- Arianespace（法国）：Ariane

但由于发射器退役和美国限制，以上四个发射器只有一种可用。另外，也有一些小型发射器，如 Rocket Lab 的 Electron 和 Virgin Orbit 的 Launcher One 也有成功将有效载荷送入轨道的经验，预计到 2024 年后，这些发射器每年预计有超过 10 次的发射。

由于新型运载火箭的开发时间很长，所以在 2024 年，能满足一年十次以上发射任务的，可能只有一种中型运载火箭。这将对卫星通信的发展造成极大限制并可能造成重大风险，因为由于发射资源短缺，可能导致发射价格上涨，并有可能造成发射延误，进而阻碍 NGSO 系统阶段性成果的实现。

5.10.3.2 监管挑战

全球卫星系统由联合国通过国际电联 ITU 进行管理。ITU 每三到五年举办一次 WRC（World Radiocommunication Conference，全球通信会议），WRC 也是这三到五年协议审查的高潮期。在两个 WRC 召开之间的几年时间里，工作组（Working Parties，WPs）和研究小组（Study Groups，SGs）合作进行分析，并就频谱分配、许可、干扰管理和未来研究等主题提出建议。参加 WRC 的国家代表包括：

- 国家监管机构（如美国 FCC 和英国的 Ofcom）
- 行业参与者（如 Intelsat、EchoStar、OneWeb、ViaSat、SpaceX 及 Amazon 等）
- 政府实体（如国家电信信息管理局、国家航空航天局等）

在过去十年中，国际和地方监管围绕着多个主题展开。主要的讨论有 [11]：

- 通过限制功率通量密度，保护现有地面网络
- 通过基于时间统计的功率通量密度，缓解地球静止系统的干扰
- 部署时间要求，定义向 ITU 提交发射申请卫星的最晚发射要求
- 5G/6G 技术中，空中/地面频谱分配及协调规则

在美国，频谱的使用是受 FCC（Federal Communications Commission，联邦通信委员会）监管。与地面系统不同，FCC 对卫星频谱分配采取“批次处理”（Processing Round）机制，其处理流程如下图所示（慧智微注）：



图 5.174: FCC 对卫星频谱的“批次处理”（Processing Round）流程

在 2016、2017、2020 和 2021 年，FCC 举行了多次“批次处理”，覆盖了 Ku/Ka 以及 V 波段。在这些处理过程中，有 20 多个申请者提交了市场准入申请，以获得总计 70,000 多颗卫星的授权。自 2019 年底以来，随着 NGSO 中固定卫星系统（FSS）服务的引入，在轨 NGSO 卫星数量增加了 31.4 倍。这种数量的激增也导致 FCC 批准的时间越来越长，以运营商提出申请到 FCC 采取行动计算，2016 年的 Ku/Ka 波段处理时间平均为 2 年，而在 2017 年 V 波段的处理中，时间增长为 2.9 年，增加了近一年时间 [1]。为了解决这一问题，众议院能源和商务委员会于 12 月 8 日在《卫星和电信精简法案》和《安全空间法案》中引入立法，以改革 FCC 的许可规则 [14]。

由于新进入时间不断增加，加之频谱资源稀缺，卫星市场的参与者之间竞争激烈，导致 FCC 也不得不考虑是否要指定新的频谱共享规则 [15]。不过幸运的是，卫星领域的一些主要参与者也达成了协调协议。2022 年 6 月 14 日，OneWeb 和 Starlink 撤回提交给 FCC 针对彼此共存的投诉，并请求 FCC 集中精力尽快对新一轮次的系统进行批复 [16,17]。无独有偶，2022 年 9 月 24 日，亚马逊和 Telesat 也达成合作协议 [18]。

5.10.4 结论

在过去十年，卫星工业取得了令人瞩目的成就。由于 Starlink 和 OneWeb 只有 4,000 多颗卫星在运行，卫星工业在未来几十年可能会继续发展演进，创造历史。

同样，随着卫星网络的不断完善，用户终端、功率放大器，以及发射资源等技术的发展也势在必行。只有这样，才能使整个商业变得健康可行。

另外重要的是，监管机构的法规政策同样重要，需要及时更新过时政策，这样才能鼓励创新，使新的进入者能走进卫星这个行业。

参考文献：

1. Kriezis and W.Q. Lohmeyer, “U.S. Market Access Authorization Timeline Analysis for Megaconstellation Networks,”Olin Satellite + Spectrum Technology & Policy (OSSTP) Group Industry Reports, Web: https://www.osstp.org/fcc-analysis.
2. SpaceX, “SpaceX Launches,” Web: https://www .spacex.com/launches/.
3. “OneWeb Launch Programme,” Web: https://oneweb. net/resources /launch-programme.
4. R. Barry, P. Bustamante, D. Guo, M. Honig, W.Q. Lohmeyer, I. Murtazashvili, S. Palo, “Spectrum Rights in Outer Space: Interference Management for Mega-constellations,” Telecommunications Policy Research Conference (TPRC), Washington, D.C., September 16-17, 2022.
5. U. Gupta., A. Tan, J. Liu, W.Q. Lohmeyer, “A Description and Comparison of Modern Flat Panel Antenna Technology for Ku-/Ka-Band User Terminals in Low Earth Orbit (LEO) Satellite Communications Systems,” Microwave

Journal, Vol. 69, No. 9, September 2021.

6. U. Gupta, A. Tan, W.Q. Lohmeyer, J. Liu, “Antenna Performance Specifications for Ka-/Ku-Band User Terminals in Low Earth Orbit (LEO) Satellite Communications Systems,” Microwave Journal, Vol. 69, No. 9, September 2021.
7. M. Kan, “SpaceX has been selling Starlink Dishes at a Huge Loss,” PC Mag, April 6, 2021, Web: https://www. pc-mag.com/ news/spacex-has-been-selling-starlink -dishes-at-a-huge-loss-despite-499-price.a) Tan and W.Q., Lohmeyer, “Modern Low-Cost Phased Array Technologies and Accompanying Fixed Satellite Service (FSS) Regulatory Requirements,” IEEE International Symposium on Phased Array Systems and Technology, Waltham, Mass., October 11-14, 2022.
8. W.Q. Lohmeyer, R.J. Aniceto and K.L. Cahoy, “Communication Satellite Power Amplifiers: Current and Future SSPA and TWTA Technologies,”International Journal of Satellite Communications and Networking, doi:10.1002/sat.1098.
9. Akash Systems, “A Fundamental Advantage in Technology: The Power of Synthetic Diamond,”Web: https://akashsystems.com/te
10. J. Pahl, Interference Analysis, Wiley, 2016.
11. “Federal Communications Commission,” USAGov, U.S. Government,
12. “Amendment of the Commission’ s Space Station Licensing Rules and Policies,” 18 FCC Rcd. 10760, 2003.
13. “House Committee Leaders Introduce Bipartisan Bills to Update Satellite Rules,” Space News, December 9, 2022, Web: https://spacenews. com/house -committee- leaders -introduce- bipartisan-bills-to -update- satellite-rules/.
14. “Expediting Initial Processing of Satellite and Earth Station Applications; Space Innovation,” IB Docket Nos. 22-411, 22-271, November 30, 2022.
15. “Request for Orbital Deployment and Operating Authority for the SpaceX Gen2 NGSO Satellite System,” IBFS File Nos. SAT-LOA-20200526-00055 and SAT-AMD-20210818-00105 December 1, 2022.
16. “Starlink and OneWeb Reach Spectrum Coordination Plan,”Space News, June 14, 2022, Web: https://spacenews.com/starlink-and-oneweb-reach-spectrum-coordination-plan/.
17. “Amazon and Telesat Coordinate their Planned NGSO Constellations,” Space News, September 21, 2022, Web: https://spacenews. com/amazon -and-telesat-coordinate -their-planned -ngso-constellations/.

在线阅读

本文摘自慧智微微信公众号，扫描以下二维码，在线阅读

5.11 “有线”与“无线”通信，有什么异同？

作者：彭洋洋 博士

作为一名射频领域的从业者，自读书的时候起，我们好像便不自觉地将“通信”与“无线通信”紧密地联系在一起。在我们的学习历程中，“通信”一词仿佛自然而然地与“无线通信”划上了等号，而提及“有线通信”时，则需特意加以说明。



图 5.176: 国防工业出版社出版的有关通信的书籍

在过去几十年的通信技术发展中，移动通信迅猛崛起，也将学术界与产业界的目光紧紧聚焦在了手机上。作为移动通信的杰出代表，手机中广泛应用的蜂窝网络、Wi-Fi、蓝牙、NFC、UWB 以及卫星通信等技术，都是依托射频实现的“无线”通信。这一现象不仅彰显了无线通信技术的强大能力，也加强了其在通信领域的代表性地位。

慧智微起初深耕于手机应用的射频芯片，所以在日常交流中自然而然地聚焦于“无线”通信技术。然而，近年来，随着路由器、GPON 厂商对慧智微 Wi-Fi6/Wi-Fi7 FEM 产品的广泛采用，我们逐渐感受到，通信领域并非无线通信的单一舞台，有线通信同样无可替代。路由器、GPON 厂商的产品核心功能在于将“有线”的光信号或电信号转换为“无线”的射频信号。在与这些客户专家及工程师的深入交流中，我们也逐渐感受到有线通信与无线通信思考角度的不同。

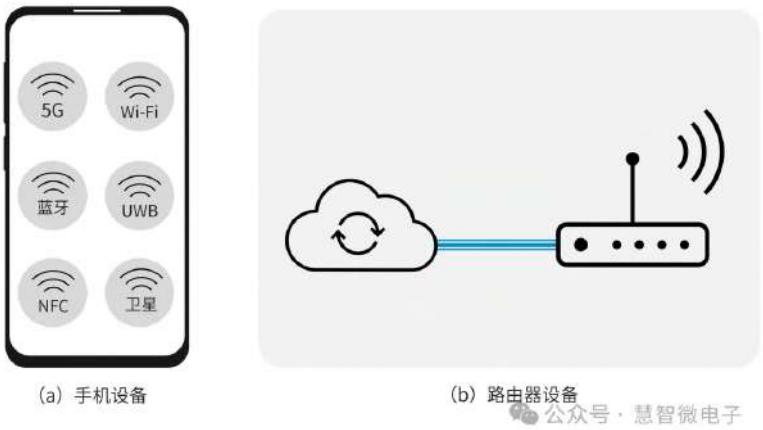


图 5.177: 手机是无线协议的集合；路由器设备是有线通信到无线通信的界面

在与专家和工程师们的深入交流中，我们感受到了有线通信的技术特点与发展脉络，也体会到区分通信中的“有线”与“无线”绝非咬文嚼字这么简单。在一种通信技术中的创新，可能是基于其独特的属性应运而生。

比如因为无线通信中频谱的宝贵，所以对高阶调制技术有了极致探索（比如 Wi-Fi7 中的 4096QAM），另外，频率资源的珍贵也使得无线通信对多频段聚合和 MIMO 技术的不断精进。

有线与无线是实现通信的两种方式，有时在一种通信技术中遇到的难题，换用另一种通信技术却能迎刃而解。通过对比这两种技术，我们能够更加清晰地洞察每种技术的独特优势、适用场景及应用边界。

因此，本文旨在通过对比有线通信与无线通信的技术特点，探讨并增进对这两种通信技术的理解。

5.11.1 通信技术的发展历程

5.11.1.1 通信技术的发展概述

通信技术的诞生与发展，源自于人类对沟通的渴望与追求。

自古以来，人类便不断探索各种通信方式，从最初的语言交流、文字记载，到后来的飞鸽传书、烽火狼烟，无不彰显着人类突破时空限制、追求自由通信的梦想。而现代通信技术的飞速发展，则让这一梦想成为了现实。

有线通信技术，作为通信技术的重要分支，自其诞生以来便承载着人类对于高效、稳定通信的期望。

随着时间的推移，有线通信技术的媒介经历了多次技术革新，从最初的铜线，到后来的同轴电缆，再到如今广泛普及的光纤，每一次迭代都伴随着通信速率的显著提升。

随着人类对通信速率要求的不断提高，有线通信技术也在不断演进，最新的光纤通信技术已经能够实现高达 800Gb/s 以上的传输速率，为人们的日常生活和工作带来了前所未有的便捷与高效。



图 5.178: 有线通信的应用

相比于有线通信技术，无线通信技术的诞生相对较晚。在 1831 年，法拉第发现电磁感应现象，这一里程碑式的发现为人类打开了无线通信的大门，让人们开始憧憬一种更加自由、便捷的通信方式。随后，得益于麦克斯韦、赫兹、马可尼等杰出科学家与工程师的不懈努力，无线通信逐渐从理论走向现实。

自 20 世纪 90 年代起，随着蜂窝移动通信技术的迅猛发展，无线通信也迎来了发展的黄金时期。科学家们和工程师们不断创新，研发出众多新技术，将无线通信的性能推向了物理极限。时至今日，5G 及 Wi-Fi7 网络已经能够提供高达 40Gb/s 的通信速率。



图 5.179: 无线通信的应用

在发展过程中，关于哪种技术将成为未来发展主流，以及哪种技术将取代另一种技术的讨论屡见不鲜。然而，随着技术的不断进步，人们逐渐认识到，有线与无线通信技术各有其独特的优势和局限。

有线技术，如光纤，能够提供极高的带宽和卓越的传输质量，一旦部署完成，数据传输成本极低，但其初始部署成本却相对较高。

相比之下，无线宽带技术，如 Wi-Fi、5G 等，则为用户带来了更加灵活便捷的接入方式。然而，由于无线通信频谱的共享特性，当多用户同时接入时，传输速率可能会受到影响，且多频率间的干扰问题也不容忽视。

人类追求的是快速、低成本、无处不在的通信接入服务。在选择宽带接入技术时，需要综合考虑带宽需求、成本、部署难度、覆盖范围等多个因素。而不同地区、不同场景下的技术选择和成本、物理限制、技术约束均存在差异，因此，没有哪一种通信技术能够独占市场。

未来，随着物联网、大数据和人工智能等新兴技术的蓬勃发展，通信技术将面临更多的挑战和机遇。有线和无线两种通信方式将相互补充，共同快速发展，以满足人类日益增长的通信需求。

5.11.1.2 有线通信的发展

有线通信的研究与发展历史可追溯至 200 年前的 19 世纪初。1835 年，塞缪尔 • F • B • 莫尔斯（Samuel F. B. Morse）成功研制出第一台有线电报机，标志着有线电报的诞生。相比之下，无线通信的首次应用则要到 1900 年前后，有线通信的发展因此领先了近一个世纪。



图 5.180: 莫尔斯 1835 年发明的有线电报机

有线电报发明后，迅速被应用于军事、民用等多个领域。1866 年，横跨大西洋的海底电缆成功铺设，将欧美大陆紧密相连，有线电报通信因此成为全球性的通信方式。

紧接着，在 1876 年，亚历山大 • 格拉汉姆 • 贝尔发明了世界上第一部电话，标志着电话技术的诞生。经过近百年的蓬勃发展，到 20 世纪中期，电话线路已在全球范围内普及，成为人们生活中必不可少的通信工具。



图 5.181: 20 世界中期的电话广告，当时电话已进入寻常百姓家

从 1835 年至 1970 年左右的这一百多年间，利用电报、电话的承载，人类逐步发现并在全球范围内应用了“有线通信”这一通信方式。而 1970 年之后，随着互联网的飞速发展，有线通信的发展踏上了高速车道。

1969 年，美国国防部国防高级研究计划署（DARPA）资助建立了名为 ARPANET（阿帕网）的网络。ARPANET 首次实现了计算机之间的互联，通过通信线路将不同地点的计算机主机连接起来。它被誉为世界上第一个计算机网络，也是今天因特网的前身，标志着计算机互联通信迈入了一个崭新的时代。

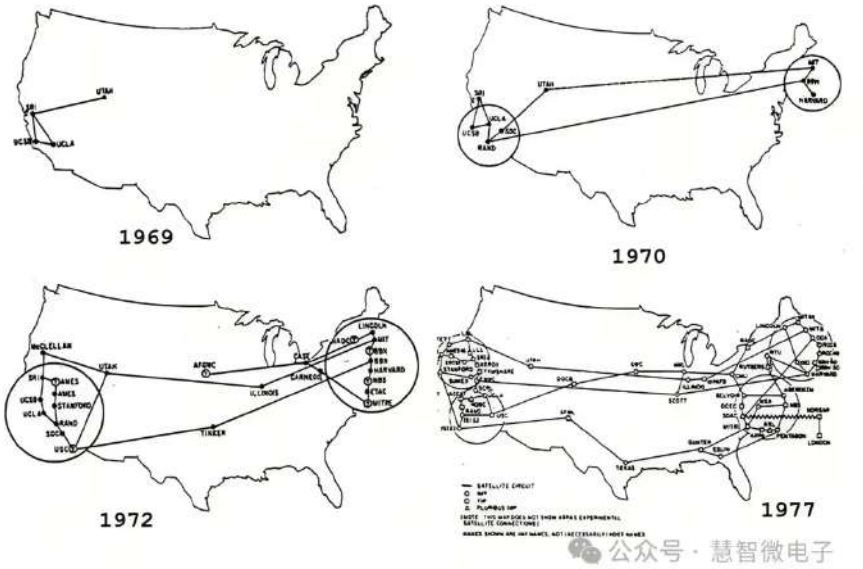


图 5.182: 最早的有线计算机网络 ARPANET 的诞生与发展

计算机互联需要依赖传输媒介，在各种计算机高速互联需求的推动下，有线传输媒介开始迅速迭代发展。在过去的五十年里，随着有线通信速率需求的不断提升，以下几种技术方案得到了广泛应用：

1. 基于有线电话线的拨号上网技术；
2. 基于同轴电缆的有线电视网络技术；
3. 基于以太网线的互联网网络技术；
4. 基于光纤介质的光纤通信技术。

- **互联网初生：基于有线电话线的拨号网络**

1969 年，互联网概念初露端倪之时，全球有线电话网络已发展得相当完善。在大多数大型城市，电话线路已广泛覆盖，国家间也建立了国际长途电话线路和交换中心，实现了跨境通信。同时，电话通信技术日益成熟，不仅语音通信清晰，还具备了基本的呼叫控制和管理功能。

正是得益于有线电话网络的硬件基础，在互联网开始普及的阶段，电话线因其广泛的分布和较高的可靠性，成为了实现信息互联的理想载体。利用已广泛铺设的电话线网络进行互联网部署，可以显著降低互联网建设的经济成本。于是，基于电话线的有线通信互联网逐步发展起来。

早期，通过有线电话线进行有线网络连接时，数据需先转化为模拟信号进行传输，这一过程称为“调制”。数据到达接收端后，模拟信号再被还原为原始数据，这一过程称为“解调”。调制和解调的过程合称为 Modem（中文译为“调制解调器”），也常被人们亲切地称为“猫”。

在 Windows 早期的拨号界面中，还可以清楚的看到有线电话线实现的拨号网络拓扑结构。

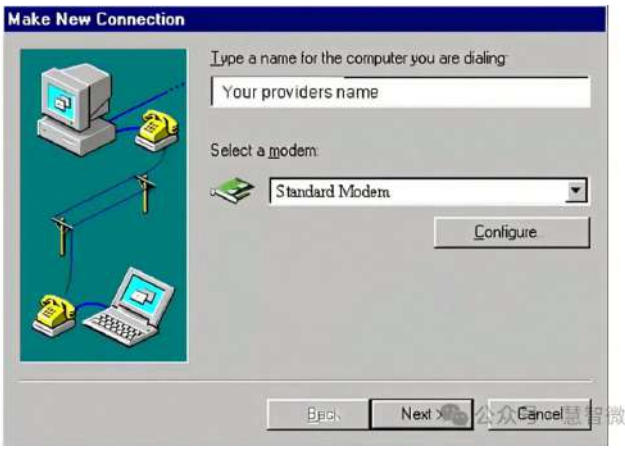


图 5.183: 早期 Windows 电脑拨号上网界面，及有线拨号网络拓扑

早期的电话线原本是为传递声音而设计的，因此在进行数据传输时，信号也被当作一种类似声音的信号来处理。拨号、传输等整个过程，都与人们打电话的方式极为相似，只不过这时的通话对象从人变成了“猫”。有过那个时期上网经历的朋友，对那段拨号上网的声音一定还记忆犹新。

然而，拨号上网由于占用的是普通电话线，因此也存在一些明显的局限：

1. 上网和打电话不能同时进行，一旦上网，电话线就被“占线”了。
2. 传输速率慢，最高仅能达到 56Kb/s，仅能满足基本的文本浏览需求。
3. 稳定性不够理想，容易受到电话线质量的影响。

尽管现在看来，有线电话线在通信方面存在诸多不便，但在当时，它确实为每一个使用者打开了连接世界的大门。一根电话线，一个“猫”，就让每个人与世界紧密相连，全球互联网时代就此拉开序幕。



图 5.184: 早期拨号上网“猫”的广告

- **探索新介质：基于同轴线的有线电视网络**

  在电话线实现互联网连接之后，人类深切体会到了全球互联给社会和生活带来的巨大变革。

  随后，人们不再满足于仅传输文字，对图像、视频的传输需求日益凸显。此时，基于电话线的网络已无法满足这些更高的传输需求，于是人类开始寻找新的通信介质。

  有线电视网络自 1970 年起迅速爆发和普及。随着全球经济的复苏，人们对生活品质的要求不断提高，对新闻、娱乐、教育等信息的获取需求也日益增长。有线电视网络通过同轴电缆将电视节目传输到千家万户，让人们在家中就能享受丰富多彩的娱乐内容，极大地满足了人们的生活需求。随着电视技术的不断进步和成本的降低，有线电视迅速普及，有线电视网络也成为了除电话线路之外，另一根深入千家万户的网络线路。



图 5.185: 有线电视是寻常人家的第二个有线网络

有线电视网络在设计之初便兼顾了宽带信号传输的需求，以应对视频传输的挑战。它采用的是同轴电缆，这种电缆具有较低的信号衰减特性，使得信号能够传输得更远。此外，同轴电缆的屏蔽结构有效抵御了外

部干扰，确保了信号传输的稳定性。



图 5.186: 用于连接有线电视的同轴电缆

随着人类对互联网网络速率要求的不断提升，电话线的拨号上网方式已难以满足需求。在此背景下，有线电视网络凭借其优势，成为了新的有线通信载体之选。基于有线电视网络，可以实现数十 Mb/s 乃至数百 Mb/s 的高速信号传输，为用户带来了前所未有的上网体验。

当通过有线电视网络上网时，信号同样需要经过调制与解调的过程，以适应有线电视网络的传输特性。值得一提的是，调制后的信号并不会占用有线电视原有的带宽，因此用户在观看电视节目时不会受到任何影响，对“猫”的存在几乎毫无察觉。

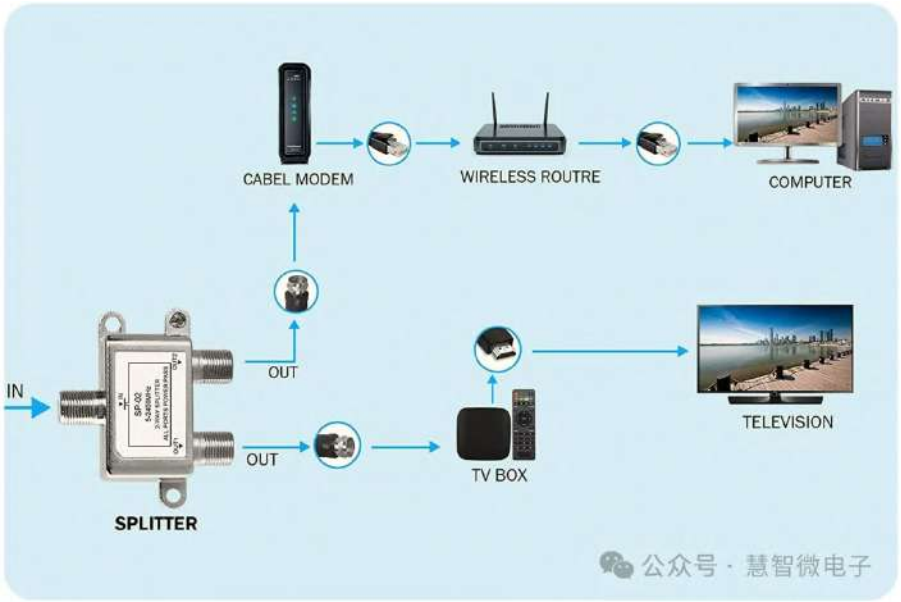


图 5.187: 有线电视宽带网络架构示意图

有线电视网络作为有线通信的一种方式，被使用了相当长的时间，并且至今仍然广泛应用于众多家庭。在许多新建小区，当光纤等其他高速有线网络尚未部署时，有线电视服务商提供的宽带服务往往成为了居民们的最佳选择。



图 5.188: 广电集团早期推出的基于有线电视的宽带服务

- **电话线的飞跃：铜质双绞线的升级**

  随着有线电视网络在宽带互联领域的广泛应用，有线电话线产业也迎来了改造升级的新篇章。

  1989 年，贝尔实验室研发出 DSL（数字用户线路）技术，让普通电话线也能高速传输数据。随着互联网的蓬勃发展，DSL 技术迅速普及，成为家庭和小型企业接入互联网的首选之一。

  DSL 技术巧妙利用现有电话线路，通过数字调制与解调，在不影响语音通话的同时，提供高达几十甚至几百 Mb/s 的通信速率。

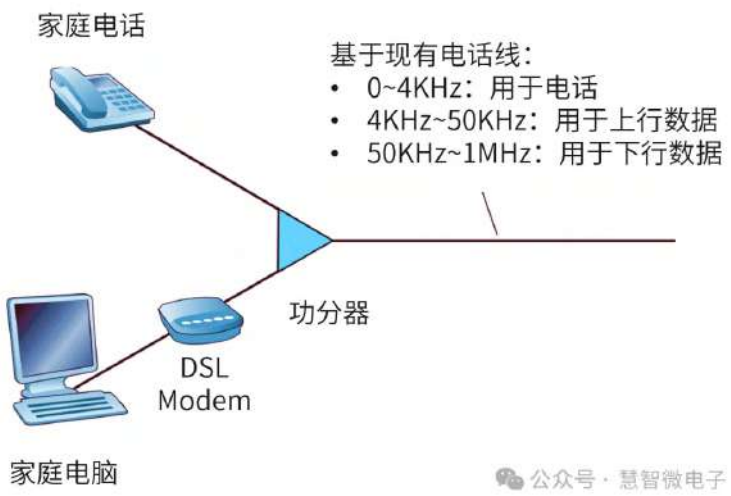
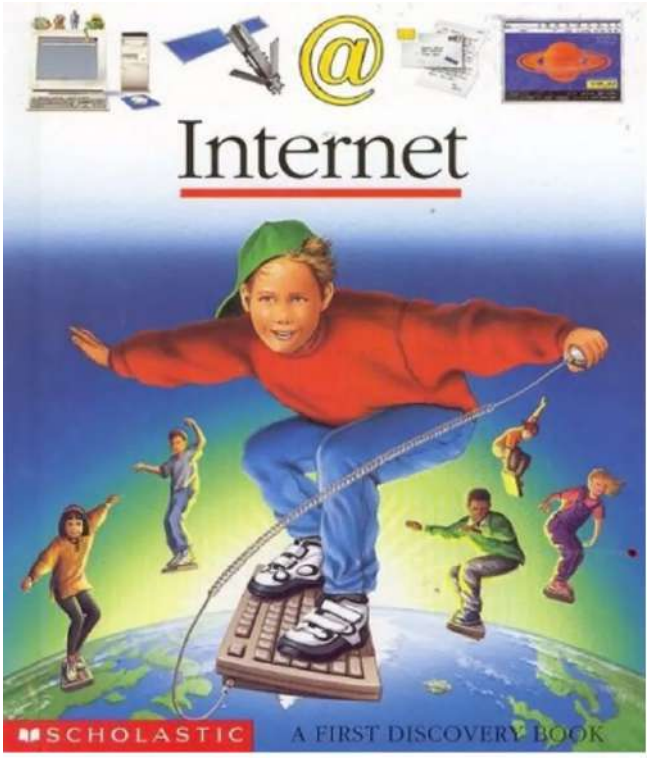


图 5.189: DSL 技术对于原有电话线的改进

借助 DSL 技术，用户彻底告别了繁琐的拨号过程，网络保持持续在线状态，从此无需在上网和打电话之间做出选择，用户体验得到了显著提升。此刻的上网体验，真正让人感受到了“冲浪”般的畅快。



公众号 · 慧智微电子

图 5.190: DSL 技术发明后，一些杂志对“上网冲浪”做的形象化宣传

为了满足用户不同的上传和下载需求，以及家庭用户与企业用户之间数据传输的差异，DSL 技术不断演进，衍生出了 ADSL（非对称数字用户线）、HDSL（高速数字用户线）、VDSL（甚高速数字用户线）等多种类型。每种类型都针对特定的应用场景，发挥着各自的优势。



公众号 · 慧智微电子

图 5.191: 中国电信在 2002 年推出的 ADSL 上网卡

值得一提的是，DSL 技术依然沿用着与电话线相同的铜质线缆。为了追求更高的传输速率，线缆从单线逐渐发展双绞线。在当今的铜质线缆传输中，双绞线结构已成为主流。

双绞线，由两根绝缘铜导线相互绞合而成，这种独特结构使得外部有干扰磁场时，感应电流可以自己对消，有效减少了电磁干扰（EMI）和串扰，显著提升了数据传输的稳定性和可靠性。

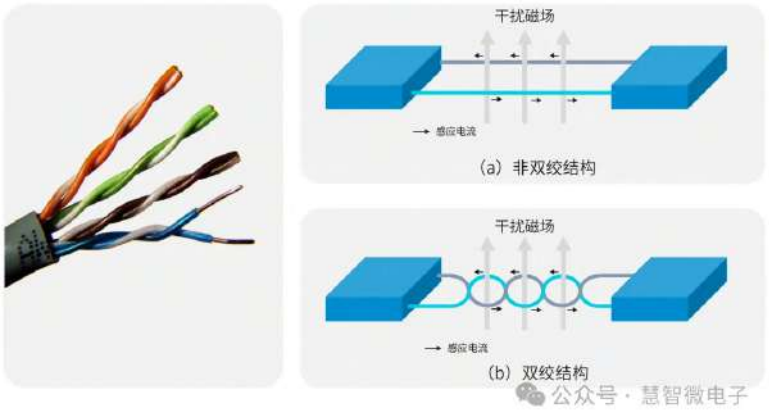


图 5.192: 双绞线结构及其技术原理

双绞线在高速数据传输领域的发展历程，经过从模拟到数字、从低速到高速、从简单到复杂的演进。

- 早期，一类线至四类线（CAT1 至 CAT4）主宰着市场，它们主要用于语音传输，在 1990 年之前广泛应用于各个领域。
- 进入 2000 年之前，五类线（CAT5）诞生，支持 100MHz 传输信号带宽，实现了 100Mb/s 的传输速率，成为数字信号有线传输的重要里程碑。
- 2001 年，超五类线（CAT5e）标准应运而生，最高支持 1Gb/s 的传输速率，至今仍是应用最广泛的网线之一。
- 随后，六类线（CAT6）将最大带宽提升至 250MHz，并加入十字形骨架减少串扰，达到了 1Gb/s 的稳定传输速率。
- 超六类线（CAT6a）更是将最大带宽扩大到 500MHz，稳定支持 10Gb/s 的传输速率。
- 七类线（CAT7）的带宽增加至 600MHz，稳定支持 10Gb/s 的传输速率，在特定环境下甚至能支持到 40Gb/s 的速率。
- 而最新的八类线（CAT8），带宽更是飙升至 2000MHz，稳定实现 40Gb/s 甚至更高的速率传输，引领着未来数据传输的新潮流。

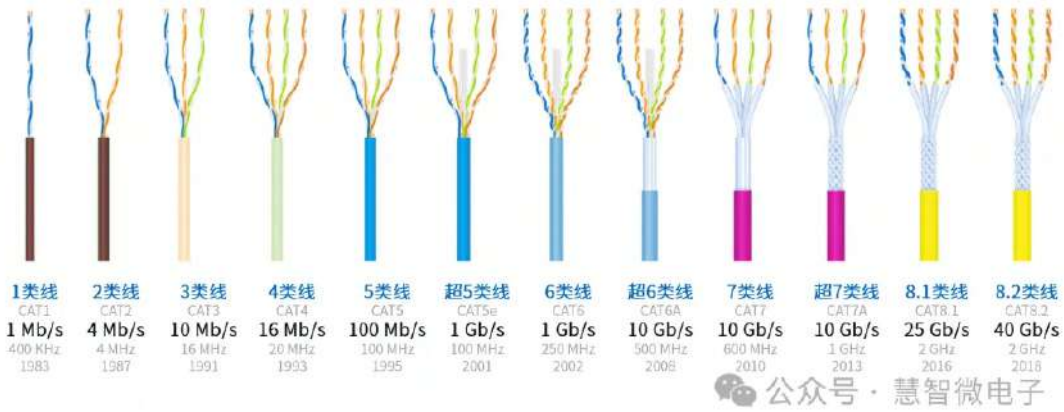


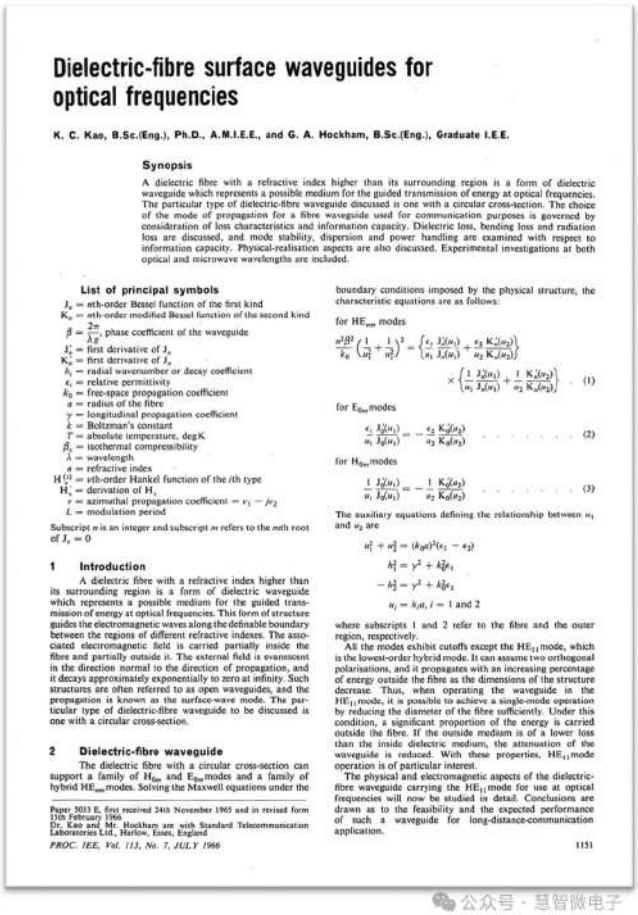
图 5.193: 不同双绞线类型及速率

- **全新网络时代：光纤通信的崛起**

如果说 ADSL 网络、有线电视宽带网络是在原有电话、电视网络基础上的升级，那么光纤通信网络则是为全球宽带有线通信量身打造的全新网络体系。

1966 年，华裔科学家高锟发表了一篇里程碑式的论文《光频率介质纤维表面波导》，首次阐述了光导纤维在通信领域应用的基本原理。在那个时代，光纤的传播损耗高达每公里 1000dB，但高锟教授却预言，光导

纤维的损耗可以降低至每公里 20dB 以下，从而实现通信的飞跃。



Dielectric-fibre surface waveguides for optical frequencies

K. C. Kao, B.Sc.(Eng.), Ph.D., A.M.I.E.E., and G. A. Hockham, B.Sc.(Eng.), Graduate I.E.E.

Synopsis

A dielectric fibre with a refractive index higher than its surrounding region is a form of dielectric waveguide which represents a possible medium for the guided transmission of energy at optical frequencies. The particular type of dielectric-fibre waveguide discussed is one with a circular cross-section. The choice of the mode of propagation for a fibre waveguide used for communication purposes is governed by consideration of loss characteristics and information capacity. Dielectric loss, bending loss and radiation loss are discussed, and mode stability, dispersion and power handling are examined with respect to information capacity. Physical-realisation aspects are also discussed. Experimental investigations at both optical and microwave wavelengths are included.

List of principal symbols

J_ν = νth-order Bessel function of the first kind
K_ν = νth-order modified Bessel function of the second kind
$\beta = \frac{2\pi}{\lambda_g}$, phase coefficient of the waveguide
J'_ν = first derivative of J_ν
K'_ν = first derivative of J_ν
h = radial wavenumber or decay coefficient
ϵ_r = relative permittivity
k_0 = free-space propagation coefficient
a = radius of the fibre
γ = longitudinal propagation coefficient
k = Boltzman's constant
T = absolute temperature, degK
β_c = isothermal compressibility
λ = wavelength
n = refractive index
$H_\nu^{(1)}$ = νth-order Hankel function of the ith type
H'_ν = derivation of H_ν
ν = azimuthal propagation coefficient $= \nu_1 - j\nu_2$
L = modulation period

Subscript m is an integer and subscript m refers to the mth root of $J_\nu = 0$

1 Introduction

A dielectric fibre with a refractive index higher than its surrounding region is a form of dielectric waveguide which represents a possible medium for the guided transmission of energy at optical frequencies. This form of structure guides the electromagnetic waves along the definable boundary between the regions of different refractive indexes. The associated electromagnetic field is carried partially inside the fibre and partially outside it. The external field is evanescent in the direction normal to the direction of propagation, and it decays approximately exponentially to zero at infinity. Such structures are often referred to as open waveguides, and the propagation is known as the surface-wave mode. The particular type of dielectric-fibre waveguide to be discussed is one with a circular cross-section.

2 Dielectric-fibre waveguide

The dielectric fibre with a circular cross-section can support a family of H_{0m} and E_{0m} modes and a family of hybrid $HE_{\nu m}$ modes. Solving the Maxwell equations under the boundary conditions imposed by the physical structure, the characteristic equations are as follows:

for $HE_{\nu m}$ modes

$$\frac{a^2\beta^2}{k_0}\left(\frac{1}{u_1^2}+\frac{1}{w_2^2}\right)^2 = \left\{\frac{\epsilon_1}{u_1}\frac{J'_\nu(u_1)}{J_\nu(u_1)}+\frac{\epsilon_2}{u_2}\frac{K'_\nu(u_2)}{K_\nu(u_2)}\right\} \times\left\{\frac{1}{u_1}\frac{J'_\nu(u_1)}{J_\nu(u_1)}+\frac{1}{u_2}\frac{K'_\nu(u_2)}{K_\nu(u_2)}\right\} \quad (1)$$

for E_{0m} modes

$$\frac{\epsilon_1}{u_1}\frac{J'_0(u_1)}{J_0(u_1)} = -\frac{\epsilon_2}{u_2}\frac{K'_0(u_2)}{K_0(u_2)} \quad (2)$$

for H_{0m} modes

$$\frac{1}{u_1}\frac{J'_0(u_1)}{J_0(u_1)} = -\frac{1}{u_2}\frac{K'_0(u_2)}{K_0(u_2)} \quad (3)$$

The auxiliary equations defining the relationship between u_1 and u_2 are

$$u_1^2 + u_2^2 = (k_0a)^2(\epsilon_1 - \epsilon_2)$$
$$h_1^2 = \gamma^2 + k_0^2\epsilon_1$$
$$-h_2^2 = \gamma^2 + k_0^2\epsilon_2$$
$$u_i = h_ia,\ i = 1 \text{ and } 2$$

where subscripts 1 and 2 refer to the fibre and the outer region, respectively.

All the modes exhibit cutoffs except the HE_{11} mode, which is the lowest-order hybrid mode. It can assume two orthogonal polarisations, and it propagates with an increasing percentage of energy outside the fibre as the dimensions of the structure decrease. Thus, when operating the waveguide in the HE_{11} mode, it is possible to achieve a single-mode operation by reducing the diameter of the fibre sufficiently. Under this condition, a significant proportion of the energy is carried outside the fibre. If the outside medium is of a lower loss than the inside dielectric medium, the attenuation of the waveguide is reduced. With these properties, HE_{11} mode operation is of particular interest.

The physical and electromagnetic aspects of the dielectric-fibre waveguide carrying the HE_{11} mode for use at optical frequencies will now be studied in detail. Conclusions are drawn as to the feasibility and the expected performance of such a waveguide for long-distance-communication application.

Paper 5033 E, first received 24th November 1965 and in revised form 15th February 1966
Dr. Kao and Mr. Hockham are with Standard Telecommunication Laboratories Ltd., Harlow, Essex, England

公众号 · 慧智微电子

图 5.194: 高锟教授 1966 年所发表的论文

高锟教授的这一理论为光纤通信的发展指明了方向，并因此改变了全人类的通信技术。2009 年，他因此项荣获诺贝尔奖，也成为了通信界的传奇人物。

高锟教授 1933 年出生于上海金山，后在香港、英国求学深造。1964 年，在英国求学期间，他提出了一个革命性的设想：用光代替电流，用玻璃纤维代替导线来构建电话网络。这一想法在当时被视为天方夜谭，甚至遭人嘲笑。然而，高锟教授并未放弃，他坚信自己的理念，经过多年的不懈努力，光纤技术终于成为全球范围内的主流技术，他也因此被誉为“光纤之父”。

除了科学家的身份，高锟还是一位杰出的教育家。1970 年，他加入香港中文大学，参与筹办电子学系。1987 年至 1996 年间，他更是担任了香港中文大学的校长。然而，2003 年，高锟教授被确诊患有阿兹海默病，记忆力和表达能力逐渐衰退。即便如此，2009 年当他荣获诺贝尔奖时，依然通过视频发表了感人的获奖感言，展现了他的坚韧与不屈，也让人为之动容。



图 5.195: 高锟教授（右）及夫人展示诺贝尔奖章

1966 年，高锟教授奠定了光纤传输的理论基础，随后产业界积极投身研究，不断取得技术突破。

- 1970 年，康宁公司成功研制出传输损耗每千米仅 20dB 的石英光纤，这一成果有力证明了光纤作为通信介质的巨大潜力。
- 1972 年，随着光纤制备工艺中原材料提纯、制棒和检验技术的持续提升，光纤的传输损耗进一步降低至每公里 4dB，标志着光纤通信时代的正式开启。
- 1976 年，美国西屋电信公司在亚特兰大成功进行了世界上首个 44.736Mb/s、传输距离达 110 公里的光纤通信系统试验，这一里程碑式的成就使光纤通信向实用化迈出了坚实的一步。
- 同年，日本电报电话公司也将光纤损耗降低到了 0.5dB/km，而现代通信光纤的损耗已经可以进一步降低到 0.17dB/km 至 0.2dB/km 的极低水平。

1980 年代，光纤通信迈入了大规模商用化的崭新阶段。得益于光纤制备工艺的持续改进，光纤的传输损耗不断降低，传输距离和带宽则显著提升。与此同时，光放大器、光开关、光复用器等各类光纤通信设备和系统应运而生，为光纤通信的广泛应用奠定了坚实基础。

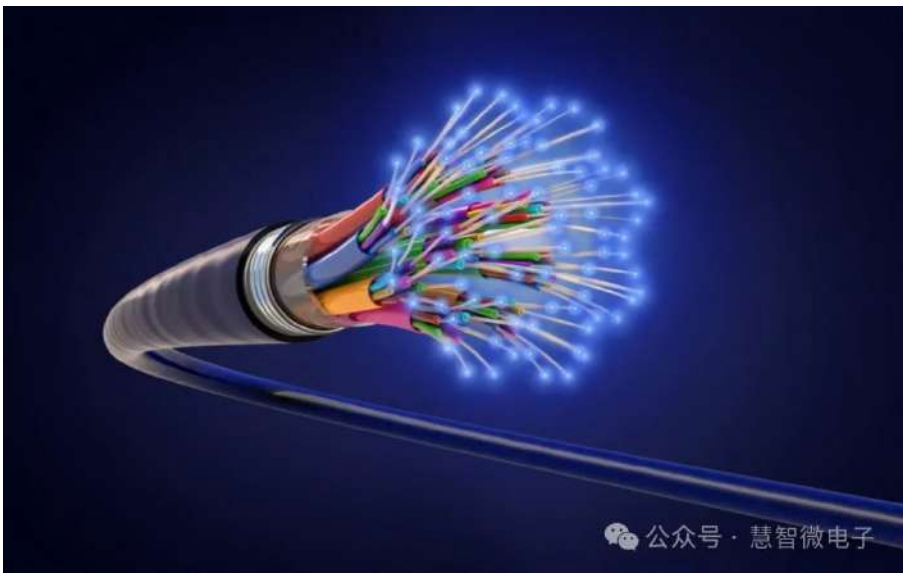


图 5.196: 用于有线通信的光纤

历经近半个世纪的发展，光纤网络产业和技术已日趋成熟。如今，商用光纤通信系统的传输速率已高达

100Gb/s、400Gb/s，甚至更高，在互联网、电视、电话等多个领域得到了广泛应用。随着技术的不断革新和应用需求的持续增长，预计未来几年内，光纤通信系统的传输速率将跃升至 Tb/s 级别。

相较于传统铜线网络，光纤网络展现出诸多优势：传输速率更快（从几十 Gb/s 提升至几百 Gb/s 甚至更高）、传播距离更远（从几百米延伸至数百公里）、损耗更低，且抗干扰性更强。因此，光纤网络已成为长距离通信、骨干网通信、数据中心等大容量通信传输中的核心支撑技术。

近年来，AI 人工智能技术得到了大规模推广与应用，推动了数据规模的持续膨胀。AI 模型在训练、推理等环节中需处理海量数据，这对数据传输的速度和带宽提出了前所未有的高要求。光纤技术，凭借其高速、大容量的独特优势，成为了支撑 AI 技术发展的核心基础设施。

在数据中心互联（DCI）领域，AI 等应用的核心诉求之一是实现数据中心间的高速数据传输，以满足云计算、大数据、人工智能等对高带宽、低延迟网络的迫切需求。为应对这一挑战，光纤有线通信中的光模块技术不断迭代升级，800G、1.6T 等超高速光模块已逐步应用于 DCI 场景，为数据传输提供了强有力的支撑。

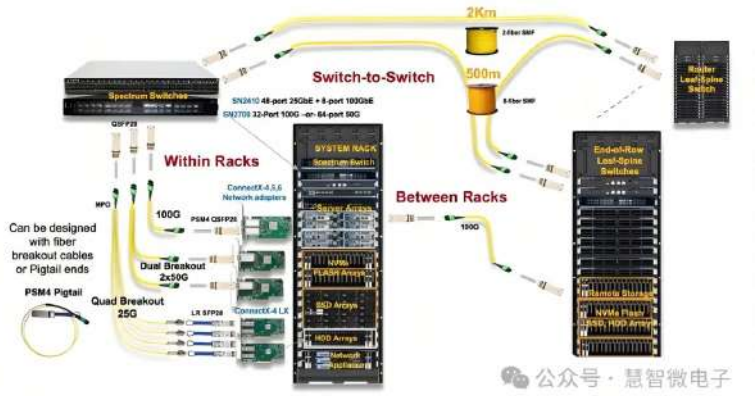


图 5.197: 数据中心内部用于数据的传输和交换的有线光纤网络

AI 技术的兴起与爆发，为有线通信中的光纤技术带来了新的发展契机与挑战。光纤技术持续创新与发展，以满足 AI 技术对数据传输的更高标准。同时，光纤技术也为 AI 技术的蓬勃发展提供了坚实保障。AI 应用与有线光通信相辅相成，共同驱动着整个社会的数字化转型与智能化升级。

5.11.1.3 无线通信的发展

如果说有线通信是人类对于网速无限追求所绘就的鸿篇巨制，那么无线通信就是人类挣脱有线束缚的灵动之舞。

无论有线通信的速率如何日新月异，它始终无法摆脱一个根本性的局限：需要物理介质进行传输。这意味着，人类必须先铺设好“有线”的网络，才能实现通信。这无疑大大限制了人类的自由活动空间。

正是人类对通信“自由”的不懈追求，催生了无线通信技术的迅猛发展。

1831 年，英国科学家法拉第首次揭示了电磁感应现象，这一重大发现不仅为电磁理论的发展奠定了坚实基础，更为无线通信技术的崛起埋下了伏笔。1964 年，麦克斯韦完善了电磁理论。1887 年，德国科学家赫兹通过严谨的实验，证实了电磁波的真实存在。紧接着，1899 马可尼也实现了无线电报的传输，这些里程碑式的成就共同标志着无线通信技术的正式诞生。



图 5.198: 无线通信的重要发展历程碑事件

随着无线电通信技术的日益成熟，其应用范围也迅速拓展。20 世纪初，无线电报在海上通信中崭露头角，极大地提升了航海的安全性和效率。随后，广播技术的横空出世，更是让无线通信走进了寻常百姓家，成为人们获取新闻、娱乐等重要信息不可或缺的渠道。

如今，无线通信技术已经取得了长足的发展，其在人类生活中的应用也愈发广泛。从仅覆盖几米的蓝牙耳机，到能覆盖几百米的无线局域网；从连接几公里基站的手机通信，到实现数万公里跨洋通信的卫星技术，无线通信已经深深渗透到人们生活的方方面面，成为现代社会不可或缺的一部分。

尽管都是依托射频技术实现通信，但不同的无线通信技术在其发展路径和技术思路上却各具特色。根据通信距离的不同，无线通信技术主要可以分为以下几类代表性技术：

- **短距离通信：**以无线局域网为代表，满足近距离的数据传输需求。
- **长距离通信：**蜂窝网络技术，实现广域覆盖的移动通信。
- **远距离通信：**卫星通信技术，跨越天际，实现全球范围内的通信连接。
- **短距离通信：有线网络的延申与扩展**

  虽然无线通信网络在无线电报、海上救援等领域已有应用，但真正走进大众视野，并受到广泛关注，还得归功于无线局域网（WLAN，Wireless Local Area Network）技术的蓬勃发展。

  无线局域网技术的起源可以追溯到 20 世纪 70 年代，当时美国夏威夷大学的研究人员从美军二战期间使用的无线电传输技术中汲取灵感，成功创建了第一个基于封包式技术的无线电通信网络——ALOHA NET，这标志着无线局域网的雏形初现。

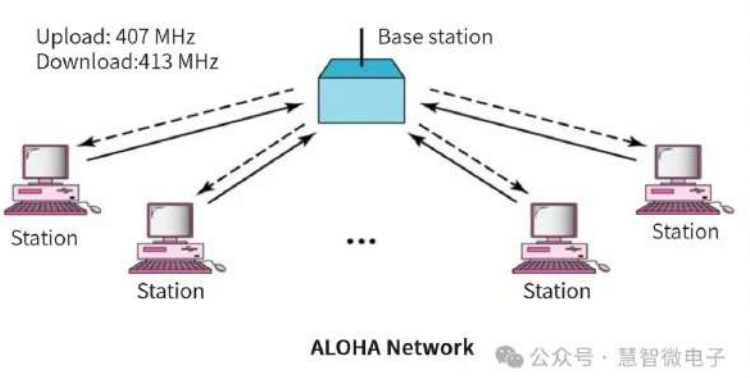


图 5.199: 早期的无线通信网络 ALOHA NET 的设想图

进入 20 世纪 80 年代，随着个人电脑的广泛普及和移动办公需求的日益增长，人们迫切需要一种能够摆脱有线网络束缚，实现设备间无线互联的技术。无线局域网技术凭借其无需布线、灵活便捷、易于扩展等显著优势，迅速赢得了人们的青睐和关注。

1990 年，电气和电子工程师协会（IEEE）成立了 802.11 WLAN 工作委员会，专注于无线局域网标准的制定与推广工作。经过数年的努力，1997 年 IEEE 正式发布了首个 WLAN 标准——IEEE 802.11，该标准为

无线局域网的物理层和 MAC 层设定了统一规范，极大地促进了无线局域网技术的迅猛发展。随着标准的持续完善与更新，无线局域网技术的性能不断提升，功能也日益丰富，充分满足了用户对高速、稳定、安全网络连接的需求。

1999 年，苹果公司推出了 iBook 笔记本电脑，这是无线网络技术首次被融入笔记本电脑中的创举。这一创新举措为后来设备的移动联网奠定了坚实的基础。在 iBook 的发布会上，苹果公司的传奇领袖乔布斯为了展示电脑确实没有连接任何网络线，在台上用呼啦圈绕着电脑旋转，这一魔术般的操作引发了场下观众的长时间惊叹，而这一幕也成为了科技史上的经典瞬间。



图 5.200: 1999 年苹果 iBook 发布中的经典一幕，乔布斯用呼啦圈证明笔记本的无线连接

无线局域网（WLAN）实质上是有线互联网在终端接入段的延伸与补充。在有线网络广泛铺设的基础上，WLAN 技术凭借无线传输的优势，填补了有线网络难以覆盖或不便布线的区域空白。它将有线网络的信号无缝传递至用户身边的每一个角落，为用户提供了便捷、灵活的网络连接体验。通过 WLAN，用户可以摆脱线缆的束缚，随时随地畅享互联网服务，极大地提升了生活的便捷性和工作效率。

随着科技的飞速发展，为了满足人类日益膨胀的数据传输需求，IEEE 802.11 协议不断演进，从 802.11a、802.11b 逐步升级到 802.11ac，再到如今的 802.11ax（即 Wi-Fi 6）以及最新的 802.11be（即 Wi-Fi 7）。每一次更新不仅见证了技术的巨大飞跃，更带来了数据传输速率的显著提升。从最初的几 Mb/s 跃升至如今高达 46Gb/s 的峰值速率，Wi-Fi 正以前所未有的速度满足现代社会对高速无线通信的迫切需求。

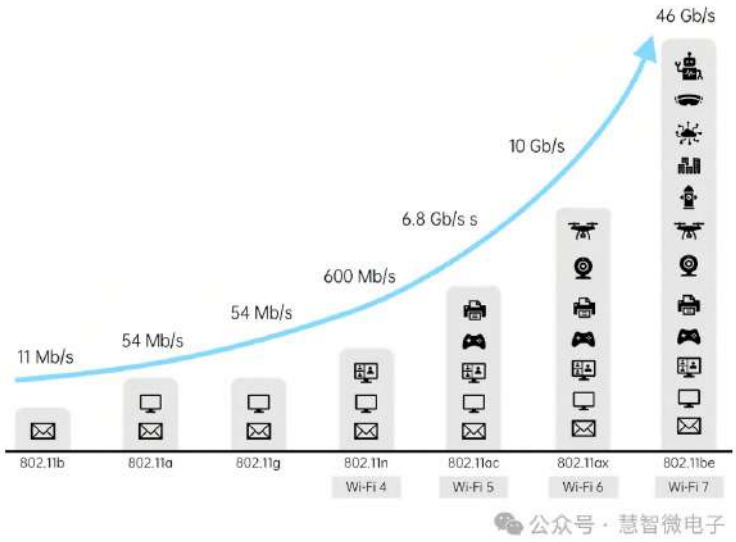


图 5.201: Wi-Fi 速率的飞跃发展

Wi-Fi 虽是短距离通信技术的杰出代表，但短距离通信领域还有其他多种技术，蓝牙、UWB、星闪等多种通信技术同样各具特色。这些技术为设备间的短距离无线连接提供了丰富多样的选择，满足了不同场景下的特定通信需求。

以星闪技术为例，这是一项由华为牵头，并联合多家企业共同研发的新型无线短距离通信技术。星闪技术以其低功耗、高速率、低时延的显著优势，能够支持多达数百台量级设备的互联，这一连接能力远超传统的蓝牙和 Wi-Fi 技术。星闪独特的设计架构，通过灵活的通信接口，既满足了高速率、大传输量、高质量连接的需求，又兼顾了低功耗、轻量级的连接场景。同时，星闪技术还融入了循环前缀正交频分复用（CP-OFDM）波形、Polar 信道编码等多种先进通信技术，进一步提升了通信的性能和可靠性。凭借这些优势，星闪技术有望在智能汽车、智能终端、智能家居、工业智造等多个行业领域展现广泛的应用前景。



图 5.202: 星闪作为新一代近距离无线连接技术的特点

- **长距离通信：开启自在互联新时代**

无线局域网虽已极大地改变了我们的互联网接入方式，让联网体验实现了质的飞跃，但其连接范围仍局限于几米之内，主要服务于室内环境。一旦走出无线网络覆盖区，便会失去网络连接，这显然无法满足人类对随时随地、自由互联的渴望。

要实现这种自在互联，有线通信技术显然力不从心，还是需要借助于无线通信技术。于是，无线电话系统应运而生，为人类的通信方式带来了新的变革。

然而，人类并不需要手机与全球每一部手机都直接相连。因为我们在短时间内的移动范围相对有限，通常只有几公里。如果在这段时间内，我们的手机能够始终保持与一个稳定的无线网络连接，并在移动到下一个区域时，顺利切换到下一个无线网络，那么对于我们来说，就仿佛一直处于联网状态，实时在线。因此，构建一个能够连续覆盖、无缝衔接的全球无线网络，成为了实现人类自在通信的可行技术方案。

这个技术实现的方案就是蜂窝网络。

20 世纪 40 年代，贝尔实验室凭借射频技术，推出了首个无线电话系统。该系统通过射频信号，在固定的基站与移动的车载电话之间建立通信联系，尽管当时需要人工接线员进行呼叫转接，且每个城市仅支持 3 名用户同时连接，但其“固定基站”+“移动终端”的通信模式却为后来的蜂窝网络普及奠定了坚实基础。随后，美国、欧洲、日本等地纷纷着手发展和改进无线电话系统，大多数系统都沿用了“固定基站”+“移动终端”这一核心模式。人们将覆盖区域细分为多个小区，每个小区均配备一个基站和一个控制器，基站负责连接并服务其覆盖范围内的移动终端。而不同基站之间，则通过有线电话网实现互联互通。这种覆盖网络因其形似蜜蜂窝的结构，而得名蜂窝网络；同时，由于其如细胞般的分布特点，也被称作 Cellular（细胞）网络。



图 5.203: 蜂窝网络结构示意图

20 世纪 70 年代至 80 年代，射频无线通信系统逐步迈向全球化阶段，涌现出一系列标准化、兼容化的 1G 通信系统。到了 1991 年，蜂窝网络迎来重大升级，全球较为统一的 2G 标准 GSM（Global System for Mobile Communications，全球移动通信系统）应运而生。GSM 标准以其高效的频谱利用率、高质量的语音和数据传输能力，以及低成本的设备服务等优势，赢得了全球众多国家和地区的广泛支持和采用，成为当时最流行的 2G 系统。基于 GSM 标准的手机实现了全球漫游，让人们能够无障碍地与全球伙伴保持联系。

尝到了全球射频无线标准统一的甜头后，全球各国和地区在蜂窝通信协议的发展上保持了紧密的合作与步调一致。大约每 10 年，蜂窝通信协议就会迎来一次重大演进：2001 年左右诞生了 3G，2011 年左右迎来了 4G，而到了 2019 年左右，5G 技术应运而生。随着通信协议的不断进步，通信速率也在持续提升，但全球各国和地区始终保持着一致的发展步伐。

经过近 50 年的蓬勃发展，蜂窝网络的通信速率实现了飞跃式提升。在 5G 蜂窝网络中，无线通信已经可以实现 20Gb/s 以上的峰值速率。

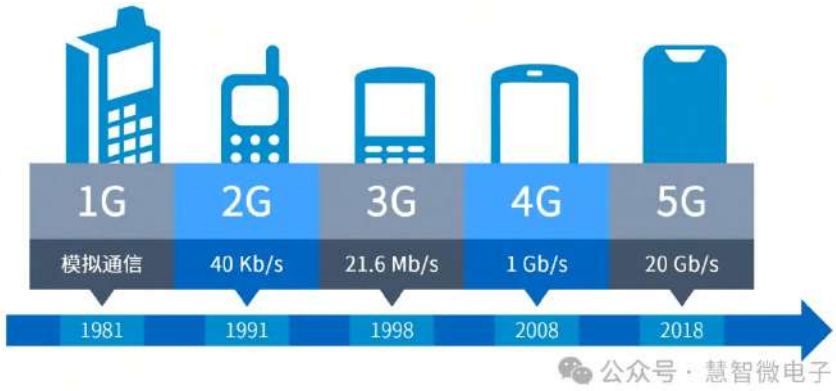


图 5.204: 蜂窝网络的技术演进

- **远距离通信：卫星通信开启全球直连新篇章**

蜂窝网络虽已带来近乎全球的无线互联体验，但其运行基础仍依赖于几公里范围内的有线通信基站部署。这意味着，蜂窝网络主要覆盖城市、乡村等人口密集区域，而对于全球广袤的沙漠（占比约 10%）和浩瀚的海洋（占比约 71%），蜂窝网络则力不从心。

这些区域虽人口稀少，但对通信的需求却尤为迫切。人类要探索未知，拓展边界，就必须掌握在这些区域的通信能力。而预先铺设有线电缆显然不切实际，因此，无线远距离通信成为了这些区域的唯一可行方案。在这些应用场景中，卫星通信成为实现全球直连的关键技术。

卫星通信是一种利用卫星进行信息传输的技术，自 1957 年苏联成功发射第一颗人造卫星以来，便开启了

人类卫星时代的新篇章。1962 年，AT&T 公司发射的卫星首次横跨大西洋，实现了大洋两岸的电话和电视转播，这标志着商用卫星通信技术的正式诞生。此后，卫星通信技术迅猛发展，被广泛应用于电视转播、电话通信、电报传输、传真服务等多种通信业务领域。

根据卫星在轨高度的不同，卫星移动通信可分为 GEO（静止轨道）卫星系统、MEO（中轨道）卫星系统和 LEO（低轨道）卫星系统三大类。

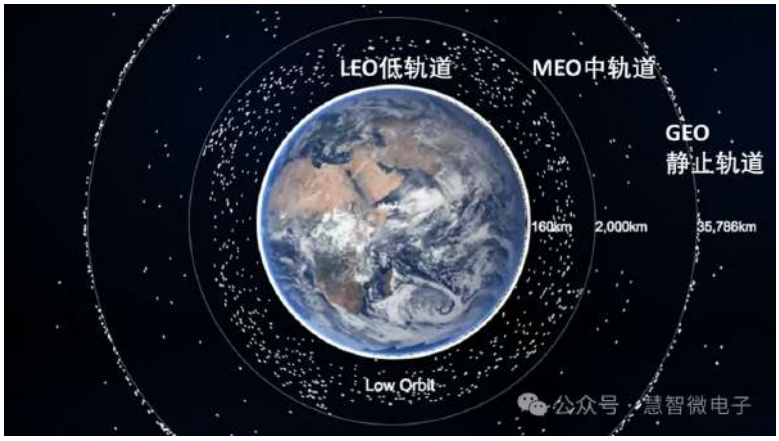


图 5.205: 卫星的分类

GEO（静止轨道）卫星系统运行在赤道上方 35,800 公里处，与地球保持同步，仅需少数几颗卫星即可实现全球覆盖（除极低地区外）。然而，由于其运行高度较高，导致传播延时相对较大。Inmarsat 系统便是 GEO 卫星系统的典型代表，它由 13 颗卫星组成，为国际船舶、飞机等提供电话、传真和数据业务，目前拥有约 30 万用户。

MEO（中轨道）卫星系统的轨道高度介于 2,000 至 30,000 公里之间，由于无法与地球保持同步，其发展速度相对 GEO 和 LEO 系统较慢。Odyssey 系统是 MEO 卫星系统的典型代表，由 12 颗卫星组成，服务全球 280 万用户，但通信速率相对较低。

近年来，LEO（低轨道）卫星系统备受关注。其轨道高度在 1,500 公里以下，虽然需要较多的卫星才能实现全球覆盖，但其传播损耗小、时延短，特别适合多用户全球互联，因此成为宽带卫星通信的重点发展方向。卫星通信作为无线通信的一种重要方式，使得终端的连接距离突破了数万公里的限制，真正实现了覆盖整个地球的无线互联。它将无线通信的无限互联特性发挥到了极致，为人类的通信事业带来了前所未有的变革。

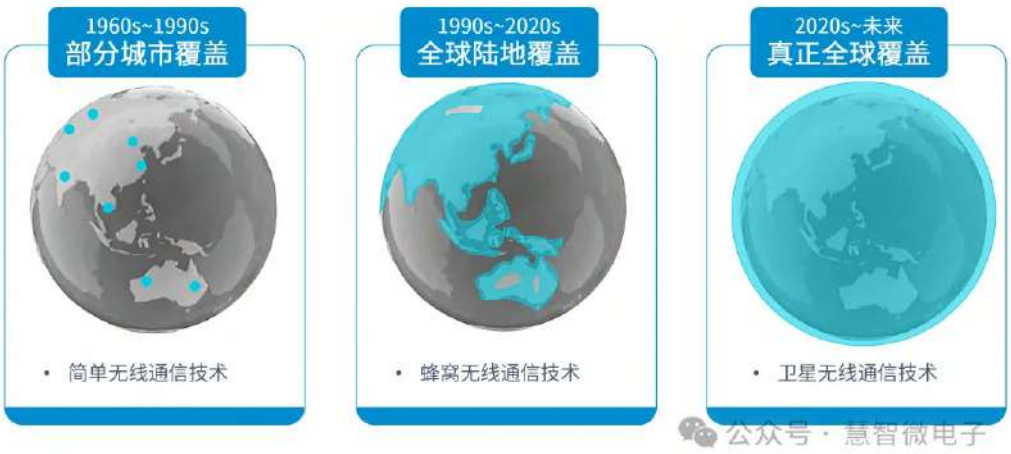


图 5.206: 卫星无线通信将带来人类互联的变革

5.11.2 有线与无线通信的对比

有线通信和无线通信，虽然都是用于将信息从一个地点传递到另一个地点的通信方式，但两者之间存在着诸多显著差异。这些差异不仅影响了它们的设计思路，也决定了它们各自的演进逻辑。

具体来说，有线通信与无线通信的不同主要体现在以下几个方面：

1. 传输媒介不同；
2. 传输特性各异；
3. 应用场景有别。

5.11.2.1 传输媒介的差异

有线通信与无线通信之间最为本质的区别在于它们的传输媒介。

有线通信依赖于金属导线、光纤等有形媒质来传输信息，而无线通信则是通过电磁波在空间中传播信号。这一传输基础的不同，奠定了两种通信技术的根本差异。

这一差异直接导致了若干显性的技术特点。例如，有线通信需要预先进行网络部署，而无线通信则无需此步骤。

此外，传输媒介的差异还对技术的发展产生了深远的影响：

- 有线通信对传输媒介具有独占性，而无线通信则需共享空中的传输媒介；
- 有线通信可以拥有多个传输通路，而无线通信在不进行信息处理复用的情况下，在一定距离范围内通只有一条传输路径；
- 有线通信具有较强的抗干扰能力，而无线通信则需对干扰进行更多考虑和处理；
- 有线通信中信号传输的功率可预测性较高，而无线通信由于路径复杂，损耗更为不可控。

由于上述根本差异，无线通信与有线通信之间的区别远非仅仅将网线替换为无线信号那么简单。这导致了两种通信技术在研究重点上常常存在巨大的差异。

以我个人在无线通信领域的经验为例，我们关注的是如何在有限带宽内实现信号的纯净传输，如何设计滤波器以有效抑制带外干扰，确保我们发射的信号仅占据指定的频段，不对其他频段造成干扰。我们的工作始终围绕着“窄带”信号处理展开。

而在与有线光通信行业的同行交流时，我发现工程师的思考方式截然不同。大家致力于拓宽电路的带宽，做的是“宽带”电路，以充分利用有线媒介的全部传输能力。这是因为有线媒介是独占的，无需与其他用户共享。

此外，由于无线通信网络中的所有用户需要共享同一网络资源，这促使无线通信技术中大量研究了复用技术，如时分复用、频分复用、码分复用等，竭尽所能地提高无线网络的用户容量。而在有线通信中，这种复用技术的应用则相对不那么极致。

另外，在干扰处理，功率控制算法等方面，二者也表现出明显的不同。



图 5.207: 有线通信与无线通信传输媒介的对比

5.11.2.2 传输特性的差异

有线通信与无线通信由于传输媒介的不同，展现出截然不同的传输特性。

在使用带宽、调制方式、传输速率这些与传输特性紧密相关的方面，两者存在显著差异。

就使用带宽而言，有线通信能够利用的带宽范围广泛，可达几 GHz 甚至几十 GHz。而无线通信虽然总频谱宽度也达到数十 GHz，但由于需要与众多无线通信应用共享，因此单个通信技术下的单个用户通常只能使用几十 MHz 的带宽，甚至在一些频谱资源紧张的地区或运营商网络中，信号带宽可能仅限于几 MHz。

就调制方式来说，在无线通信中，频谱资源极为宝贵，因此调制方式成为了重点研究的领域，并取得了诸多改进。

例如，最新的 Wi-Fi7 协议已经采用了 4096QAM 的信号调制技术，每个符号位可以携带 12bit 的信号信息。而在光纤通信中，广泛使用的仍是 NRZ（非归零）和 PAM4（四电平脉冲幅度调制）技术，每个符号位仅携带 1~2bit 的信号信息。此时，如果传递同样 12bit 的信息，NRZ 和 PAM4 调制分别需要 12 个以及 6 个时钟周期传递完成，而 4096QAM 只需要 1 个时钟周期就可以传递完成。

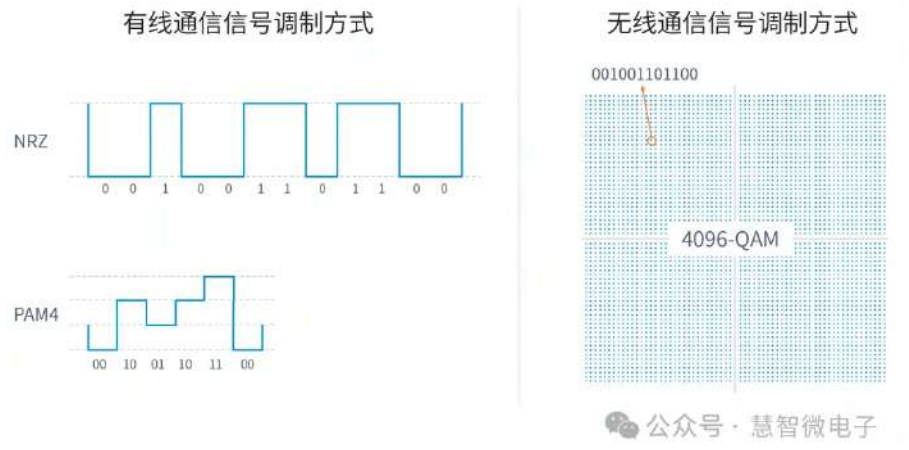


图 5.208: 有线及无线不同的调制方式的对比

就传输速率而言，随着通信技术的不断进步，有线通信和无线通信在速率上都实现了显著提升。然而，由于有线通信具有信道独享的优势，其速率始终保持在明显高于无线通信的水平。

值得注意的是，无线通信的峰值速率通常是在用户独享信道时的理论值。在实际无线网络环境中，由于多用户接入和非理想特性的影响，实际体验速率往往远低于理论速率。例如，尽管 5G 和 Wi-Fi7 技术理论上可达到数十 Gb/s 以上的传输速率，但在日常手机使用中，实际速率通常仅在 100Mb/s 以下。

下图展示了过去 50 年中，有线通信与无线通信代表技术的速率演进情况。

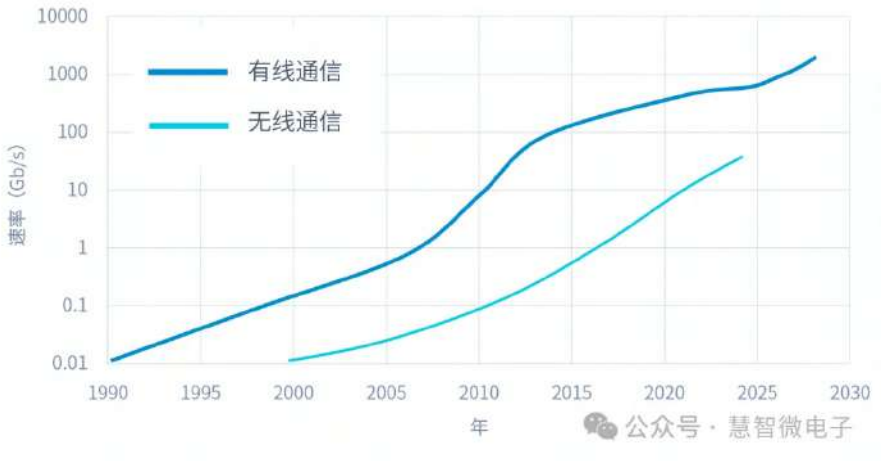


图 5.209: 有线通信与无线通信演进速率比较

5.11.2.3 应用场景的差异

尽管在速率比拼上，无线通信可能不及有线通信，但这并不意味着无线通信就此“落败”。衡量通信也不是只有速率这唯一指标。

从过去的 1G 到如今的 5G，无线通信技术经历了飞速的迭代与发展，传输速率实现了百万倍的提升。然而，速率并非衡量通信技术的唯一指标。通信技术还具备诸多其他特性，这些特性相互融合，使得在某些特定场景中，无线通信技术能够脱颖而出，成为无可替代的选择。

- **有线通信：高可靠、大容量通信的首选**

  有线通信凭借其出色的稳定性和大容量传输能力，已成为高可靠性、大容量通信领域的首选。

  通过光纤、同轴电缆、双绞线等物理导线进行信号传输，有线通信确保了信息在传输过程中的低损耗、高保真度。特别是在需要高速率数据传输的场景中，有线通信展现出了无可匹敌的优势。

  无论是固定通信网络的建设，还是对通信质量和可靠性要求极高的 AI 数据中心服务器等应用场景，有线通信都以其独特的优势，赢得了业界的广泛认可，成为首选的通信方式。

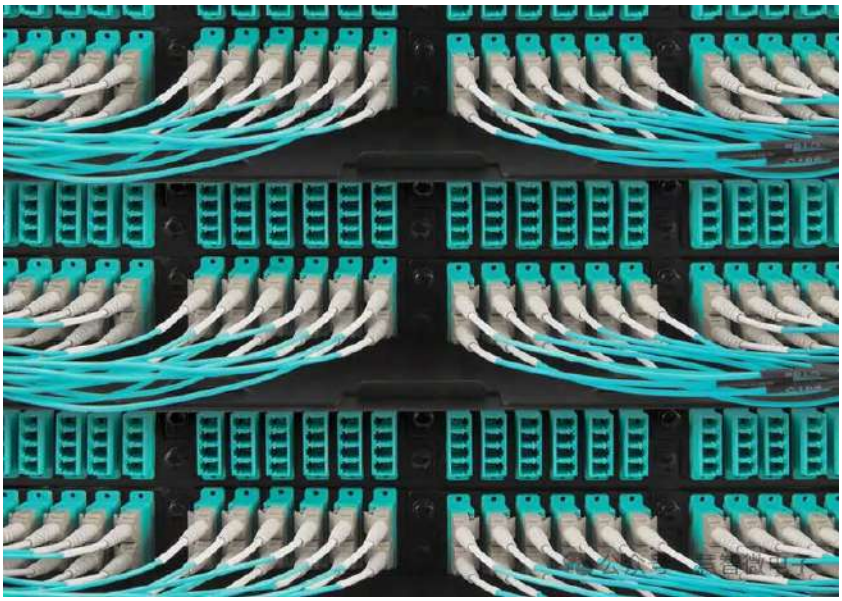


图 5.210: 用于数据中心有线通信的光纤互联

- **无线通信：打造多样化应用场景**

无线通信则以其灵活便捷的特性，开创了丰富多彩的应用场景。

无线通信使终端摆脱了线缆的束缚，实现了信息的空中自由传递。在移动通信、物联网等领域，无线通信发挥出了其独特的优势。无论是日常的个人通讯，还是智能设备的互联互通，无线通信都以其高效、广泛的覆盖范围，成为了一种灵动且不可或缺的通信方式。



图 5.211: 无线通信可实现不同距离的丰富覆盖

5.11.3 “有线”与“无线”通信未来展望

有线通信与无线通信技术，作为现代通信领域的两大核心支柱，各自秉持着独特的优势与无限的发展潜力，未来必将绽放出更加璀璨的光芒。

有线通信技术，以其卓越的稳定性和大容量传输能力，长期以来一直是固定网络和数据中心的基石。在未来，随着 AI 技术的蓬勃发展，对高速、低延迟、高可靠性的通信需求将愈发迫切。这将推动有线通信技术不断创新与突破，尤其是在光纤通信领域，更高速率、更低损耗的光纤传输技术将不断涌现，为 AI 数据中心、云计算等高性能应用场景提供强有力的支撑。同时，有线通信技术也将与新兴技术如量子通信等深度融合，探索出更多前所未有的通信方式，迎来有线通信技术的再次爆发。

而无线通信技术，作为连接世界的纽带，其灵活性和广泛性早已深入人心。未来，无线通信技术还将继续以惊人的速度发展，不断拓宽其应用边界。随着 5G 技术的全面商用和 6G 技术的研发推进，无线通信技术将实现更高速率、更低时延、更广泛覆盖的连接，为移动通信、物联网、车联网等领域带来革命性的变化。同时，无线通信技术也将与人工智能设备、大数据等新兴技术深度融合，催生出更多创新应用和服务，让社会变得更加智能、便捷和多姿多彩。未来社会，无线通信将无处不在，成为推动社会进步和发展的重要力量。

另外，有线与无线通信技术各有千秋，未来一些技术可以相互借鉴，共同发展。例如无线通信可借鉴有线通信的大带宽传输优势，探索更高频率、更宽频谱的通信技术。而有线通信则能从无线通信的高阶调制、多用户复用技术中汲取灵感，提升光纤资源利用效率。二者相互促进，可能可以提供更高效、更智能的通信解决方案，共同推动人类世界向更加智能、便捷的方向发展。



图 5.212: 有线与无线通信还将继续发展

5.11.4 结语

在技术研究中，通过对比不同技术，可以更深刻理解每种技术的背后的本质。

有线通信与无线通信，作为通信技术的两大分支，正是适合拿来对比研究的对象。它们各自承载着信息传输的重任，却以不同的方式实现着这一目标。通过对比这两种技术，不仅可以看到它们在传输媒介、传输特性以及应用场景上的显著差异，更深刻理解了它们各自的优势与局限，也可以更深刻的理解技术原理。

在对比的过程中，也可以发现，技术的发展并非孤立存在，而是可以相互借鉴、共同进步的。有线通信与无线通信在演进过程中，不断从对方身上汲取灵感，实现了技术的突破与创新。这种相互借鉴、共同发展的模式，为未来的技术发展提供了更多的可能性。

展望未来，有线通信与无线通信将继续携手并进，共同推动通信技术的蓬勃发展。随着 AI、大数据等新兴技术的不断涌现和融合，未来的通信技术将更加高效、智能、便捷。而在这个过程中，有线通信与无线通信的对比研究将继续发挥重要作用，引领人类走向更加广阔的通信世界。

参考文献

1. 无线重构世界，电子工业出版社
2. Wi-Fi7 协议，IEEE P802.11be™/D3.0, 2023-1
3. 5G NR 协议，3GPP TS 38.101-1 V18.6.0, 2024-06
4. Wi-Fi 市场探析：今日应用与未来趋势，慧智微
5. 传遍朋友圈的“iPhone 13 支持卫星通信”，支持的究竟是什么？，慧智微
6. Historical Trends in Wireline Communications 60 Improvement in Speed in 20 Years，Razavi，2015
7. 国际星闪无线短距通信联盟，https://www.sparklink.org.cn/

在线阅读

本文摘自慧智微微信公众号，扫描以下二维码，在线阅读

5.12 MIPI RFFE 协议解读

作者：石头

MIPI （Mobile Industry Processor Interface）协议是 MIPI 联盟（MIPI Alliance）提出的用于标准化移动终端系统各器件间通信的通信协议。MIPI 联盟 2003 年成立，最早的初衷是为了标准化显示接口。



图 5.214: MIPI 联盟最早于 2003 年成立

经过多年发展，截止 2019 年 12 月，MIPI 联盟已经发布 48 份标准，应用领域也扩展至汽车、工业及 AR/VR 等领域。MIPI 联盟拥有包括终端厂商、器件厂商、平台厂商以及测试厂商在内的 339 个会员单位。MIPI 协议成为手机终端各器件通信的主流标准协议。MIPI 联盟工作覆盖范围及涉及的技术不应用领域如图5.215及图5.216所示。

MIPI RFFE（MIPI RF Front-end，MIPI 射频前端）是 MIPI 联盟 RFFE 工作组在 2010 年推出的用于移动终端射频前端控制的控制接口标准。

在 MIPI RFFE 推出之前，射频前端的控制解决方案复杂，如果用并行的 GPIO 口进行控制，需要接口过多，系统成本无法承载；一些厂商开始自定义串口控制，不过由于通信协议复杂，需要考虑复杂的软件控制和时序控制，实现困难，并且自定义串口也不利于不同厂商器件的通信。

2010 年，MIPI 联盟推出用于射频前端控制的 MIPI RFFE 协议。MIPI RFFE 协议总线由一根电源线（VIO）以及两根控制线（SCLK 和 SDATA）构成，实现简单，易于部署，可实现时序范围要求内的近实时控制。图5.217为 MIPI RFFE V1.0 中所提出的系统架构。

在过去 10 年中，无线终端通信协议从 3G 演进至 4G、5G，MIPI 协议也不断演进。目前，MIPI 协议已演进至 V3.0 版本，支持多种 Trigger 模式，适应 5G 通信系统中更多器件以及更严苛的时序控制要求。各个 MIPI 版本演进时增加的主要特性如图5.218。

慧智微在 MIPI RFFE 协议的 V1.x 时代加入 MIPI 联盟 RFFE 工作组，对 MIPI RFFE 协议技术演进长期关注。以下将对 MIPI RFFE 协议的技术进行讨论。

5.12.1 MIPI RFFE 协议解读

5.12.1.1 MIPI RFFE 系统架构

如图5.219是 MIPI RFFE 的系统架构：

一个 MIPI RFFE 的架构通常包含：

1. Signal Bus：包含 SCLK 和 SDATA 两条信号线

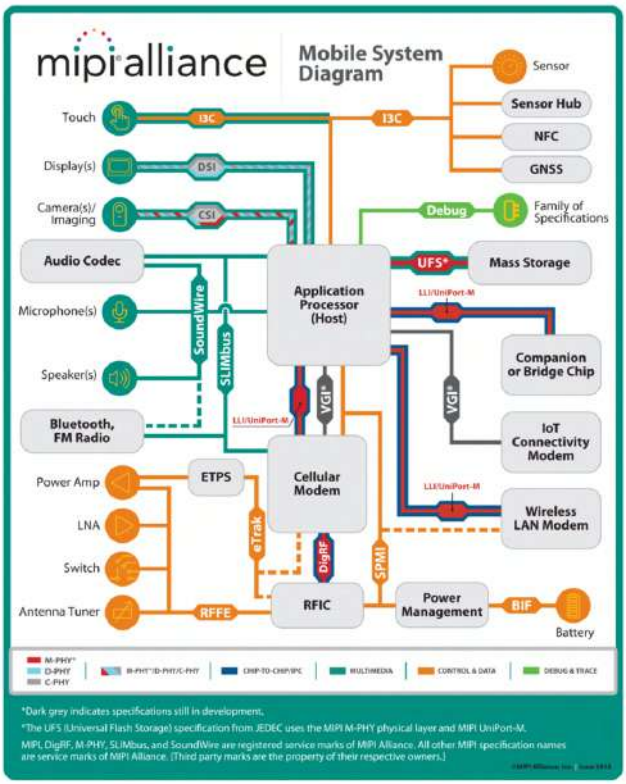


图 5.215: MIPI 联盟工作覆盖范围



图 5.216: MIPI 联盟所涉及的技术领域及应用领域

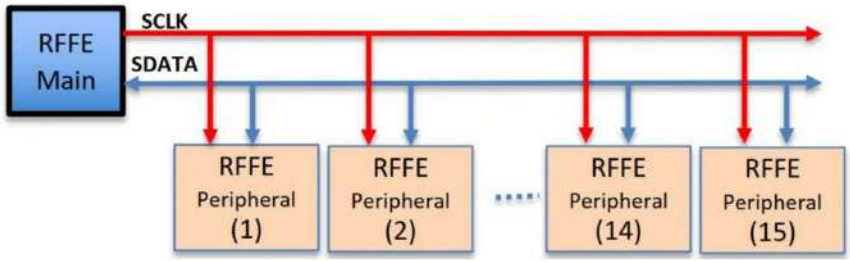


图 5.217: MIPI RFFE V1.0 中所提出的系统架构

MIPI RFFE版本	推出时间	主要更新
V1.0	2010年7月	• 首次发布 • 26-MHz • 支持3个Trigger
V1.1	2011年11月	• 漏洞修复 • 无重大更新
V2.0	2014年12月	• 同步读取 • 速度扩展至52MHz • 支持多主（Multi-Master） • 扩展预留寄存器空间
V2.1	2018年4月	• 遮罩写入（Masked-Write） • 扩展Trigger • 扩展寄存器
V3.0	2020年4月	• 更多的扩展Trigger支持 • 计时Trigger（Timed Trigger） • 映射Trigger（Mappable Trigger）

图 5.218: 各个 MIPI 版本演进时增加的主要特性

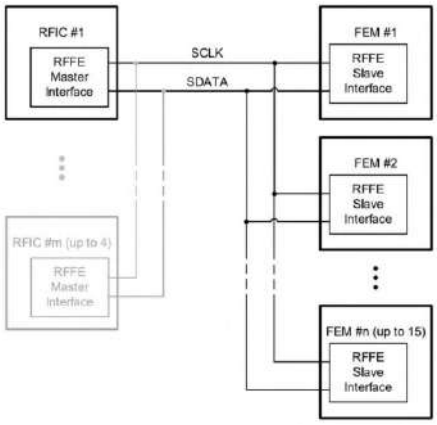


图 5.219: MIPI RFFE 系统架构（来自 MIPI RFFE V3.0）

2. RFFE Master Device：控制系统 MIPI signal bus 的器件，一般为手机系统中的 RFIC（transceiver），单个 MIPI 系统最多可支持 4 个 Master device
3. BOM：Bus Own Master，任一时刻只能有一个 BOM 被激活
4. RFFE Slave Device：MIPI 受控器件，一般为射频前端模组，例如 PA、LNA、Switch 等，单个 MIPI 系统最多可支持 15 个 device

一个 Slave 器件的 MIPI 管脚除了 SCLK 和 SDATA 还有一个 VIO，这些管脚功能如图5.220。

一个 Slave 器件会有三种状态，分别为 Active、Shutdown 和 Startup。这三种状态的转换关系如图5.221。

1. VIO 控制 MIPI 系统是否 Shutdown
2. 当 VIO=high 时，MIPI 进入 Startup 状态，随后自动进入 active 状态
3. 如需对 MIPI 寄存器进行重置，可发 1C 寄存器中 PWR_MODE[0]=0b1

管脚	功能
VIO	• 提供电压电平参考及电压源 • 所有相同MIPI bus上的器件需要使用一个相同的VIO
SCLK	• 一般由BOM驱动产生 • 提供时钟信号以及同步功能 • 时钟信号并不是一直都在发，只有在有command sequence要发送时才会触发 • 时钟速度目前v3.0最大可支持52MHz，未来可能还会提升
SDATA	• 一般由BOM驱动，也可由slave驱动 • 双向信号 • 主要提供控制信号和数据的传送

图 5.220: MIPI 管脚功能

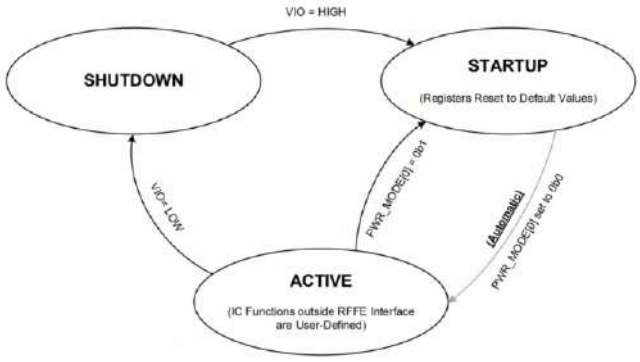


图 5.221: Slave 器件的三种状态：Active、Shutdown 和 Startup

5.12.1.2 MIPI RFFE 控制指令

MIPI 通信主要通过控制 Slave 器件的相关寄存器，以达到控制器件的目的。在 MIPI RFFE 协议中定义了多种寄存器，寄存器包含：

1. 系统定义寄存器：系统规定有特殊应用的寄存器，如上文所提到的 1C 寄存器有 PWR_MODE 控制等
2. 用户定义寄存器：用户可根据具体应用定义寄存器功能

MIPI 究竟如何使用控制指令来控制这些寄存器？图5.223是 MIPI RFFE 各个操作的控制指令结构。如图5.223所示，一个操作的控制指令基本会包含以下几部分：

1. SSC（Sequence Start Condition）：指令开始标志
2. Command Frame：一般包含 SA（Slave Address）、操作类型对应代码、以及奇偶校验码
3. Data Frame：不一定有，同时也包含奇偶校验码
4. BPC（Bus Park Cycle）：指令结束标志

以一个最常见的控制指令寄存器写入（Register Write）为例做介绍如图5.224及图5.225。按照图所示，如果以时钟速率 52M 进行计算的话，完成一次寄存器写入的指令时间大约为：1/52M*24≈0.5µs

根据控制指令表，我们找到有一种更高阶的寄存器写入操作——Masked Write。Masked Write 提供了一种可以可以只改寄存器部分 bit 位的 value 的方案。例如要更新某一寄存器其中的 D1/2/3/7 这几个 bit 的值，其他 bits 不做更新，那么就可以采用 Masked Write 功能，具体指令结构和操作示意如图5.226及图5.227。

5.12.1.3 MIPI RFFE 的 ID 配置

MIPI RFFE 中定义了多种 ID，ID 作为 MIPI 中 Slave 器件的识别代号有着非常重要的作用。

MIPI RFFE 包含的相关 ID 以及对应寄存器地址、说明如图5.228。包含：

1. MID（Manufacturer ID）：每个厂商在加入 MIPI Alliance 之后都会被分配一个专门的 MID，可以从官网查询到
2. PID（Product ID）：由厂商自行管理给不同产品配置不同的 PID，在 MIPI RFFE V2.1 引入了 Extend-PID，这样 PID 就包含 16bits 供配置
3. USID（Unique Slave ID）：MIPI bus 主要通过 USID 对 Slave 器件进行寻址，USID 共 4bits 16 组状态，但一个 MIPI bus 上最多 15 个 Slave 器件，MIPI RFFE 定义 SID=0b0000 时为广播 ID，即 BSID（Broadcast Slave ID）

MID 和 PID 所对应的寄存器都是 read-only，而 USID 一般有一个默认的值，同时可以支持可编程。原因是 MIPI 指令是以 USID 进行寻址的，当一个 MIPI bus 上的所有 Slave 器件存在两个或以上的 USID 是相同的时候，会导致无法区分，因此可编程的 USID 至关重要。

图5.229是可编程 USID 的机制图。更新 USID 的步骤流程为：

- 通过当前 USID 写入 PID 寄存器 (0x1D)，写入值需要和想要更新器件的 PID 值相同

Register Address	Register Name	Number of Bits	Notes
0x0000	REGISTER_0	0 to 8	User-defined register purpose and content
0x0001 – 0x001B	User-Defined Registers	0 to 8	User-defined register names, purpose, and content
0x001C	PM_TRIG[7:0]	8	State/Mode Bits and Triggers Register Bits 7:6 – PWR_MODE<1:0> Bits 5:0 – TRIG_REG<5:0>
0x001D	PRODUCT_ID[7:0]	8	Bits 7:0 – PRODUCT_ID<7:0> Product Identification – Selected by Device Manufacturer
0x001E	MANUFACTURER_ID[7:0]	8	Manufacturer Identification, Assigned by MIPI Bits 7:0 – MANUFACTURER_ID<7:0>
0x001F	USID[7:0]	8	USID Register Bits 7:4 – Manufacturer Identification, MANUFACTURER_ID<11:8> Bits 3:0 – USID<3:0> Unique Slave Identifier (USID), Fixed or Programmable
0x0020	EXT_PRODUCT_ID[7:0]	8	Bits 7:0 – EXT_PRODUCT_ID<15:8> – Selected by Device Mfr. (Optional for Use)
0x0021	REV_ID[7:0]	8	Bits 7:0 – REVISION_ID<7:0> User-defined organization for device revision ID information
0x0022	GSID0-GSID1[7:0]	8	Bits 7:4 – GSID0<3:0> (Primary GSID – Optional) Bits 3:0 – GSID1<3:0> (Secondary GSID – Optional)
0x0023	UDR_RST[7:0]	8	Bit 7 – User-Defined Registers Software Reset Bits 6:0 – Spare (Reserved)
0x0024	ERR_SUM[7:0]	8	Bits 7:0 – User-defined Error Logging
0x0025	Reserved	8	Reserved for future RFFE Use (Do Not Use)
0x0026	Reserved	8	Reserved for future RFFE Use (Do Not Use)
0x0027	Reserved	8	Reserved for future RFFE Use (Do Not Use)
0x0028	Reserved	8	Reserved for future RFFE Use (Do Not Use)
0x0029	Reserved	8	Reserved for future RFFE Use (Do Not Use)
0x002A	Reserved	8	Reserved for future RFFE Use (Do Not Use)
0x002B	BUS_LD[7:0]	8	Bits 7:4 – Spare (Reserved) Bits 3:0 – BUS_LD<3:0> Slave SDATA Bus Load (Optional for Use)
0x002C	TEST_PATT[7:0]	8	Bits 7:0 – TEST_PATT<7:0> Slave Fixed Test Pattern – Value = 0xD2 (Optional for Use)
0x002D	EXT_TRIG_A_MASK[7:0]	8	Extended Triggers Masks Block A Bits 7:0 – EXT_TRIG_A_MASK<7:0>
0x002E	EXT_TRIG_A_REG[7:0]	8	Extended Triggers Register Block A Bits 7:0 – EXT_TRIG_A_REG<7:0>
0x002F	EXT_TRIG_B_REG[7:0]	8	Extended Triggers Register Block B Bit 7 – Reserved Bits 6:0 – EXT_TRIG_B_REG<6:0>
• • • • • •			
0x0040 – 0xFDFF	User-Defined Registers	0 to 8	User-defined register names, purpose, and content
0xFE00 – 0xFFFF	Reserved Registers	0 to 8	Reserved for future RFFE Use (Do Not Use)

图 5.222: MIPI RFFE V3.0 所定义的寄存器（完整版请参与协议文档）

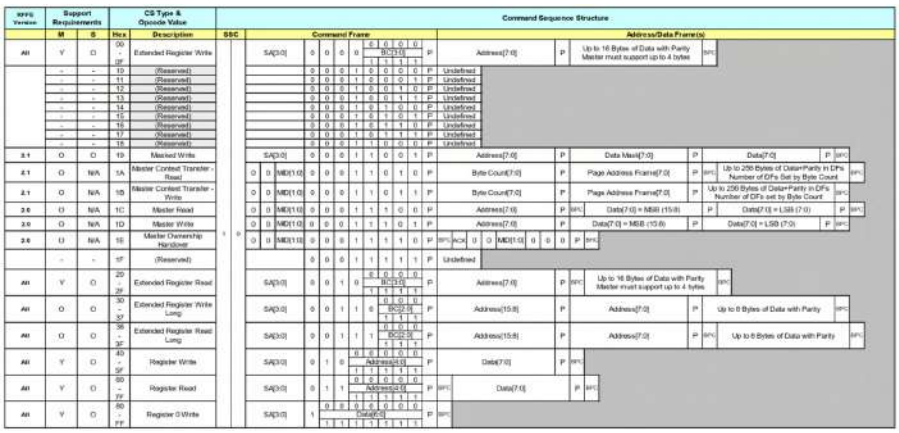


图 5.223: MIPI RFFE V3.0 中所定义的控制指令结构

Description	SSC		Command Frame											Data Frame				
Register Write	1	0	SA[3:0]	0	1	0	0 0 0 0 0 / Address[4:0] / 1 1 1 1 1					P		Data[7:0]			P	BP
Register Read	1	0	SA[3:0]	0	1	1	0 0 0 0 0 / Address[4:0] / 1 1 1 1 1					P	BP	Data[7:0]			P	BP

图 5.224: MIPI RFFE 中基本的写入及读取指令

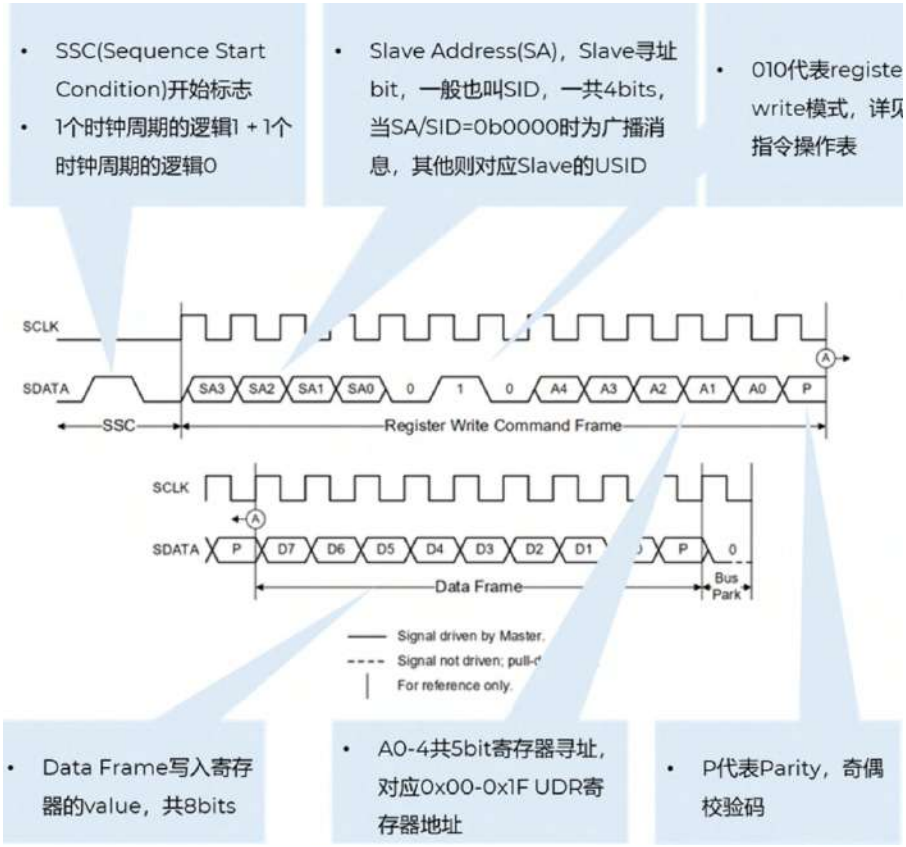


图 5.225: Register Write 操作指令解释

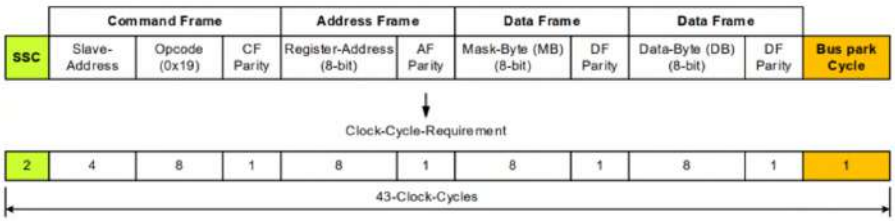


图 5.226: Masked Write 指令结构和操作

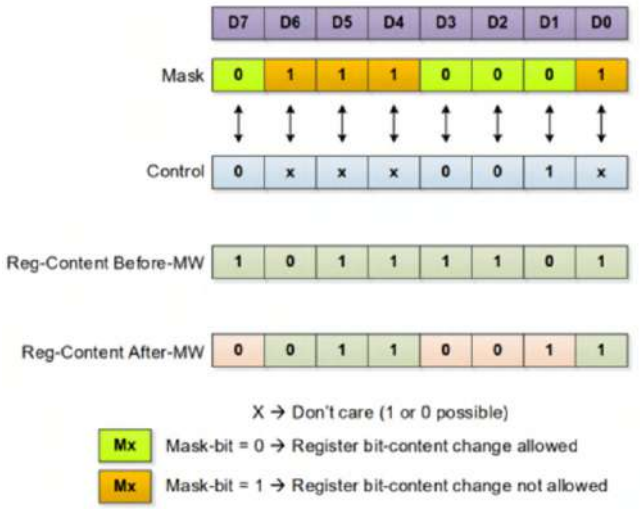


图 5.227: Masked Write 指令结构和操作

Register Address	Bits	Register Name	Notes
0x001D	[7:0]	PRODUCT_ID[7:0]	This is a read-only register. However during the programming of the USID using any of the USID Program Procedures a Write Command Sequence is performed on this register even though the write operation does not change its value.
0x001E	[7:0]	MANUFACTURER_ID[7:0]	This is a read-only register. However during the programming of the USID using any of the USID Program Procedures a Write Command Sequence is performed on this register even though the write operation does not change its value.
0x001F	[7:4]	MANUFACTURER_ID<11:8>	These bits are read-only. However during the programming of the USID using any of the USID Program Procedures a Write Command Sequence is performed on this register even though the write operation does not change its value.
	[3:0]	USID<3:0>	These bits are read-only for a device not supporting programmable USID; however, for programmable USID they are read/write. Performing a write to this register using the described programming procedures will program the USID in devices supporting this feature. These bits store the USID of the device.
0x0020	[7:0]	EXT_PRODUCT_ID[7:0] (== EXT_PRODUCT_ID<15:8>)	This is a read-only register, and implementation is optional as a means to extend the 8-bit PRODUCT_ID to a 16-bit EXT_PRODUCT_ID. Use of this register is optional in USID programming, despite whether EXT_PRODUCT_ID is used or not. During the programming of the USID using USID Program Procedure 3 a Write Command Sequence is performed on this register even though the write operation does not change its value.

图 5.228: MIPI RFFE 中所定义的多种 ID 说明（来自 MIPI RFFE V3.0）

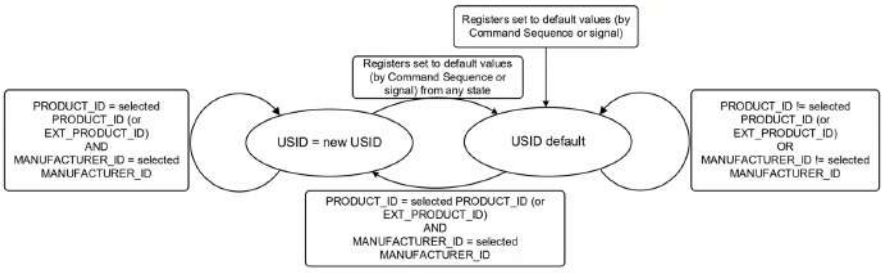


图 5.229: USID 的编程机制（来自 MIPI RFFE V3.0）

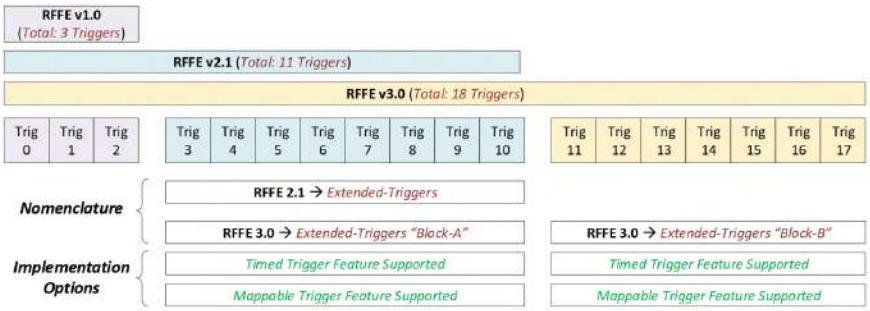


图 5.230: MIPI RFFE 中 Trigger 功能的演进

Command Sequence Type	Paylaod	Required SCLK Cycles	Lat @ 26 MHz (µs)	Lat @ 52 MHz (µs)
Register-Write	1-Byte	25	0.96	0.48
Extended Register Write	1-Byte	34	1.31	0.65
Extended Register Write Long	1-Byte	43	1.65	0.83

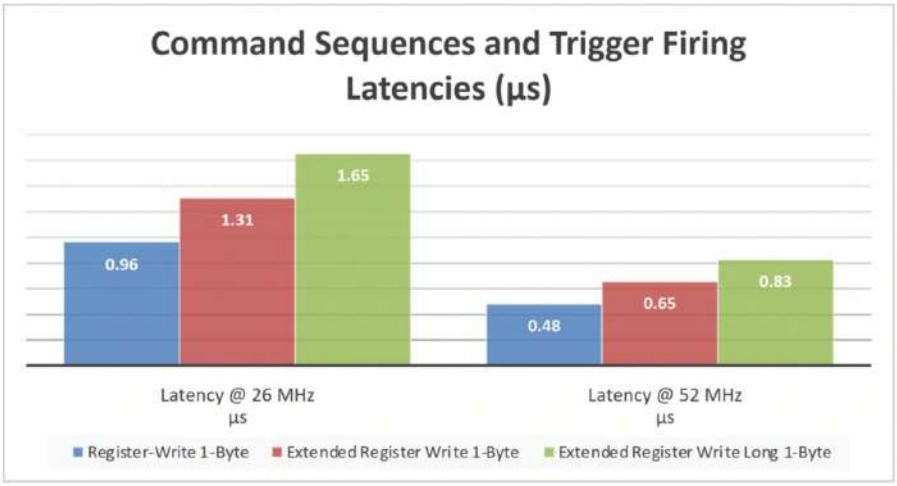


图 5.231: MIPI RFFE V3.0 指令写入时间

- 通过当前 USID 写入 MID 寄存器 (0x1E)，写入值需要和想要更新器件的 MID 值相同
- 通过当前 USID 写入 USID 寄存器 (0x1F),写入要新的 USID 值,当写入 PID/MID 与想要更新的器件 PID/MID 值一致时，新的 USID 值即生效，反之 USID 仍维持原来的值

简单来讲就是只有当 MID/PID 以及默认 USID 对应上之后才能够对 USID 进行更新，这就要求：当一个 MIPI bus 上挂有多个相同产品（MID/PID 相同）时，为了做可编程的区分，一般会在此产品增加 ID pin 脚，其接高电平或 GND 分别对应不同的默认 USID。

5.12.2 MIPI RFFE Trigger 模式

近年来，通信协议演进至 5G。5G MIMO 功能需要控制射频前端多颗器件同时工作，同时 5G 系统用于控制的时间变短，控制接口需要以更短的时间，完成控制。

为了实现更快的控制速度，MIPI RFFE V2.0 中将时钟速度由 26MHz 提升至 52MHz，采用 52MHz 通信后，单个寄存器件读写的时间减少一半。但这还不足以满足 5G 通信系统的时序要求。为此，MIPI V3.0 中着重升级了 Trigger 模式功能。不仅 Trigger 的数量由 V1.0 的 3 个、V2.1 中的 11 个，增加至 V3.0 中的 18 个，而且还增加了 Timed Trigger 和 Mappable Trigger 功能。图5.230为 MIPI RFFE 中 Trigger 功能的演进。

通过 Trigger 模式与 52MHz 时钟的结合，MIPI V3.0 写入指定的时间可以缩小至 1µs 以下。MIPI V3.0 中，指令写入的时间如图5.231所示。

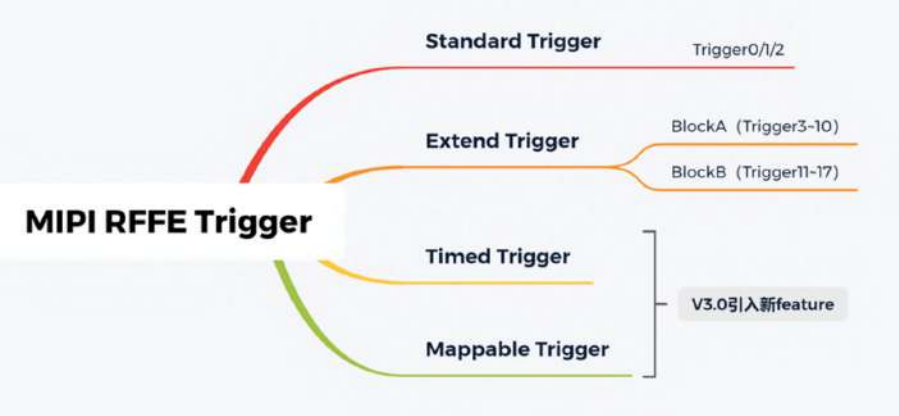


图 5.232: MIPI RFFE 中的 Trigger 功能分类

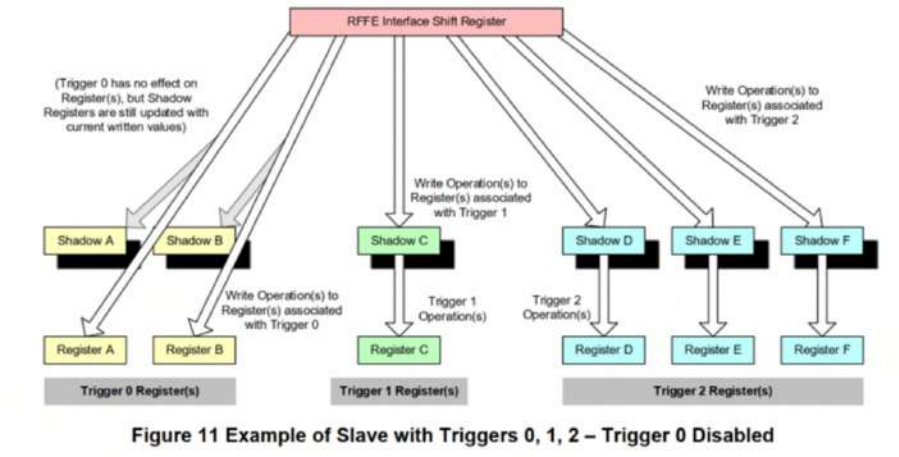


Figure 11 Example of Slave with Triggers 0, 1, 2 – Trigger 0 Disabled

图 5.233: MIPI RFFE 中的 Trigger 功能

Trigger 功能的引入

当要控制一个射频芯片时，可能需要写入多个寄存器，通常可以按照串行方式每个寄存器轮流写入及生效，但由于每个寄存器生效存在时间差（以 52MHz 为例，每个寄存器写入时间为约 0.5μs），这期间可能导致芯片工作不正常。为了解决这个问题，MIPI 协议中针对 UDR 寄存器引入的 Trigger 模式，即通过 Trigger 功能，触发多个寄存器同时生效。

Trigger 作为 MIPI RFFE 的重要 Feature，随着 5G 对更多器件的控制越来越复杂，V3.0 也引入了更多的 Trigger 功能，下图是简单的 Trigger 分类，本章节将一一做解释。图5.232为 MIPI RFFE 中的 Trigger 功能分类。

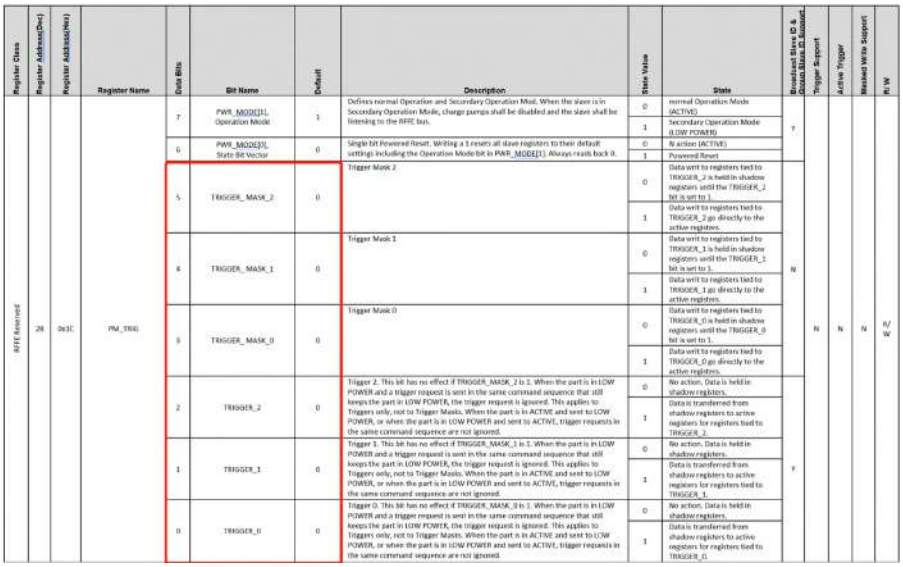
Standard Trigger

MIPI 的 Trigger 机制如图5.233，为了支持 Trigger 模式，MIPI 协议中引入了 Shadow 寄存器，用于存储未经触发的寄存器 value。

1. 当 Trigger0 disable 时，此时写入寄存器 A/B 的值直接生效，同时 A/B 寄存器对应的 shadow register 也同时更新为写入的值；
2. 当 Trigger1/2 enable 时，此时寄存器 C/D/E/F 的写入值首先只能更新在对应的 shadow register 中，只有当 Trigger1 执行写入操作时，寄存器 C 才能正式生效，同理寄存器 D/E/F 也需要相同 Trigger2 操作。
3. 每个寄存器只能对应 1 个或 0 个 Trigger。

在实际操作中，MIPI RFFE 中共有三个 Standard Trigger，分别为 Trigger0/1/2，受 Reg0x1C[5:0] 寄存器（见图5.233）控制：

1. 0x1C[5:3]：Trigger Mask，当 Trigger Mask 为 0 时，新写入值只更新到 shadow register



Register Class	Register Address(Dec)	Register Address(Hex)	Register Name	Data Bits	Bit Name	Default	Description	State Value	State	Broadcast Store ID & Broadcast Allowed	Trigger Support	Active Trigger	Masked Write Support	R/W
RFFE Reserved	28	0x1C	PM_TRIG	7	PWR_MODE[1], Operation Mode	1	Defines normal Operation and Secondary Operation Mod. When the slave is in Secondary Operation Mode, charge pumps shall be disabled and the slave shall be listening to the RFFE bus.	0	normal Operation Mode (ACTIVE)	Y				
								1	Secondary Operation Mode (LOW POWER)					
				6	PWR_MODE[0], State Bit Vector	0	Single bit Powered Reset. Writing a 1 resets all slave registers to their default settings including the Operation Mode bit in PWR_MODE[1]. Always reads back 0.	0	N action (ACTIVE)					
								1	Powered Reset					
				5	TRIGGER_MASK_2	0	Trigger Mask 2	0	Data writ to registers tied to TRIGGER_2 is held in shadow registers until the TRIGGER_2 bit is set to 1.	N				
								1	Data writ to registers tied to TRIGGER_2 go directly to the active registers.					
				4	TRIGGER_MASK_1	0	Trigger Mask 1	0	Data writ to registers tied to TRIGGER_1 is held in shadow registers until the TRIGGER_1 bit is set to 1.					
								1	Data writ to registers tied to TRIGGER_1 go directly to the active registers.					
				3	TRIGGER_MASK_0	0	Trigger Mask 0	0	Data writ to registers tied to TRIGGER_0 is held in shadow registers until the TRIGGER_0 bit is set to 1.		N	N	N	R/W
								1	Data writ to registers tied to TRIGGER_0 go directly to the active registers.					
				2	TRIGGER_2	0	Trigger 2. This bit has no effect if TRIGGER_MASK_2 is 1. When the part is in LOW POWER and a trigger request is sent in the same command sequence that still keeps the part in LOW POWER, the trigger request is ignored. This applies to Triggers only, not to Trigger Masks. When the part is in ACTIVE and sent to LOW POWER, or when the part is in LOW POWER and sent to ACTIVE, trigger requests in the same command sequence are not ignored.	0	No action. Data is held in shadow registers.	Y				
								1	Data is transferred from shadow registers to active registers for registers tied to TRIGGER_2.					
				1	TRIGGER_1	0	Trigger 1. This bit has no effect if TRIGGER_MASK_1 is 1. When the part is in LOW POWER and a trigger request is sent in the same command sequence that still keeps the part in LOW POWER, the trigger request is ignored. This applies to Triggers only, not to Trigger Masks. When the part is in ACTIVE and sent to LOW POWER, or when the part is in LOW POWER and sent to ACTIVE, trigger requests in the same command sequence are not ignored.	0	No action. Data is held in shadow registers.					
								1	Data is transferred from shadow registers to active registers for registers tied to TRIGGER_1.					
				0	TRIGGER_0	0	Trigger 0. This bit has no effect if TRIGGER_MASK_0 is 1. When the part is in LOW POWER and a trigger request is sent in the same command sequence that still keeps the part in LOW POWER, the trigger request is ignored. This applies to Triggers only, not to Trigger Masks. When the part is in ACTIVE and sent to LOW POWER, or when the part is in LOW POWER and sent to ACTIVE, trigger requests in the same command sequence are not ignored.	0	No action. Data is held in shadow registers.					
								1	Data is transferred from shadow registers to active registers for registers tied to TRIGGER_0.					

图 5.234: MIPI table from SmarterMicro S55255-12 L-PAMiF datasheet

Register Address	Register Name	Number of Bits	Notes
0x002D	EXT_TRIG_A_MASK[7:0]	8	Extended Triggers Masks Block A Bits 7:0 – EXT_TRIG_A_MASK<7:0>
0x002E	EXT_TRIG_A_REG[7:0]	8	Extended Triggers Register Block A Bits 7:0 – EXT_TRIG_A_REG<7:0>
0x002F	EXT_TRIG_B_REG[7:0]	8	Extended Triggers Register Block B Bit 7 – Reserved Bits 6:0 – EXT_TRIG_B_REG<6:0>
0x0030	EXT_TRIG_B_MASK[7:0]	8	Extended Triggers Masks Block B Bit 7 – Reserved Bits 6:0 – EXT_TRIG_B_MASK<6:0>

图 5.235: Extend Trigger 寄存器

2. 0x1C[2:0]：Trigger Reg，当 Trigger Mask 为 0，且 Trigger Reg=1 时，shadow register 的 value 会转移到 destination register

举例，假设寄存器 Reg01/02 被设置为 Trigger0 控制，那么如何进行 Trigger 控制：

- Trigger0 enable：
  - 设置 Trigger0 enable，写入 0x1C 寄存器值为 0bxxxx0xx0
  - 写入 reg01/02 值，值只能在 Shadow 寄存器中生效
  - 设置 Trigger0 生效，写入 0x1C 寄存器为 0bxxxx0xx1
  - Reg01/02 生效
- Trigger0 disable：
  - 设置 Trigger0 disable，写入 0x1C 寄存器值为 0bxxxx1xx0
  - 写入 reg01/02 值，直接生效

Extend Trigger

为了支持更多的寄存器 Trigger 数量，MIPI RFFE V2.1/3.0 分别各引入了 8 个 Trigger，工两组 Block A 和 BlockB，统称为 Extend Trigger。

Extend Trigger 的控制和 Standard Trigger 一样，对应的寄存器为 0x2D ~30，如图5.235。

需要注意的是，下文要讲的 Timed Trigger 和 Mappable Trigger 只能应用于 Extend Trigger，不能用于 Standard Trigger。

Timed Trigger

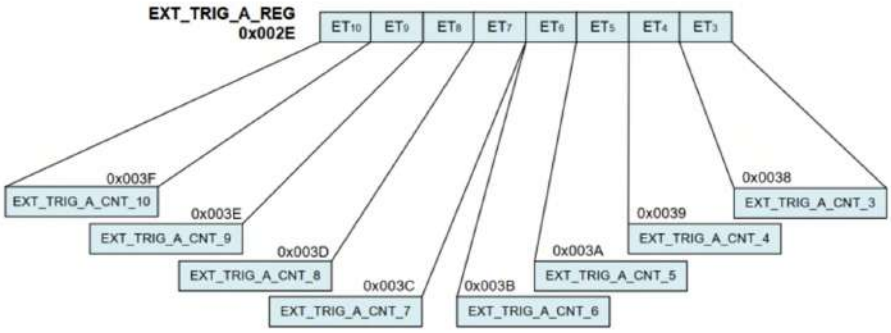


图 5.236: Counter register

OFDM Symbol Duration

Parameter / Numerlogy (u)	0	1	2	3	4
Subcarrier Spacing (Khz)	15	30	60	120	240
OFDM Symbol Duration (us)	66.67	33.33	16.67	8.33	4.17
Cyclic Prefix Duration (us)	4.69	2.34	1.17	0.57	0.29
OFDM Symbol including CP (us)	71.35	35.68	17.84	8.92	4.46

图 5.237: 写入寄存器时长

Timed Trigger 顾名思义是在时间上对 Trigger 进行操作，具体就是提供寄存器在延时一定时间后生效，主要机制是对 Trigger 设置延时计数器生效。

对应的延时计数器其实是由一个延时的计数寄存器进行控制，如下图是 BlockA Extend Trigger 对应的 Counter register。如图5.236。

Timed Trigger 的工作机制如下：

1. 当 counter register 的值 =0，则此 Extend Trigger 属于正常 Trigger
2. 当 counter register 的值为非零时，寄存器的数则跟随时钟开始倒计时，直到值 =0，Trigger 生效

Mappable Trigger

如上介绍，我们了解到每个寄存器有且最多只能配置对应一个固定的 Trigger，但随着 5G 的加入，射频前端越来越复杂，例如以下场景：

在 5G 手机中，可能有超过 8 路接收通路同时工作（CA 场景），此时就需要同时控制 8 路 LNA 进行快速的 AGC，每路 LNA 至少需要发送一个寄存器 (~0.5us)，然而 5G 的 AGC 切换是在 symbol 之间进行，而 symbol 与 symbol 之间只有一个 CP 的时间长度（当 SCS=60kHz，时间为 1.17us，如图5.237）来完成写入寄存器 +LNA gain settle 时间。因此要同时串行发送以上多路 LNA 是无法完成的。

Mappable Trigger 提供这样一个可能，对寄存器不配置固定的 Trigger，而是可编辑的 Trigger，这样通过把所要控制的所有寄存器都配置到相同的 Trigger，那么就可以在正式生效前坐 pre-setting，然后再同时 Trigger。

Mappable Trigger 具体操作如下：

1. 定义一个寄存器叫 mTrig set，4 个 bits 对应一个 BlockA/B Extend Trigger，如图5.238；
2. 一个 UDR 寄存器最多只能配置一个 mTrig set，或一个 Fixed Trigger，不同同时配置两种 Trigger；
3. 当某个寄存器被配置为 mTrig setA，且此时 mTrig setA 被写入成 0b0010 时，查表对应的 Extend Trigger 是 ET05，则 Extend Trigger05 为当前寄存器及所有配置为 mTrig setA 的寄存器对应的 Trigger；
4. 当 mTrig setA 的值更新后，对应的 Trigger 也会同时更新；
5. mTrig 不会影响 Timed Trigger，二者相互独立。

5.12.3 MIPI RFFE 协议的未来发展

随着协议演进，射频前端功能越来越复杂；同时射频前端器件与基带控制间的关系越来越紧密，如图5.239高通公司表示未来通信系统中 Modem 将越来越多的与射频前端进行联合设计。以上两点都对 RFFE 与 Modem 之

Table 9 mTrig Set Assignment for Block AB Mapping

Extended Trigger Block	mTrig Set Assignment Value <3:0>	UDR Set Trigger Mapping
Block A	0b0000	Extended Trigger 3
	0b0001	Extended Trigger 4
	0b0010	Extended Trigger 5
	0b0011	Extended Trigger 6
	0b0100	Extended Trigger 7
	0b0101	Extended Trigger 8
	0b0110	Extended Trigger 9
	0b0111	Extended Trigger 10
Block B	0b1000	Extended Trigger 11
	0b1001	Extended Trigger 12
	0b1010	Extended Trigger 13
	0b1011	Extended Trigger 14
	0b1100	Extended Trigger 15
	0b1101	Extended Trigger 16
	0b1110	Extended Trigger 17
	0b1111	Write to Destination (No Trigger Assignment)

图 5.238

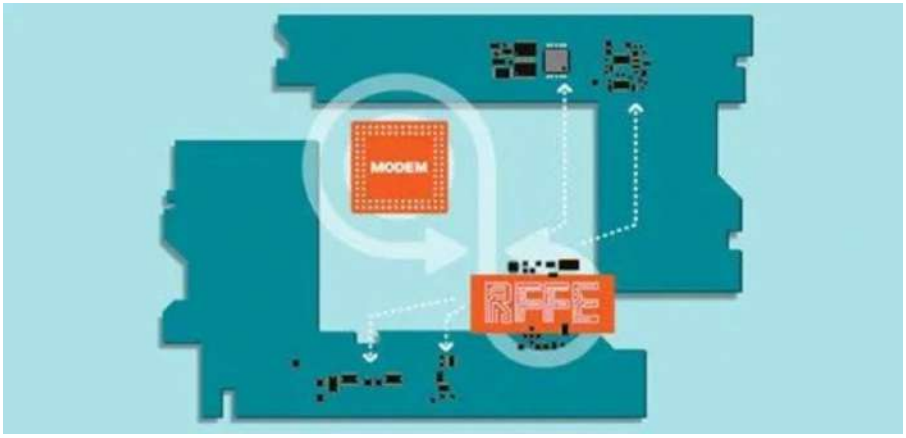


图 5.239: 高通公司将 Modem 与 RFFE 进行联合设计

间的接口：MIPI RFFE 接口提出更高的要求。

在 MIPI RFFE V3.0 发布之后，目前 MIPI 联盟 RFFE 工作组在讨论之中的更新有：

1. 加入 5G 通信中的新特性，甚至 6G 演进中的新需求。
2. 考虑 5G NR FR2（毫米波）频段的需求和挑战。
3. 会员单位提出的其他新需求。

可以预计，MIPI RFFE 的未来将更加强大。预计未来在应用中使用到的 MIPI RFFE 功能有：

1. MIPI RFFE 继续成为移动终端的标准通信接口，并且应用领域不断扩大。
2. Trigger 模式的灵活运用，以满足协议苛刻的时序要求。
3. 更多寄存器控制支持，以满足日益复杂的器件接口控制功能。

以上为慧智微在与 MIPI 联盟 RFFE 工作组工作，以及公司 5G 器件开发过程中对 MIPI RFFE 技术的了解，您在使用过程中对 MIPI RFFE 协议技术有什么理解或者问题，欢迎扫描文章底部二维码，参与在线留言讨论。

参考文献

1. MIPI RFFE V3.0 协议：mipi_RFFE_specification_v3-0.pdf
2. MIPI RFFE V2.1 协议：mipi_RFFE_specification_v2-1.pdf
3. MIPI Alliance website: https://www.mipi.org/
4. 慧智微产品规格书：SmarterMicro S55255-12 UHB-PAMiF Datasheet v1.74.pdf
5. http://www.sharetechnote.com/

在线阅读

本文摘自慧智微微信公众号，扫描以下二维码，在线阅读

5.13 为了更流畅地刷抖音，射频工程师都做了什么

作者：剑文

移动通信的发展一直围绕着信号传输速率的提升演进。2G 时代几十 Kbps 的速度只能支持看文本小说，3G 时代几 Mbps 的速度已经可以逛论坛斗图了，到了 4G 时代几百 Mbps 的高速网络让人们可以随心所欲的刷抖音，开视频会议。而 5G 时代，基于 VR/AR 的元宇宙蓝图已经展开。

为了让人们能更流畅的上网，射频通信工作者们可谓是操碎了心。

跟据香农定理，信号传输速率：

$$R = B\log_2(1+\frac{S}{N})$$

其中 R 为信号速率，B 为信号带宽，S/N 为信噪比。

考虑到现在的调制编码方式一个码元可以携带多 bit 信息，MIMO 技术在同一频段同时传输不同数据，因此可以将香农定理修正为：

$$R = mB\log_2(1+\frac{S}{N})\log_2 M$$

其中 m 为 MIMO 通道数，M 为码元状态数。

从上式不难看出，信号速率主要跟以下三项有关：

- 信号带宽
- 信噪比
- 调制方式

5.13.1 信号带宽

5.13.1.1 向高频大带宽出击

2G 时代 GSM、GPRS、EDGE 布局在 2GHz 以下，只有 0.2MHz 的带宽，能传输的数据实在有限。3G 时代的 WCDMA、TD-SCDMA、CDMA2000 将频率扩展至 2GHz 以上，已经开始支持单载波 5MHz 带宽。4G 时代频率已经延伸至 2.7GHz，单载波最高达到 20MHz。5G 时代将频率延伸至 5GHz 甚至毫米波，直接将单载波带宽扩展到 100MHz 以上，显著提高了信号传输速率。

		典型带宽B（MHz）	调制方式
2G	GSM	0.2	GMSK
	GPRS	0.2	GMSK
	EDGE	0.2	8PSK
3G	WCDMA	5	BPSK、QPSK、16QAM
	TD-SCDMA	1.6	QPSK、8PSK、16QAM
	CDMA2000	1.25 、3.75	BPSK、QPSK、8PSK、16QAM
4G	FDD LTE	1.4、3、5、10、15、20	QPSK、16QAM、64QAM、256QAM
	TD-LTE	1.4、3、5、10、15、20	QPSK、16QAM、64QAM、256QAM
5G	NR	5、10、15、20、25、30、40、50、60、80、100	QPSK、16QAM、64QAM、256QAM

图 5.241: 2/3/4/5G 典型带宽与调制方式

带宽的增加和高频的扩展，都为射频前端，特别是射频 PA 提出了新的挑战。PA 需要采用新的架构和工艺来支持更高的频率，更大的带宽。

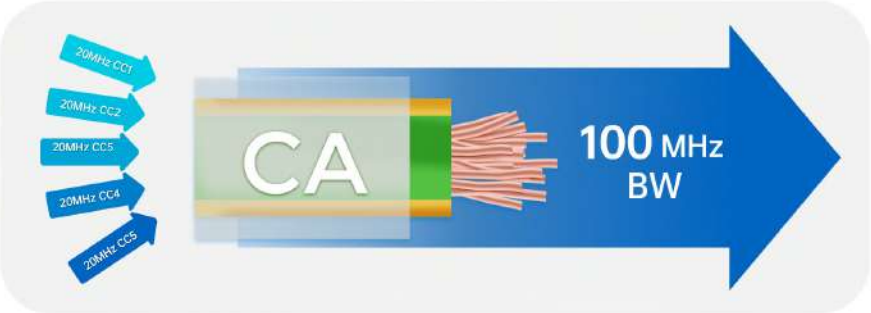


图 5.242: CA 示意图

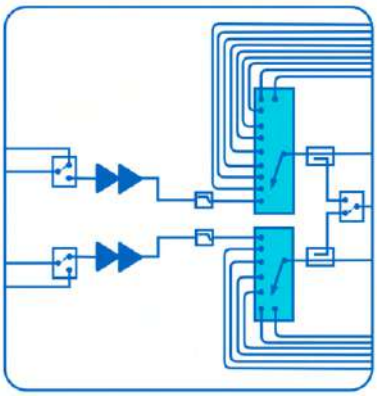


图 5.243: 慧智微 S2592x Phase3 TXM

5.13.1.2 物尽其用

Sub 3G 的频段相当拥挤，有效带宽 2.2GHz 的频谱内分布了 50 多个 Band，除了 B40/41 这种高频，其他频段因为中心频率低，带宽天生较窄。为了在有限的带宽内让这些频段尽可能地发光发热，为香农定理的第一项做出贡献，通信工作者们发明了载波聚合技术（CA，Carrier Aggregation），如图5.242。

带内连续 CA，是将较为宽带的一个频段内没被用上的那一段再利用起来，一起传输，相当于将带宽加倍。

带内非连续 CA，与前者相比，两段载波有一定的频率间隔，但一般不大。仍然是在同一个 Band 内操作。

带间 CA，将两个不同的频段同时传输，可以理解为一辆五菱宏光和一辆 AE86 一起秋名山过弯漂移，传输带宽瞬间飙升。

由于带间 CA 需要两个频段同时工作，所以射频前端对天线选择模块提出了新的要求。一种方式是采用 diplexer 分频器，将一根天线下的两段频率通过滤波分成两路同时传输；另一种则是拆分天线，将两段频率分别接不同天线。因此天线选择开关也需要相应的调整，由原来的 SPXT 改为 DPXT 或者 3PXT。支持带间 CA 的发射模组我们也称之为 Phase 3 TxM，如图5.243。

因为两个 PA 同时工作难度较大，因此在上行时增加带宽往往采用带内 CA 的方式。对于下行，可以灵活应用拆分天线、diplexer 分频和开关双开，最多可以支持 5 载波 CA，带宽直接拉满。

5.13.2 信噪比

带宽和调制两项，系统都已经规定好了，信噪比 S/N 是通信工作者少有的能通过个人能力为人民造福的地方。

5.13.2.1 降低噪声提升单路信噪比

信号在传输过程中的每一步，都会增加新的噪声，不管是经过天线连接口、开关、滤波器、PCB 走线，还是经过 LNA，噪声系数 NF 只增不减。但是接收机内部也存在固有噪声系数，一般为 2 ~9dB，如果能将平台自带

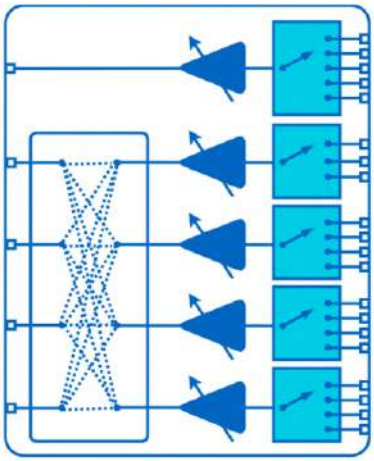


图 5.244: 慧智微 S1520X 系列 LNA BANK

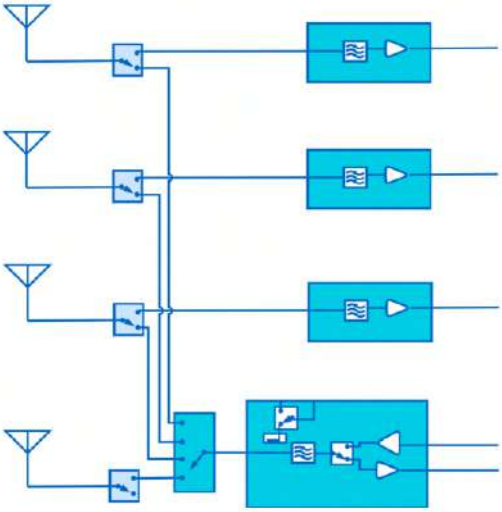


图 5.245: 慧智微 S552XX+S1572X 系列 Sub 6G 模组

的噪声系数削弱，系统信噪比是不是就能提高呢？因此射频工作者发明了外置 LNA，即低噪声放大器。将 LNA 靠近天线放置可以有效抑制 LNA 后端的插损和噪声系数，相比不加时，系统 NF 少下降约 2dB，等同于将系统信噪比提升 2dB。外置 LNA 对提升接收灵敏度功效显著，4G/5G 手机上现在基本是标配。LNA Bank 的系统框图如图5.244所示。

5.13.2.2 增加分集增强信号

3G 以前的时代，信号都是一发一收，不求能上网，只求可以打电话。但是到了 4G 时代，互联网应用层出不穷，人们也需要除了打电话之外的其他精神娱乐，对网速的需求非常迫切，因此诞生了接收分集 DRx（Diversity Receive），传输和主接收 PRx（Primary Receive）一样的数据，信号加倍而噪声因为不相干被部分消除，增强了信噪比。

后来短视频和网剧的兴起对流量又提出了新的要求，通信工作者们提出 MIMO 技术，多条路径一起传输不同数据，从早期的 2*2 MIMO，到现在的 4*4 MIMO，无需增加带宽即可极大的提升信号传输速率。当然 MIMO 也对接收器件的数量和隔离度提出了新的要求。如图5.245所示。

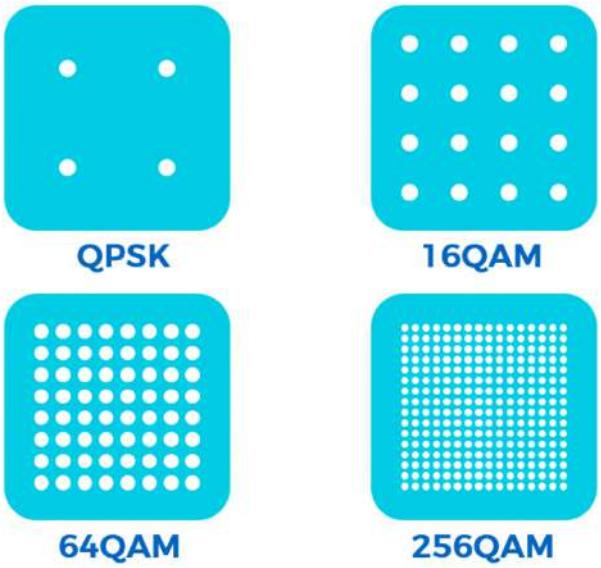


图 5.246: 调制星座图

5.13.3 调制方式

早期的 BPSK 调制一个码元符号只能携带 1bit 数据，后来出现 QPSK，一个码元符号可以携带 2bit 数据，再到后来的 256QAM，一个码元符号可以携带 8bit 数据。在前方帮忙趟坑开路的 Wi-Fi 老大哥已经商用 1,024QAM 了，一个码元符号携带 10bit 数据，接下来的 Wi-Fi 7 据说要到 4,096QAM，一个码元符号携带 12bit 信息。等 Wi-Fi 支持 4,096QAM 调制的时候，咱们移动通信也会大规模支持 1024QAM，到时候网速又可以再上一个台阶。调制星座图如图5.246所示。

5.13.4 物理通道速率计算

前面讲完理论基础，定性之后咱们来定量看一下物理通道的理论速率极限。

协议分配的带宽两侧都会留好保护间隔，因此可以传输数据的带宽要打个折扣，一般射频通信工程师会按 RB 计算实际有效数据带宽。LTE 的 1RB= 15KHz * 12 = 180KHz，满带宽 100 RB 则有 18MHz 有效带宽。而 Sub 6G NR 的子载波间隔可以到 30KHz，因此 1RB=30KHz*12=360KHz，满带宽有 273 RB 即 98.28MHz 有效带宽。

按香农公式理论计算，LTE 20MHz 带宽下 64QAM 单路传输信道理论速率极限为：

$$R = \frac{18 \times 10^6 \times \log_2)(1+1) \times \log_2 64}{10^6} = 108\ Mbps$$

如果有看官对此计算有疑虑，咱们不妨从码元传输角度再来计算一遍。

RB 结构示意图如图5.247所示。1ms 的时长可以发送 2 个 slot，一个 slot 满带宽发 100RB，一个 RB 由 12 个子载波、7 个码元符号组成，一个码元符号 64QAM 下携带 6bit 信息，因此 LTE 20MHz 带宽下 64QAM 单路传输理论速率极限为：

$$R = \frac{1000 \times 2 \times 100 \times 12 \times 7 \times \log_2 64}{10^6} = 100.8\ Mbps$$

以上计算结果与香农公式理论计算结果 108Mbps 相差 7.2Mbps，我说什么来着，你小子的算法果然有问题！

桥豆麻袋！实际编码的时候码元符号最前面会加上循环前缀（CP，Cyclic Prefix）。在 OFDM 符号前面加上 CP 可以消除符号间的干扰，但这个并不携带我们要传的数据，因此真正传输的数据还要去掉 CP 占用的时间。每个 symbol 前面都有 CP，除第一个 CP 占 160 个采样周期，其他 6 个 CP 只有 144 个采样周期，而一个 OFDM 符号有 2048 个采样周期，因此 CP 在时隙中占比为：

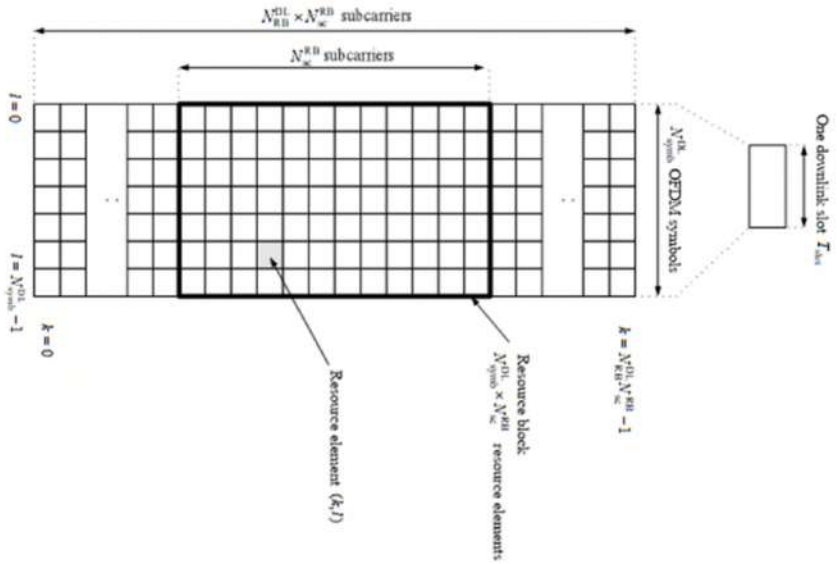


图 5.247: RB 结构示意图

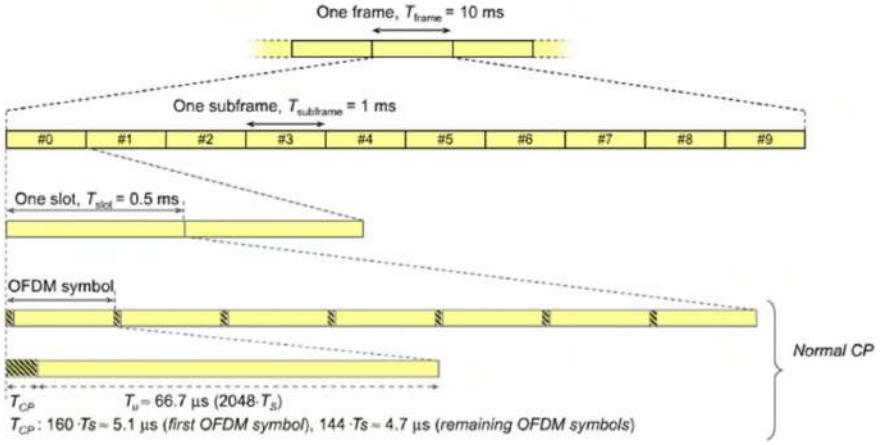


图 5.248: 时域子帧结构图

$$\frac{160+144\times 6}{160+144\times 6+2048\times 7}=\frac{1}{15}$$

这个就是物理信道中被“浪费”的带宽。时域子帧结构如图5.248所示。

我们再将按频域计算的结果，去掉时域中被 CP 占据“浪费”的部分，108*14/15=100.8(Mbps) 。

与香农公式理论计算完美吻合！

由于传输的符号中，还有大量的控制信息，不能用来传数据。这些资源开销在协议规定为 5G 下行 0.14，上行 0.08，4G 上下行均为 0.2177。此外 256QAM 调制最大码率为 948/1024≈0.92578。

因此 4G 下行有效传输数据比例为：

$$(1-0.2177)\times 0.92578=72.4\%$$

5G 下行有效传输数据比例为：

$$(1-0.14)\times 0.92578=79.6\%$$

发到用户手上的真实数据还需要乘上这个系数，即 $R'=R\times 79.6\%$。

手机真正使用时的网速，业内更专业地称之为吞吐率，受多种因素影响。



图 5.249: 实测 4G（左）与 5G（右）下载速率

在实际使用中，想必大家经常遇到 5G 手机有时候刷剧也很卡，缓冲半天；4G 手机有时候反而网速飞快，APP 秒下。一方面是因为基站分配的带宽限制，另一方面则是因为高端 4G 手机配置很高，网速一点都不输 5G。

拿一台支持 5 载波 CA 的高端 4G 手机为例，理想状态下，下行 256QAM 调制，4*4 MIMO，5CC CA 聚合为 100MHz 带宽。此时物理信道理论极限数据率：

$$R = \frac{4 \times 90 \times 10^6 \times \log_2(1+1) \log_2 256 \times \frac{14}{15} \times 0.724}{10^6} = 1,946\ Mbps$$

一台普通的 4G 手机，正常情况下行 64QAM 调制，2*2 MIMO，10MHz 带宽。此时物理信道理论极限数据率：

$$R = \frac{2 \times 9 \times 10^6 \times \log_2(1+1) \times \log_2 64 \times \frac{14}{15} \times 0.724}{10^6} = 73\ Mbps$$

速度也是相当的快。

而 5G 手机正常使用时，下行 256QAM 调制，4*4 MIMO，10MHz 带宽。此时物理信道理论极限数据率：

$$R = \frac{4 \times 9.8 \times 10^6 \times \log_2(1+1) \log_2 256 \times \frac{14}{15} \times 0.724}{10^6} = 230\ Mbps$$

和 4G 手机不相上下。

而当 5G buff 叠满时，CA 聚合出 200MHz 带宽，4*4MIMO，256QAM 调制。此时物理信道理论极限数据率：

$$R = \frac{4 \times 196.56 \times 10^6 \times \log_2(1+1) \log_2 256 \times \frac{14}{15} \times 0.724}{10^6} = 4,600\ Mbps$$

光速下载。

当然这些都是基于没人跟你抢基站用的假设，实际使用时，每个用户被分配到的带宽相当有限。比如大型商场里人满为患，一个用户被分配到的带宽可能只有 1.4MHz，调制降到 QPSK，这时网速只有 1Mbps，你要是玩韩信一个后跳直接就闪进东皇的大。

图5.249为小米 11 国内版手机分别采用 4G 与 5G 连接测试图示，可以看到，4G 连接时下载速率为 89.47 Mbps，而 5G 连接时下载速率为 236.16 Mbps，与前文计算基本一致。用户使用时被分配的带宽一样的情况下，5G 一般调制阶数更高，MIMO 路数更多，虽无法体现数量级的差距，但还是比同等情况下的 4G 速率更快。

5.13.5 总结

咱们每天把玩的平平无奇的手机，汇聚了无数射频通信工作者的智慧结晶，为了将网速做快，方便大家追剧刷抖音，手机上充分整合利用了各种 CA、MIMO、高阶调制，层层 buff 叠在一起才有了现在高达几个 Gbps 的速度。未来射频仍将引领移动通信向更高更快更强演进。

发射机的 EVM 通常在 10% 以内，对应的 SNR 在 15 到 30dB，而计算信道理论速率极限的信噪比却为 0dB，其间接收和发射信噪比的差异欢迎大家扫码参与在线留言讨论。

在线阅读

本文摘自慧智微微信公众号，扫描以下二维码，在线阅读

5.14 蜂窝移动终端的 Cat 指的是什么？

作者：大强子

在蜂窝移动终端设备中，经常看到以“Cat x”来对标称设备的通信能力。除了最近如日中天的“Cat.1”，还听到如 Cat 4、Cat 6、Cat 12 等制式标准。

如图5.251分别为物联网模块厂商 Quectel（移远通信）的 LTE 模块产品，产品规格从 Cat 4 一直到 Cat 18。

Quectel LTE Module Family

Product	EC20 R2.1	EG06	EG12	EG18
Picture				
Form Factor	LCC	LGA	LGA	LGA
Dimensions (mm)	32.0 × 29.0 × 2.4	37.0 × 39.5 × 2.8	37.0 × 39.5 × 2.8	37.0 × 39.5 × 2.8
4G	LTE Cat 4	LTE Cat 6	LTE Cat 12	LTE Cat 18

图 5.251: 物联网模块公司 Quectel（移远通信）的 LTE 模块产品 [1]

不止是物联网模块产品，手机产品也有“Cat”的分类，如图5.252，根据 Opensignal 对历代 iPhone 的统计[2]，LTE 时代 iPhone 的通信能力由 Cat 3，历经 7 年一直提升至 Cat 19，之后再进入 5G 时代。

不同的“Cat”究竟指的是什么？具体是怎么实现的？不同“Cat”的射频前端方案有什么不同？

5.14.1 “Cat”的由来：LTE UE 终端能力等级定义

3GPP 组织制定 LTE 标准的时候，有一个初衷，就是希望 LTE 能为不同用户提供不同等级的网络服务能力。所以，LTE 里提出了 ue-Category 的概念：不同的业务使用不同的 ue-Category，网络提供不同的带宽和速率。

而“Cat”就是 Category（类别）的缩写。根据 3GPP 的定义，ue-Category 代表的就是 User Equipment（用户设备）等级，表明了终端所支持的上行和下行数据传输能力。ue-Category 可简称 UE Cat，也叫终端能力等级。以如下表格为例，3GPP 定义的主要参数如图5.253。

正如 LTE 的名字缩写的由来：Long Time Evolution（长期演进）一样，4G LTE 技术规范随着版本一次又一次的更新而演进，每次更新都会引入一些新技术和新 UE Cat 值。4G 最早版本是 Release 8（R8），后续接着还有

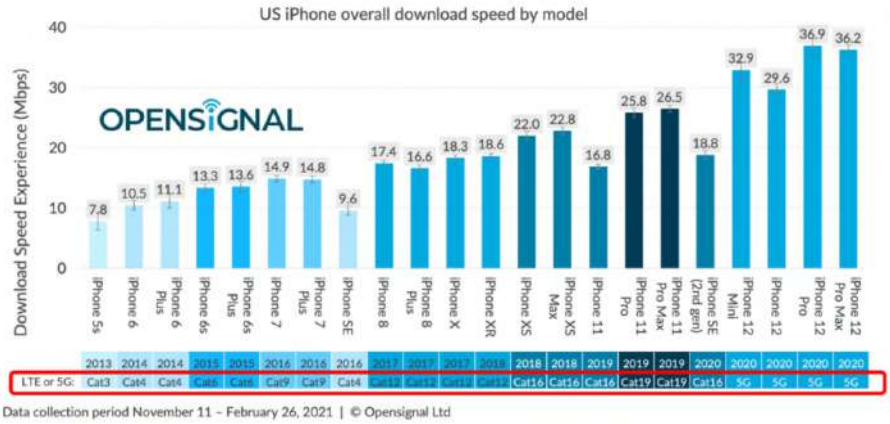


图 5.252: Opensignal 对历代 iPhone 通信能力统计 [2]

Table 4.1-1: Downlink physical layer parameter values set by the field *ue-Category*

UE Category	Maximum number of DL-SCH transport block bits received within a TTI (Note 1)	Maximum number of bits of a DL-SCH transport block received within a TTI	Total number of soft channel bits	Maximum number of supported layers for spatial multiplexing in DL
Category 1	10296	10296	250368	1
Category 2	51024	51024	1237248	2
Category 3	102048	75376	1237248	2
Category 4	150752	75376	1827072	2
Category 5	299552	149776	3667200	4
Category 6	301504	149776 (4 layers, 64QAM) 75376 (2 layers, 64QAM)	3654144	2 or 4
Category 7	301504	149776 (4 layers, 64QAM) 75376 (2 layers, 64QAM)	3654144	2 or 4
Category 8	2998560	299856	35982720	8
Category 9	452256	149776 (4 layers, 64QAM) 75376 (2 layers, 64QAM)	5481216	2 or 4
Category 10	452256	149776 (4 layers, 64QAM) 75376 (2 layers, 64QAM)	5481216	2 or 4
Category 11	603008	149776 (4 layers, 64QAM) 195816 (4 layers, 256QAM) 75376 (2 layers, 64QAM) 97896 (2 layers, 256QAM)	7308288	2 or 4
Category 12	603008	149776 (4 layers, 64QAM) 195816 (4 layers, 256QAM) 75376 (2 layers, 64QAM) 97896 (2 layers, 256QAM)	7308288	2 or 4

NOTE 1: In carrier aggregation operation, the DL-SCH processing capability can be shared by the UE with that of MCH received from a serving cell. If the total eNB scheduling for DL-SCH and an MCH in one serving cell at a given TTI is larger than the defined processing capability, the prioritization between DL-SCH and MCH is left up to UE implementation.

图 5.253: 3GPP 中所定义的 ue-Category

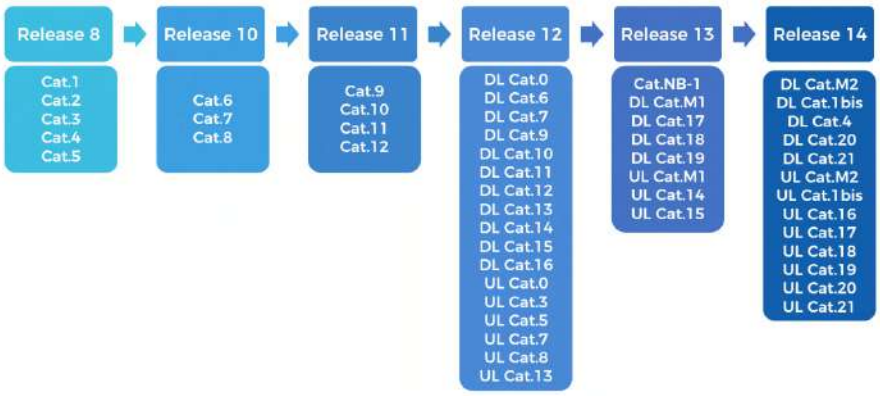


图 5.254: 3GPP 定义的 UE Cat 能力等级



图 5.255: 3GPP Release 版本和 UE Cat 类型关系

R9，R10，R11，R12，R13，R14 等等（从 R15 开始该定义 5G UE 的能力级别）。如图5.254整理了 3GPP 协议从 R8 到 R14 定义的 UE Cat 值。

从 R8 到 R11，3GPP 定义了从 Cat 1 到 Cat 12 的 UE 能力级别。到了 R12，定义的能力级别前面开始带了 DL 和 UL。DL 就是 Downlink 的缩写，表示下行，也就是表示 UE 下载的能力，UL 是 Uplink 的缩写，表示上行，也就是表示 UE 上传的能力。如图5.255所示。

在 Cat 1 到 Cat 12 定义中，上下行的能力是绑定在一起的，即定义中同时包含上行和下行能力。以 Cat 4 为例，下行支持 150Mbps，上行支持 50Mbps。要改变上下行任意一个能力，都必须重新定义，一点也不灵活。

而实际应用中，可能需要下载的速率高一点，但对于上行的需求一般，那么把下载能力和上传能力绑定在一起就没这个必要了。因此，从 R12 开始，对于 UE Category 的定义就区分开了 UL 和 DL，上传速率用 UE UL Cat X，下载速率用 UE DL Cat Y，X 和 Y 之间允许存在协议规定的配置的组合。

由 R8 至 R12，以及 R13 之后所定义的 Cat 类别及能力统计如图5.256、图5.257及图5.258三张表所示：

同时，3GPP 在 R12 及后续版本也定义了 UE DL Category 和 UE UL Category 的组合，这些组合需要满足 3GPP 中 TS 36.303 协议规范。组合配置如图5.259。

5.14.2 不同 Cat 的物理实现

对于移动终端，影响通信能力的主要参数有：

- UE 支持的 MIMO 层数
- UE 支持的载波聚合数量
- UE 支持的最高阶调制方式

以上三个参数的不同配置的就可以得到不同速率，对应不同上行和下行数据传输能力。LTE 协议中的不同 Cat 正是通过调整以上三个配置，实现不同的 Cat 的传输速率。

载波聚合：

载波聚合（Carrier Aggregation, CA）是将 2 个或更多的载波单元（Component Carrier, CC）聚合在一起以支

Cat.	Rel	Downlink				Uplink			
		最大传输速率 Mbit/s	最大支持 MIMO 层数	支持 256 QAM	载波聚合数目	最大传输速率 Mbit/s	最大支持 MIMO 层数	支持 64QAM	载波聚合数目
0	R12	1	1	no	1	1	1	no	1
1	R8	10	1	no	1	5	1	no	1
2	R8	51	2	no	1	25	1	no	1
3	R8	102	2	no	1	51	1	no	1
4	R8	150	2	no	1	51	1	no	1
5	R8	299	4	no	1	75	1	yes	1
6	R10	301	2或4	no	1或2	51	1	no	1
7	R10	301	2或4	no	1或2	102	1或2	no	1或2
8	R10	2998	8	no	5	1497	4	yes	5
9	R11	452	2或4	no	2或3	51	1	No	1
10	R11	452	2或4	no	2或3	102	1或2	no	1或2
11	R11	603	2或4	可选	2或4	51	1	no	1
12	R11	603	2或4	可选	2或4	102	1或2	no	1或2

图 5.256: R8 至 R12 对应 Cat 能力等级（上下行绑定）

Cat.	Rel	Downlink			
		最大传输速率 Mbit/s	最大支持 MIMO层数	支持 256QAM	载波聚合数目
M1	R13	1	1	no	1
M2	R14	4	1	no	1
0	R12	1	1	no	1
1bis	R14	10	1	no	1
4	R14	150	2	no	1
6	R12	301	2或4	no	1或2
7	R12	301	2或4	no	1或2
9	R12	452	2或4	no	2或3
10	R12	452	2或4	no	2或3
11	R12	603	2或4	可选	2或4
12	R12	603	2或4	可选	2或4
13	R12	391	2或4	yes	1或2
14	R12	3916	8	yes	5
15	R12	749-798	2或4	可选	2-5
16	R12	978-1051	2或4	可选	3-7
17	R13	25065	8	yes	32
18	R13	1174-1206	2或4或8	可选	2-8
19	R13	1566-1685	2或4或8	可选	2-11
20	R14	1948-2019	2或4或8	可选	3-13
21	R14	1348-1413	2或4	可选	4-9

图 5.257: R12 及之后版本定义的下行 Cat 等级

Cat.	Rel	Uplink			
DL Cat		最大传输速率 Mbit/s	最大支持 MIMO层数	支持 64QAM	载波聚合数目
M1	R13	1	1	no	1
M2	R14	6	1	no	1
0	R12	1	1	no	1
1bis	R14	5	1	no	1
3	R12	51	1	no	1
5	R12	75	1	yes	1
7	R12	102	1或2	no	1或2
8	R12	1497	4	yes	5
13	R12	150	1或2	yes	1或2
14	R13	9585	4	yes	32
15	R13	226	1或2	yes	2或3
16	R14	105	1	yes	1
17	R14	2119	4	yes	5
18	R14	211	1或2	yes	1或2
19	R14	13563	4	yes	32
20	R14	316	1或2	yes	2或3
21	R14	301	1或2	yes	2或4

图 5.258: R12 及之后版本定义的上行 Cat 等级

DL/UL Cat等级组合	
DL Cat	UL Cat
M1	M1
M2	M2
0	0
1	1
4	5
6	5,16
7	13,18
9	5,16
10	13,18
11	5,16
12	13,15,18,20
13	3,5,7,13,16,18
14	18,7
15	3,5,7,13,16,18
16	3,5,7,13,15,16,18,20
17	14,19
18	3,5,7,13,15,16,18,20
19	3,5,7,13,15,16,18,20,21
20	3,5,7,13,15,16,18,20,21
21	3,5,7,13,15,16,18,20

图 5.259: R12 及之后版本对应 DL/UL Cat 能力等级组合



图 5.260: 载波聚合技术说明

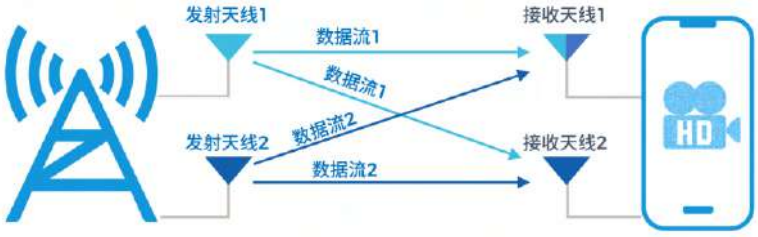


图 5.261: MIMO 技术说明

持更大的传输带宽（最大为 100MHz）。

UE 终端使用多个 4G 载波进行传输，相当于是把可用的频谱资源扩大了多倍。如图5.260以 3 载波聚合为例，单载波能达到下载 150Mbps 的速率，3 载波聚合就能达到 450Mbps 的速率。

多流传输

多流传输也被称作 MIMO（Multi input, Multi output 多输入多输出），本质利用信号在空间传播的路径不同，在多个路径同时发送多路数据，每一路数据成为一个流。常见的应用是 2 流和 4 流，UE 的速率相比单流来说直接变成 2 倍或 4 倍。如图5.261。

高阶调制

高阶调制就是在同样的信号里承载更多的数据。可以理解成，同样的一个车道，当你只有一辆车在运输时，运送的货物是有限的，当你有多辆车同时在运输货物，这样你就可以运输更多的货物。同理我们信号传输也是，单位时间内发送的数据多了，速率当然也就提升了。

在基站通信里，调制就好像在传输信道里面添加更多的信息。一般调制用 X-QAM 表示，X 越大阶数越高。从图5.262我们可以看出，256-QAM 相比 64-QAM，同样的一个码元包含的 bit 数，由 6 个变成了 8 个，在码元速率一样的情况下，速率提升了 33%。

以上三种技术的结合，就能实现不同 UE 的能力级别。

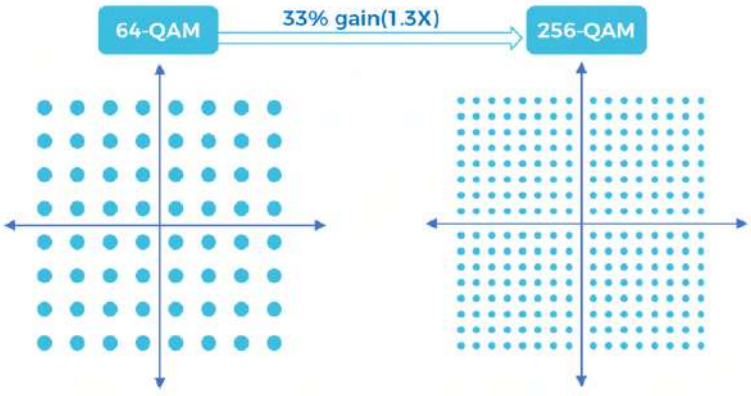


图 5.262: 64-QAM 及 256-QAM 调制技术说明

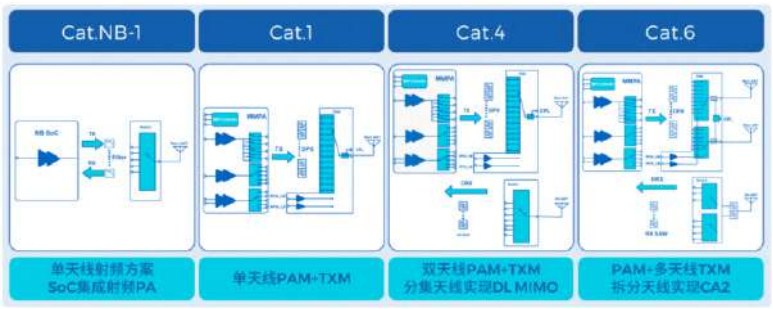


图 5.263: 典型 Cat 等级射频前端方案

需要说明的是，UE 的通信能力并不是越强越好，还需要结合成本、功耗综合考虑。比如在 5G 协议中，厂商先后提出了 5G Redcap 低能力终端，以及 5G 低功率终端方案，用来通过降低终端通信能力，来降低功耗、节省成本。

5.14.3 不同 Cat 的射频前端方案

由于不同 Cat 能力等级的技术特点，其前端射频方案的实现形式也不相同。典型的 Cat 等级对应前端射频方案如图5.263。

- Cat.NB-1：
  - 采用单天线射频前端方案
  - Half duplex 工作模式和较低的前端损耗使得对 PA 功率要求降低，甚至可以集成在 SoC 内部实现
  - 高度集成的方案还把滤波器和开关包含进去
- Cat 1（又写作 Cat.1）：
  - 采用单天线射频前端方案，无分集接收
  - 发射一般 4G MMPA+TXM 实现 Cat.1 的多频段 band 支持
  - Cat.1 only 的场景下 TXM 可用 Switch 替换，去掉 2G PA
- Cat 4：
  - 采用双天线射频前端方案，主集 + 分集
  - 主路采用 4G MMPA+TXM 实现
  - 增加分集接收实现 DL MIMO
- Cat 6 及以上：
  - 方案较为复杂，根据终端网络支持能力需增加不同方案
  - 多层数 MIMO 的需求，需要相对应天线数量实现上下行 MIMO 的支持
  - 拆分 TXM 天线实现频段间的 CA 功能，增加 DL/UL 传输速率

5.14.4 结语

进入 4G/5G 时代以来，人们越发发现不同使用场景对无线通信网络的需求是不同的，于是在 4G 时代，协议开始提出不同通信能力的标准，来满足不同场景的需求。3GPP 对不同通信能力的终端以“Cat”（缩写自 Category，类别）划分。本文介绍了不同 Cat 的由来及使用到的技术，以及实现方案。

随着 5G 到来，无线通信设备的应用场景将更加多样化，可能还会出现更多的协议细分。期待和您一起理清纷杂协议的相互关系，共建智能的无线互联世界。

针对以上问题您有更多的见解吗？欢迎扫码在线留言讨论。

参考文献

1. https://www.quectel.com/
2. https://www.opensignal.com/
3. https://www.3gpp.org
4. https://www.cablefree.net/ wireless technology/ 4g lte/lte-ue-category-class-definitions
5. 2021 Cellular Broadband Device & ModuleMarket, Techno Systems Research, 2021.

在线阅读

本文摘自慧智微微信公众号，扫描以下二维码，在线阅读

5.15 5G 物联网技术之“5G RedCap”

作者：大强子

5G 里的“RedCap”并不是“红帽子”的意思。

自 2019 年 4 月首个 5G NR 网络推出以来，5G 标准已在全球广泛采用。根据 GSMA 数据显示，到 2021 年底，全球 5G 连接数达到 6.4 亿。然而，大多数 5G 连接主要集中在消费者智能手机和固定无线接入（FWA）产品上。与智能手机采用 5G 相比，物联网（IoT）采用 5G 的速度一直相对缓慢。图5.265为 TSR 公布的非手机物联网出货统计及预测，可以看到自 2021 年之后，5G 物联网才开始有较明显的出货 [1]。

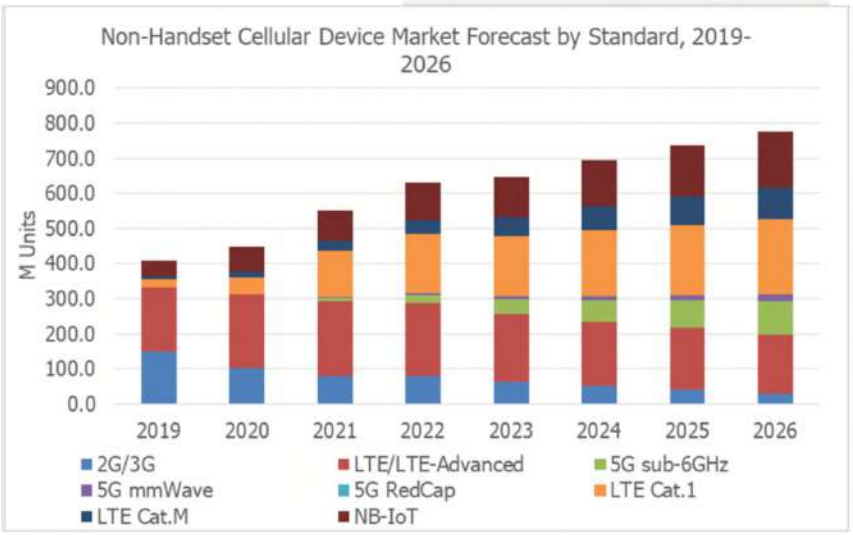


图 5.265: 非手机蜂窝物联网出货（以协议划分）[1]

与智能手机所处的 eMBB（增强移动宽带）场景不同，许多物联网应用并不需要大带宽性能。相反，不少用例要求比 5G eMBB 场景更小的带宽、更低的功耗，以及更低的成本。

为了支持这一场景物联网设备使用，第三代伙伴计划协议（3GPP）一直在开发一种名为轻型版 5G 标准：5G RedCap（5G Reduced Capability），以满足物联网这一需求。

5G RedCap 究竟是什么？为什么开发 5G RedCap？5G RedCap 商用的时间点是什么？带着以上问题，本文将对物联网 5G RedCap 技术及应用展开介绍。

5.15.1 什么是 5G RedCap

5G RedCap 指的是 3GPP 所提出的 5G 标准，与之前发布的 5G 标准相比，功能更加精简。5G RedCap 于 2019 年 6 月首次被纳入 3GPP Release17 研究项目 [2]。

5G RedCap 最早被称之为“低复杂度 NR 设备”，之后曾被命名为 NR-lite、NR-light 或 IWSN（工业无线传感网络）。**在现在大多数 3GPP 文档中，这个术语被标准化为“5G RedCap NR”，也就是** 5G Reduced Capability NR。

5.15.2 为什么是 5G RedCap

从当前 5G 覆盖的业务场景来看，5G 定义了三大场景：主要针对大带宽应用的 eMBB 场景，主要针对高可靠超低时延应用的 URLLC 场景，主要针对低速率、大连接的物联网应用的 mMTC 场景。

5G 应用场景的定义看似全面，但其实一些场景需求依然没有覆盖到。如：

- 工业无线传感网：要求对于通信服务可靠性达 99.99%，端到端时延小于 100ms，但对于速率的要求不高；

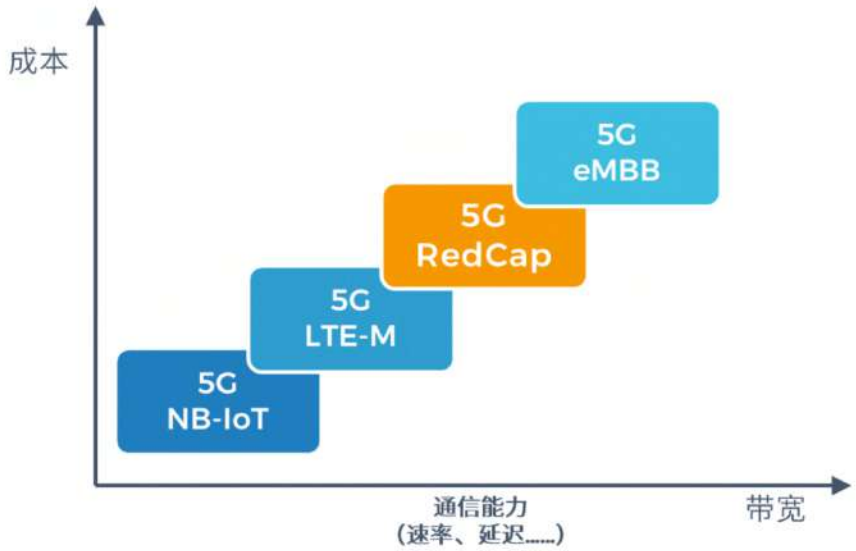


图 5.266: 5G 技术方案能力对比

- 可穿戴设备应用：对于速率的要求下行在 150Mbps，上行最高 50Mbps，属于中高速率要求，但是对于功耗、设备电池使用时长及尺寸等有高要求，需要整体方案足够精简。

显然，以上用例对于通信能力的要求要高于 NB-IoT/Cat-M，但要低于 eMBB 和 URLLC，并且方案需要足够简化。于是，针对这些物联网应用，3GPP 着手定义“精简版”5G 场景，即 5GRedCap。5G 技术方案能力对比如图5.266。

5G RedCap 的定义还有一个因素是来自于成本的考量，5G 物联网要普及开来，就需要把终端价格降下来。当前市面上 5G 模组价格在大几百甚至上千元，价格因素成为 5G 模组大规模部署的阻力。

把一些不必要的功能去掉就可以带来模组价格的降低。2021 年 12 月中国移动召开移动物联网研讨会，其会议主题就是关于 5G 模组成本优化演进，重点提到了 5G RedCap 的规划和布局，即利用 5G Rel-17 协议中的 RedCap 新功能助力模组终端成本优化 [3]。根据预测，5G RedCap 的模组价格未来会控制在 100 ~200 元（人民币）之间，相较于当前 5G 模组近 1000 元人民币的单价有明显的降低。

综上，5G RedCap 背后的基本想法是：为物联网应用定义一种新的、不那么复杂的 NR 设备，它可以提供比 LTE Cat-M 和 NB-IoT 更快的数据传输速度，但又比现在部署的传统 5G NR 设备简单便宜得多。一句话：够用就好。

此外，3GPP RedCap 工作项目组正在努力复用当前 5G 基站硬件的规范制定，同时为 RedCap 设备提供与 5G NR 设备相同的覆盖范围。其目标是使 RedCap 设备易于部署，只要基于对现有移动网络运营商（MNOs）NR 网络进行软件升级，就能够为 RedCap 设备提供连接。

5.15.3 5G RedCap 有什么不同

根据最新 3GPP Release17 协议标准，5G RedCap 在降低设备复杂度、降低成本方面主要有以下几个功能特点：

- 最大带宽更窄
- 下行接收分支数更少
- 最大 MIMO 层数更少
- 调制阶数更低
- 双工方式支持半双工
- 功耗更低

图5.267总结了 5G RedCap 设备与传统 Release 15 NR 设备的支持能力的对比。

	FR1频段		FR2频段	
	传统5G	5G RedCap	传统5G	5G RedCap
最大带宽	100MHz	20MHz	200MHz	100MHz
最小接收分支数	2或4(根据频段决定)	1（传统5G需要2路接收的情景） 1或2（传统5G需要4路接收的情景）	2	1
最大下行MIMO层数	2或4(根据频段决定)	1（RedCap需要1个分支数的情景） 2（RedCap需要2个分支数的情景）	2	1
最大下行调制阶数	256QAM	64QAM	64QAM	64QAM
双工方式	FD-FDD，TDD	HD-FDD，FD-FDD，TDD	TDD	TDD

图 5.267: 5G RedCap 设备与传统 Release 15 NR 设备的支持能力的对比

带宽的减少、上下行 MIMO 层数的减小以及最大下行链路调制阶数的减小都有助于降低基带复杂度。减少 TX/RX 上下行天线数量，并允许在 FDD 频段支持 HD-FDD 半双工工作模式，有助于降低天线和前端射频器件的成本。下面将详细地描述这些减少的功能特性。

支持带宽

传统 5G 设备需要在 FR1 频率范围内最大支持 100MHz 信号带宽，在 FR2 频段内最大支持 200MHz 信号带宽。而对于 RedCap，最大支持信号带宽分别减少到 20MHz 和 100MHz。带宽减小降低了系统复杂度，然而这种带宽的减少并不影响 5G RedCap 对于网络的接入，也最大限度地减少了对网络和设备部署的影响。

最小接收分支数

接收分支数可以理解成接收天线数量。通过减少接收分支的数量，可以减少接收天线的数量，节约成本。接收天线的最小个数要求与频段有关：在 FR1 频段，对于传统 NR 设备需要配置至少 2 个或 4 个接收天线，RedCap 设备只需要配置 1 个或 2 个接收天线。在 FR2 频段，对于传统 NR 设备需要配置至少 2 个接收天线，RedCap 设备只需要配置 1 个接收天线。

最大下行 MIMO 层数

RedCap 设备支持的下行 MIMO 层数是和天线数量相关的，最大能够支持的下行 MIMO 层数等于天线数量。接收天线数量的减小相应也降低了最大下行 MIMO 层数的能力，带来了基带信号处理复杂度的降低。

最大调制阶数

在 FR1 频段，传统 NR 设备需要支持下行 256QAM。对于 RedCap 设备，当前要求最大支持 64QAM 调制方式。

双工方式

5G RedCap 对于 TDD 频段要求还是一样的，主要在 FDD 频段要求有所放松。在 FDD 频段，传统 NR 设备需要支持 FDD 频段内的全双工（Full duplex FDD，FD-FDD）工作模式，前端需要采用双工器（Duplexer）来实现收发同时工作。对于 RedCap 设备，FD-FDD 的支持是可选的，可以采用半双工（Half duplex FDD，HD-FDD）的工作方式，这样射频前端就省去了双工滤波器的需要，可以采用开关进行收发的切换。

功耗

Rel-17 在 RRC 控制，空闲和非活动模式中增加了对功耗优化的增强技术，这一点在 5G RecCap 中同样也有体现，例如在低移动设备中放宽监控机制，采用增强型非连续接收（eDRX），减少寻呼（Paging）周期，达到降

	FR1 FDD	FR1 TDD	FR2
成本降低	~65%	~58%（RedCap设备支持2X接收天线）	~48%
		~71%（Redcap设备支持1X接收天线）	

图 5.268: RedCap 设备相比传统 5G 成本对比 (来自 3GPPTSG-RAN WG1 Meeting,R1-2008837)

	速率	时延	可靠性	电池寿命	设备尺寸
可穿戴	下行5-50Mbps 上行2-5Mbps	宽松	N/A	至少几天 或1-2周	紧凑型
工业无线传感器	<2Mbps	<100ms	0.9999	至少几年	N/A
视频监控	2-4Mbps（经济型） 7.5-25Mbps（高端型）	<500ms	99%-99.9%	N/A	N/A

图 5.269: 可穿戴、工业无线传感和视频监控要求（来自 3GPPTR38.875）

低终端功耗的目的。

根据以上 5G RedCap “减能力”的研究，相比传统 5G 设备而言，5G RedCap 预计可以使 5G 模组成本降低 60% 70%，如图5.268。设备复杂度的降低大幅度带来了成本的节约。

5.15.4 5G RedCap 应用场景在哪里

如前文所述，5G RedCap 在这些“减能力”的加持下，带来了设备复杂度和成本极大的优化。3GPP 文件 RP-202933 中描述了这些用例的关键需求 [4]，如图5.269所示主要包括三个应用场景：

- 可穿戴设备，如智能手表、可穿戴医疗设备、AR/VR 护目镜等
- 工业无线传感器
- 视频监控

如图5.2705G 技术示意图所示，5G 标准目前还没有可以与高性能、低时延的 eMBB、ULRRC 标准和低端大规模物联网 mMTC（NB-IoT、LTE-M）标准共存的中高速标准，RedCap 的出现正好契合了物联网这一需求。

5G RedCap 在带宽，吞吐率方面低于 eMBB，但是远高于 LPWA；在功耗和成本方面高于 LPWA，但又远低于 eMBB。5G RedCap 这一技术特性，可以广泛应用在可穿戴设备、工业传感器以及视频监控等领域。



图 5.270: RedCap 在 LPWA,eMMB 和 URLLC 中技术定位



图 5.271: 3GPP Release17 时间线

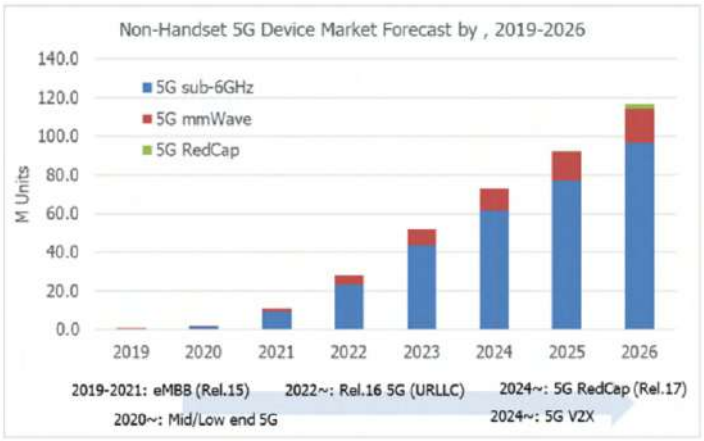


图 5.272: 非手机类 5G 设备市场预测

5.15.5 5G RedCap 商用进程

根据如图5.2713GPP R17 时间线，5G RedCap 重要的商用时间点如下：

- 2019 年 6 月 3GPP 会议上 5G RedCap 首次作为 Release17 Study Item 被提出
- 2021 年 3 月 5G RedCap 标准化项目正式立项
- 2022 年 Q2，预计 Release17 冻结，5G RedCap 标准化完成

Release17 标准化完成并不意味着 5G RedCap 马上商用，还需要芯片公司、物联网模块公司做大量的应用工作。根据业界推测，至少还要 1 ~2 年时间才能看到 RedCap 的商用产品出现。5G RedCap 芯片组预计将在 3GPP Release17 正式发布后的 2022 年底至 2023 年中首次亮相。高通和展锐、ASR 等公司有可能在早期阶段将其商业化。预计 2024 年后，LTE Cat.1/Cat.4 商业化的几家芯片供应商将陆续推出 5G RedCap 芯片解决方案。

根据 TSR 预测，如图5.272，5G RedCap 设备市场将在 2025 年开始增长，到 2026 年 5G RedCap 市场将初具规模 [1]。

5.15.6 关于 5G RedCap 的一些预测

3GPP Rel-17 协议即将于 2022 年 Q2 完成冻结，5G RedCap 也将为 5G 打开全新的市场，有望成为各家芯片厂商争夺的目标。当前或未来 5G RedCap 会如何发展，可以看到的趋势又如何？

个人有如下几个判断和预测：

5G RedCap 设备会是单模吗？

单模 5G RedCap 设备只可连接 5G 网络，无法连接到 LTE 网络。因此至少在前期阶段我们可能会看到 RedCap 设备会采用双模模式，同时支持 LTE 和 5G RedCap 技术。这样即使在没有 5G NR 覆盖的地区，也可以使用现有的 LTE 网络连接。

5G SA only 的进程决定着 5G RedCap 单模的演进，这一方面我国的 5G 网络部署走在了前面。

5G RedCap 对射频前端器件有何影响？

Rel-17 RedCap 有一个很吸引人的功能，就是在 FDD 频段引入了 HD-FDD 半双工工作方式，从而减少对双工器件的需求。如果 5G RedCap only 到来后，HD-FDD 可省 FDD 频段双工器，RedCap 设备的成本和尺寸会更有优势。但是正如前面所说，5G RedCap 设备至少初期还是会支持 LTE/RedCap 双模，双工器还是省不了。

双工器的移除将降低射频前端插损，这使得 5G PA 发射更少的功率就能够满足天线口功率要求。同时，5G RedCap 仅要求支持上行最大 20MHz/16QAM 调制方式，PA 的设计要求会有放松，成本也会有一定的降低。

5G RedCap 的 SoC 芯片中会集成射频前端吗？

对于 RedCap HD-FDD 工作方式，省去了双工器，这确实有点像 NB-IoT，后者现在已经做到了基带加射频、射频前端高度集成的单芯片方案。

对于 5G RedCap SoC，基带逻辑尺寸可能和 LTE Cat.4 差不多，射频部分要比 LTE Cat.4 要大些，从成本角度考虑，22/28nm 将会是较为合适的工艺节点，考虑到性能及芯片面积，12/14nm 也是个选择。

从目前看，为满足足够的线性功率输出，射频前端需要独立于 SoC 芯片。但随着 HD-FDD 半双工模式的引入、频段的精简，不排除在某些对集成度要求高、性能要求低的特定市场出现单芯片解决方案。

5.15.7 总结

“4G 改变生活，5G 改变世界”，5G 之所以能成为改变世界的通信模式，就是因为 5G 不仅仅考虑了适应人与人互联的移动通信场景，还考虑了用于物与物相连的万物互联场景。5G RedCap 是 3GPP 从物联网应用场景需求出发，针对于中速率、低成本场景开发的通信标准，有望在 5G 物联网中得到广泛应用。

“5G”与“物联网”是慧智微关注的智能互联应用场景，慧智微期待用创新的射频前端技术，为 5G 设备提供智慧连接，努力“化繁为简，使一切智慧互联”。

本文整理过程中得到众多同事及同行的帮助，在此表示感谢。

你对 5G RedCap 技术及未来发展有什么看法，欢迎扫描文章底部二维码，在线留言讨论。

参考文献

1. 2021 Cellular Broadband Device & ModuleMarket, Techno Systems Research, 2021.
2. https://www.3gpp.org
3. 2021 中国移动移动物联网研讨会，https://www.c114.com.cn/news/118/a1184735.html，2021 年 12 月 17 日
4. 3GPP RP-202933, New WID on support ofreduced capability NR devices, https://www.3gpp.org
5. https://www.ericsson.com/en/blog/2021/2/reduced-cap-nr
6. https://www.sierrawireless.com/iot-blog
7. https://omdia.tech.informa.com/OM018676/ 5G-RedCap is-the big missing piece of the 5G IoT puzzle
8. https://itectec.com/archive/3gpp-specification-tr 38-875
9. https://blog.csdn.net/qq_38987057/article/details/ 121026368
10. https://user.guancha.cn/main/content?id=572503
11. https://iot.ofweek.com/2021-12/ART-132216- 8500-30540154.html
12. https://iotbusinessnews.com/2022/01/24/74015 non-handset 5g-market volume will reach-117-million-units-by-2026/

在线阅读

本文摘自慧智微微信公众号，扫描以下二维码，在线阅读



5.16 车联网，需要什么样的射频前端芯片？

作者：秦沣牛娃

近年来，智能汽车市场整体爆发式增长，根据 Canalys 的统计和预测，如图5.274，2020 年全球电动汽车出货量为 310 万辆，到 2030 年将达到 3,300 万辆，接下来 10 年产生大于 10 倍的出货增长。

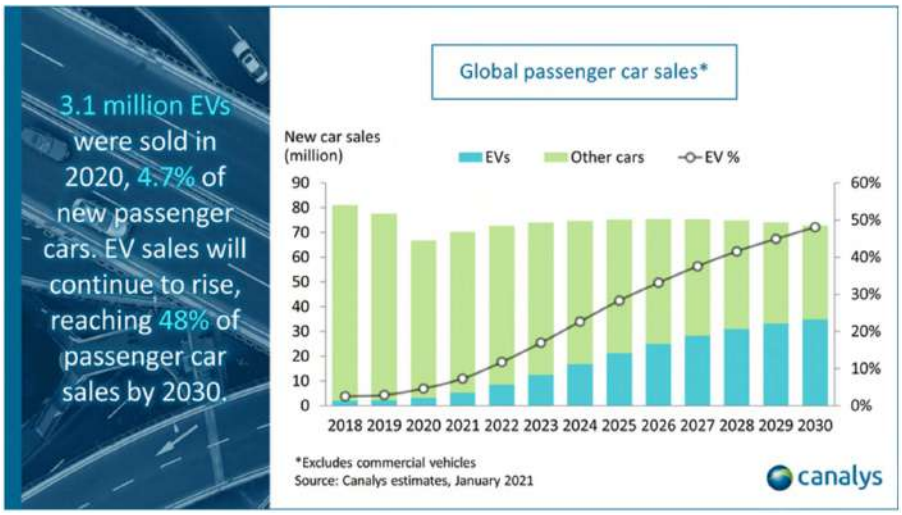


图 5.274: Canalys 对于电动汽车的统计和预测 [1]

无线通讯产业在最近几年同样经历快速发展，尤其作为新基建的 4G、5G 通信系统，短短几年在全球快速部署，帮助各行各业实现无线智能互联，提高工作效率和成绩。

由以上两大行业作为市场和技术基础，车联网行业应运而生。

车联网可以帮助汽车实现重要智能功能，根据资料统计 [2]，车联网可以分阶段实现以下功能：

1. **第一阶段：智能化、娱乐化**。包括：上网、听歌、视频等（现阶段）；
2. **第二阶段：辅助驾驶**。可以理解为应急刹车、辅助停车等；
3. **第三阶段：特定场景无人驾驶**。比如在货场使用无人驾驶，或者有固定路线的无人驾驶。考虑到安全因素，还未将无人驾驶技术在所有场景中全部实施；
4. **第四阶段：完全无人驾驶（最终目标）**。

在车联网中，需要用到的最重要技术之一就是车载无线通讯技术。车用无线连接有哪些技术？车用射频前端芯片有哪些特殊要求？带着以上问题，本文对车用射频前端芯片做一个讨论。

5.16.1 汽车无线通讯方案

用于汽车无线连接的通信技术主要有蜂窝网络技术、无线局域网技术、全球导航卫星系统、V2X（车联网）技术。以下将对这几种不同技术展开讨论。

5.16.1.1 蜂窝网络（Cellular）技术

蜂窝网络又称移动网络，由于构成网络覆盖的各通信基地台的信号覆盖呈六边形，从而使整个网络像一个蜂窝而得名。大家所熟悉的手机就是蜂窝网络的典型终端设备。

从网络系统和上层应用来看，汽车蜂窝网络系统和手机蜂窝网络系统是完全一样的，也包括了 2G、3G、4G、5G 网络。采用蜂窝网络连接的汽车一般是采用蜂窝物联网模块，将汽车与蜂窝基站相连。

出于成本和整体产业链的成熟度的考虑，大多数中低端汽车会直接应用普通消费级物联网模块进行实现（射频前端方案也套用消费级器件）。

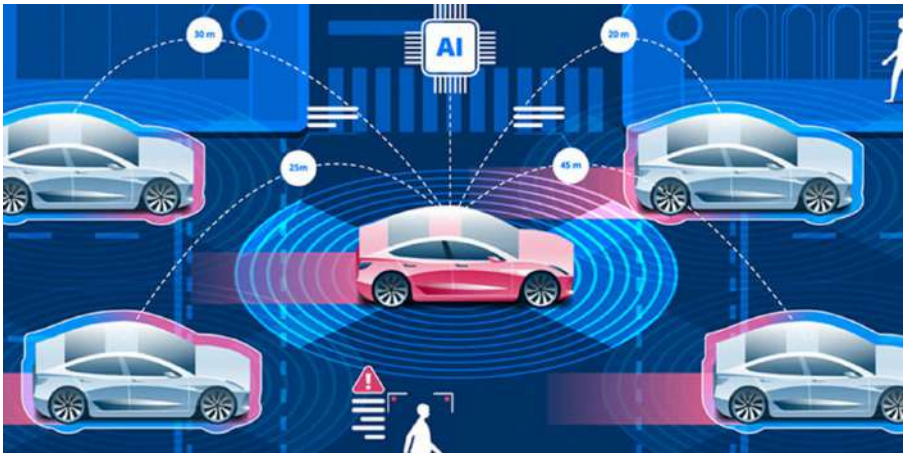


图 5.275: 汽车中的无线通讯

GNSS	GPS	GLONASS格洛纳斯	Galileo伽利略	Beidou北斗
市场	美国及部分国家	俄罗斯	欧盟	中国
主要频率	L1：1575.42MHz L2：1227.6MHz L5：1176.45MHz	L1:1602+0.5625KMHz L2:1246+0.4375KMHz (K=1-24，为每颗卫星的频率编号)	E5a/E5b: 1164-1215MHz E6: 1260-1300MHz E2: 1559-1592MHz	第一阶段： B1: 1561.098MHz B2: 1207.14MHz B3: 1268.52MHz

图 5.276: GNSS 系统及主要频段划分

5.16.1.2 无线局域网（WLAN）技术

车载 WLAN 主要是作为连接热点来应用，和无线路由器及 CPE 功能一致。一般也包括 2.4G 和 5G 两个频段，同样也包括 IEEE 802.11b/g/a/n/ac/ax 的应用，或许将来也会有 Wi-Fi 6E，Wi-Fi7 和 802.11ad 的应用。

5.16.1.3 全球导航卫星系统（GNSS）

GNSS 的全称是 Global Navigation Satellite System，全球导航卫星系统，是卫星导航系统的统称，包含 GPS、GLONASS、Galileo、Beidou 等卫星导航技术。卫星导航技术也是一种无线通讯技术。

主要的 GNSS 技术及频段划分如图5.276所示，高精度的导航定位系统是车载系统尤其对于无人驾驶汽车是关键技术。

5.16.1.4 V2X 车联网

V2X 是 Vehicle To Everything 的缩写，是指汽车与外部的所有信息连接。如果说，蜂窝通信技术、无线局域网技术以及全球导航卫星技术是汽车“借用”过来的无线通讯技术，那 V2X 就是专门为车联网打造的通讯技术。

V2X 是多种场景的统称，包含比如 V2I（Vehicle to infrastructure）、V2N（Vehicle to network）、V2V（Vehicle to vehicle）以及 V2P（Vehicle to pedestrian）等。V2X 概念最早于 1970 年被提出，目的是为了解决行车的安全问题，经过几十年发展，现在 V2X 有了更远大的目标：实现无人驾驶。进入 2020 年后，各大市场都在评估或者部署相关系统。从 5GAA（5G Automotive Association，5G 汽车联盟）给出的白皮书及各大市场的动作来看，2020 年代是 V2X 大发展时代，在此期间，整个汽车行业都会有很大的技术发展空间和市场潜力 [2]。

V2X 主要有两大技术标准：分别是 IEEE 所倡导的 DSRC 和 3GPP 所倡导的 C-V2X。目前应用中 C-V2X 有较大的技术和市场优势，后面主要针对 3GPP C-V2X 系统做相关讨论。



图 5.277: V2X 的两大分支

5.16.2 C-V2X 及射频前端方案

5.16.2.1 V2X 及其两大分支简介

从现阶段行业分布来看，V2X 主要有两大技术方向，分别为 DSRC 及 C-V2X，如图5.277，这两大方向主要是由技术传承及市场潜力来导向的。

DSRC 以欧洲市场为主，该系统是将 IEEE802.11 作为技术基础，可以理解为和 Wi-Fi 是一个系统；C-V2X 以中国和美国市场为主，该系统是将 3GPP 及其演进作为技术延申，可以理解为和 LTE、NR 是同一个系统，3GPP 组织也在主导促成统一标准。

V2X 的两种系统简介：

- DSRC

  DSRC 被定义为 IEEE 802.11p，将 IEEE 802.11 作为技术基础，和 Wi-Fi（IEEE 802.11 b/g/a/n/ac/ax）同出一脉，主要市场在欧洲及其它一些小国家，美国已于 2020 年底从 DSRC 切换到 C-V2X 发展方向。

  DSRC 最初由通用、福特、丰田、本田等车企推动，可支持车与车、车与路之间的直接通信，此后欧洲和日本等国家也都提出了各自的标准和专用频段。

  DSRC 技术的优势在于可靠性高、传输实时性强，弱点在于通信距离优势不明显且只能在 V2V 工作模式下有优势。因此，如果仅通过 DSRC 实现无人驾驶技术与车联网通信，那么就需要针对路边设施进行大规模经济投入，而这将限制 DSRC 的商业化。

- C-V2X

  C-V2X 是基于 3GPP 体系并进行延申，主要有中国、美国两大市场。

  C-V2X 最初由大唐、华为等设备商推出，3GPP 协议针对 V2X 的通信需求定义了两种通信方式：集中式（Uu）与分布式（PC5）。

  集中式也称为蜂窝式（Cellular-based），需要基站作为控制中心，车辆与基础设施（V2I）、车辆与其他车辆（V2V）、车与人（V2P）之间需要通过将数据在基站上进行中转来实现通信；分布式也称为直通式，无需基站作为支撑，让终端之间实现 V2X 通信。

  C-V2X 可以将直通模式和蜂窝模式融为一体，相较于 DSRC，具有显著的性能优势和市场前景，且现在已经演进到 5G（NR）阶段。预计 C-V2X 作为一个新兴行业，在今后十年有很大的市场潜力。

5.16.2.2 C-V2X 频谱资源及市场分布

由于 C-V2X 技术及产业近几年刚刚起步，很多国家和市场都处于中立态度去评估发展 DSRC 还是 C-V2X，且频谱划分也没有最终分配完成。

目前，3GPP 36 系列（LTE）定义的 C-V2X 频段只 B47，B47 主要是几个大市场（中国、日本、美国等）分配的频段，频率范围为 5.855-5.925GHz。B46D 虽然目前 3GPP 还未做定义，但是该频段或许未来也会覆盖全球大部分市场，B46D 频率范围为 5.725-5.825GHz；3GPP 38 系列（NR）定义的 C-V2X 频段为 n38 和 n47，n38 是用于特定业务，n47 的频率范围和主要目标市场同 B47 是一致的。

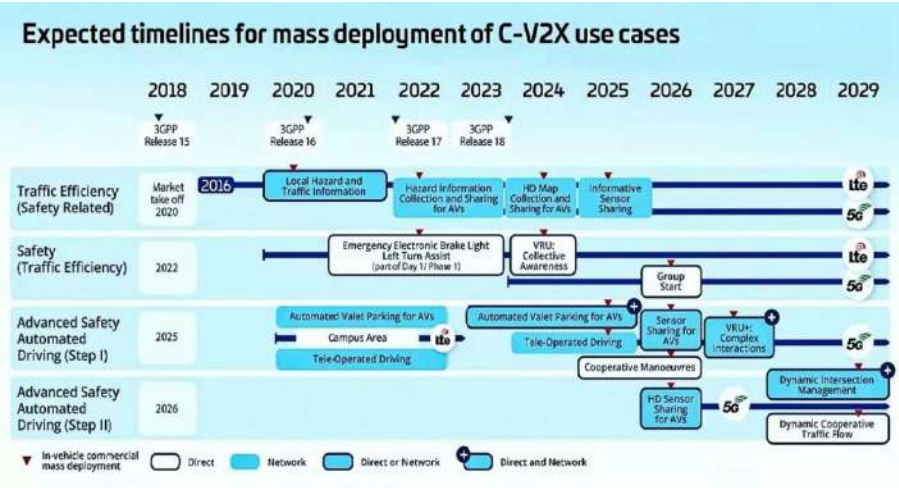


图 5.278: 5GAA 发展路标图

现阶段，中美两大市场都明确了发展方向为 C-V2X，具体频谱资源暂定为：

- USA：5.850-5.925GHz 共 30MHz 频谱资源，released in 2020 [3]
- China：5.905-5.925GHz, released in Nov, 2018 [3]
- Japan：5.850-5.925GHz [4]

注：Wi-Fi 6E 的频率范围为：5.925-7.125GHz.

5.16.2.3 C-V2X 期望发展路标图

结合如图5.278C-V2X 的 Roadmap[2]，可以将 C-V2X 的设计规划及商业应用分为以下四个主要阶段：

1. 阶段一：直到 2020 年，市场起步；
2. 阶段二：到 2022 年，辅助驾驶；
3. 阶段三：到 2025 年，无人驾驶阶段 1（特定场景）；
4. 阶段四：2025 年以后，无人驾驶阶段 2（完全无人驾驶）。

5.16.2.4 C-V2X 终端平台信息

C-V2X 终端主要采用以下平台：

- 4G C-V2X 平台
  1. QCOM MDM9150 4G C-V2X platform, released in Sept, 2017, CS in 2H,2018；
  2. Huawei Balong765 4.5G platform including C-V2X, released in Feb,2018；
  3. Autotalks V2X platform；
- 5G C-V2X 平台
  1. Huawei Balong5000 5G platform including C-V2X, released in Jan, 2019；
  2. QCOM SA515M 5G platform including C-V2X, released in 2020；
  3. MTK C-V2X platform, is not released yet.

5.16.2.5 V2X 典型射频前端方案

图5.279为一些现在已经商用的 V2X 射频前端方案：

1. SKYA21043（基于 IEEE 802.11p）：支持 DSRC 系统，是 Autotalks 平台的参考设计；
2. QPF1002Q（基于 3GPP）：PA 两级增益，支持 C-V2X 系统，是 QCOM 4G 平台 MDM9150 和 5G 平台 SA515M 的参考设计。

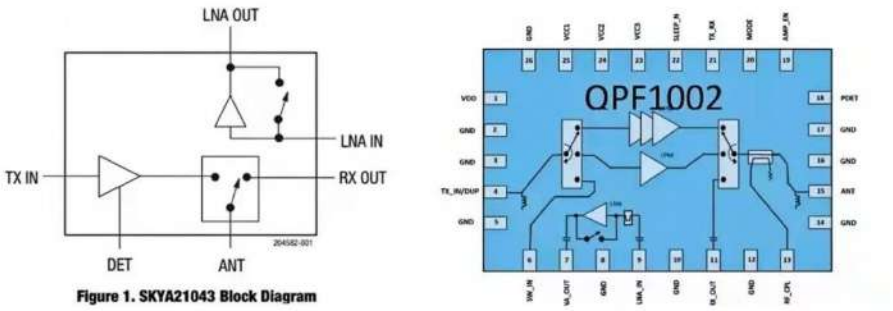


图 5.279: DSRC 和 C-V2X 射频前端方案



图 5.280: 汽车行业供应链层级

5.16.3 汽车电子及车规射频前端芯片

5.16.3.1 汽车行业供应链层级

汽车行业对上下游厂商的称谓和消费电子市场有一些差异，由于汽车产业链较长，汽车产业链一般称为 Tier1、Tier2 等。这些供应链层级如图5.280。

需要说明的是：由于现在产业链的横向和纵向的大量扩展，很多厂商都同时担任 Tier1 和 Tier2 的角色，如 Gosuncn、Neoway 等。和手机行业芯片供应商直接对接 OEM（vivo、OPPO、小米等）不同，在汽车行业射频芯片供应商一般很少接触 OEM，而是对接 Tier1 或者 Tier2 的通讯模组厂商。

5.16.3.2 车规芯片特点及要求

由于工作环境和汽车系统高安全性的应用场景，车规级芯片需至少满足以下要求：

1. 高可靠性；
2. 高安全性；
3. 零故障率；
4. 高量产一致性；
5. 长供货期（预期 5 ~7 年，甚至 10 年以上）；
6. 长使用寿命（约 15 年）。

产品阶段	参考标准	备注
设计	ISO26262 功能安全 GB/T 12750 半导体IC规范	可靠性设计技术
生产制造	TS16949 车企制造体系质量标准	生产过程质量管控
	AEC-Q 001 生产平均测试指南	测试
	AEC-Q 002 生产统计分析指南	测试良率
	AEC-Q 003 IC电气性能特性指标	
	AEC-Q 004零缺陷指南	
可靠性认证	AEC-Q100 基于失效机理的压力测试认证-针对 IC	可靠性
	AEC-Q101基于失效机理的压力测试认证-针对半导体分离器件	
	AEC-Q102基于失效机理的压力测试认证-针对光电器件	
	AEC-Q103基于失效机理的压力测试认证-针对传感器	
	AEC-Q104基于失效机理的压力测试认证-针对MCM器件（模组）	
	AEC-Q200基于失效机理的压力测试认证-针对无源器件（被动器件）	
	AQG324标准 针对功率模块，不属于AEC-Q体系，属于ECPE组织	如IGBT

图 5.281: 车规芯片设计、生产规范

5.16.3.3 车规芯片设计、生产、认证标准

由于车规芯片的特殊应用要求，车规芯片设计、生产、及可靠性认证都需要参考特定标准。车规芯片设计、生产的主要规范如图5.281。

5.16.3.4 有关 AEC-Q 认证

AEC-Q 系列认证是公认的车规元器件通用测试标准。克莱斯勒、福特和通用汽车为建立一套通用的零件资质及质量体系标准而设立了汽车电子协会（AEC）。AEC-Q 系列认证虽然不是强制性的认证制度，但目前已成为整个汽车电子行业公认的车规器件通用测试标准。芯片企业需以产品分类选择对应的 AEC-Q 标准：

- AEC-Q100:IC 芯片，包含无线通讯平台芯片组（AP/Modem/XCVR/PMIC …），射频前端器件（PA/LNA/Switch …）等有源器件；
- AEC-Q101：半导体分立器件，最常见就是二极管/MOS 管等；
- AEC-Q102：光电器件，如 LED。和射频前端产品关系不大；
- AEC-Q103：传感器，如测试胎压。和射频前端产品关系不大；
- AEC-Q104：MCM，和射频前端产品关系密切，如 PAMiD/PAMiF/L-FEM 等；
- AEC-Q200：无源（passive，被动）器件，例如 Filter/DPX/QPX 等。

具体的操作指南文档：

- AEC-Q001 Guidelines for Part Average Testing
- AEC-Q002 Guidelines for Statistical Yield Analysis
- AEC-Q003 Guidelines for Characterizing the Electrical Performance
- AEC-Q004 Zero Defects Guideline (DRAFT)
- AEC-Q005 Pb-Free Requirements
- SAE J1752/3 Integrated Circuits Radiated Emissions MeasurementProcedure

AEC-Q100 认证

AEC-Q100 文档定义了 AEC-Q100 认证涵盖的内容是 IC 芯片。对于无线通讯及射频前端行业来说，包含无线通讯平台芯片组（AP/Modem/XCVR/PMIC …），PA/LNA/Switch 等有源器件。

AEC-Q100 文档的最新版本：AEC-Q100 Ver. H 版本，release date：Sept 11, 2014。

AEC-Q100 文档 H 版本相较于 G 版本，重新定义了温度范围，包括以下四组。所有测试都是基于器件厂商提供的温度范围（From Datasheet）来进行测试的。一般车规通讯模组的要求是 Grade3，要求高的需要支持到 Grade2。根据沟通及了解，这几个温度范围主要适用于以下场景，如图5.282。

	Part Operating Temperature Grades	
	Grades	Ambient Operating Temperature Range
引擎	0	-40℃ to + 150℃
引擎盖下或者外面	1	-40℃ to + 125℃
车体外	2	-40℃ to + 105℃
车舱内	3	-40℃ to + 85℃

图 5.282: AEC-Q100 文档规定的温度范围

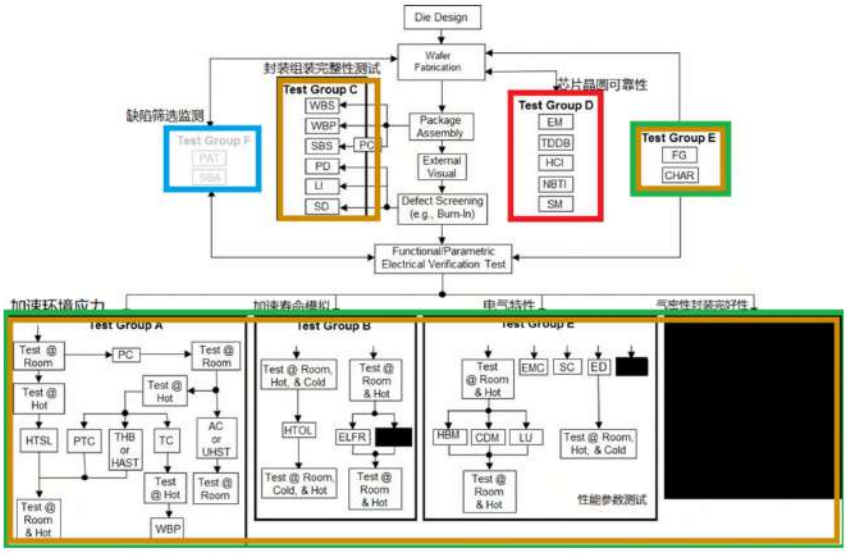


图 5.283: AEC-Q100 认证流程

AEC-Q100 认证流程

AEC-Q100 认证包括 A、B、C、D、E、F、G 七大类，40 多项测试，具体参考下图。除了 G 组关系不大之外，其余六类都和射频前端器件强相关。

由于篇幅有限，截取具有代表性的图5.283做以说明，其中基本射频指标是射频前端器件的基础，划分到 E 组电气特性中。G 组是特殊封装完整性测试，和射频前端器件关系不大。详细信息可以直接阅览 AEC-Q100 原文。

5.16.3.5 车规电子和消费电子的差异

车规电子和消费电子的差异如下：

车规器件特点：

1. 最高优先级的要求：可靠性，稳定性，使用寿命，工作温度。这几项为重要要求，同时也是 AEC-Q 的主要约束；
2. 次高优先级的要求：静电防护，冷热冲击，振动冲击。

消费器件特点：

1. 最高优先级的要求：价格，供应，性能；
2. 次高优先级的要求：功耗，兼容性，使用寿命，迭代能力，可靠性、稳定性等。

一般在消费类电子应用中，尤其中低端产品，一切以成本优先，这与车规芯片有大的不同。在有些车载设备中，比如车载 OBD 等后装组件，应用场景和售后处理都类似于消费电子，厂家在售后处理中承担的成本和风险

都是可控的，所以很多时候会采用消费级芯片进行设计。

5.16.3.6 Tier1 和 Tier2 所关心的问题

对于前装车载无线通讯产品，通常车辆拆装人工成本比较高，所以产业链都在避免因为可靠性问题等导致客退及维修。对这类问题的关注及敏感导致汽车行业 Tier1/2 在器件选择时非常谨慎，尤其对于射频前端器件。在应用中，一般每一颗器件发生失效情况都要给出理由充分的根本原因及解决方案。

在车规射频前端芯片应用中，Tier1 和 Tier2 重点关注的指标如下：

1. **功能及规格**：需要纵向（应用场景：2G/3G/4G/5G，C-V2X，DSRC，WLAN 等）与横向（实现方案：MMPA，TXM，PAMiD，L-FEM，PAMiF 等）结合去考虑基本功能和规格，其中包括射频性能；尺寸大小；工作温度范围等；
2. **高可靠性**：工作温度 Grade2：-40C ~+105℃，Grade3：-40C ~+85℃，PPM/MTBF/FIT 等要求，以及 EMC/ESD/环境耐受要求；
3. **质量认证**：AEC-Qxxx（注重结果），IATF 16949 （注重整个过程，来源于 ISO9001）；
4. **使用寿命/质保周期**：Tier1、Tier2 和 OEM 车厂都希望保证这个时间（~15 年）；
5. **供货周期**：5 ~7 年，有些甚至达到 10 ~15 年；
6. **服务支持**：提供技术支持及客退分析支持；
7. **车规射频前端器件的供应链保证**：Wafer，基板及辅助材料，封测厂等。

5.16.3.7 射频前端芯片供应商所做的保障

为了满足车规使用，射频前端芯片需要做到性能、可靠性等方面的保障，主要有：

- **性能保障**：

  完全按照 3GPP 要求射频性能及可靠性，温度范围为 Grade2：-40C to +105C ，Grade3：-40C to +85C。车规射频前端器件的应用和消费类射频前端器件的应用基本是一致的，这方面射频前端芯片可以做到良好保障。对于中高端车型来说，数据吞吐率和普通数据类设备（MBB 产品）要求是一致的，需要较高上下行吞吐率，速率要求要高于中高端手机。

- **可靠性保障**：

  0PPM 是最终目标。任何客退的发生都会增加很多成本，尤其前装车载无线产品。如果发生客退情况，哪怕小概率（即使是 1PPM），车厂都会追责到底且需要给出合理理由的分析过程及合理的整改措施。在这个过程中包括 OEM、Tier1、Tier2，芯片供应商都要投入巨大的人力资源和物力资源。由于汽车系统复杂，分析失效根因难度大。所以 0PPM 是最终目标。

- **车规级品质保障**：

  理论上消费类射频器件不能直接使用到车规终端产品上，车载产品需要满足 AEC-Q100 等车规级认证。

不过，除海外部分高端车厂外，现在车厂对用于车载通讯模块中的芯片并没有 AEC-Q100 强制要求，但通讯模块会做系列可靠性实验。这在一定程度上降低了对芯片的要求。

5.16.4 总结

“一切设备都将走向移动和互联”，汽车作为最为常见的“移动”设备，未来一定会走向智能互联。车联网，未来可期。

射频前端芯片技术将为汽车提供无线通讯技术，保障汽车的智慧连接。不过由于汽车需求的特殊性，射频前端芯片必须要经过特殊的设计、严苛的验证，才能达到车规级要求。

慧智微将不断开发高可靠、高性能射频前端芯片，为未来万物互联提供可靠保障。努力“化繁为简，使一切智慧互联”。

在此，真心感谢团队给予的大力支持并提出大量的建设性建议和意见。

您对车联网有什么看法，欢迎扫码在线留言讨论。

参考文献

1. https://canalys.com/newsroom/canalys-global electric vehicle-sales-2020
2. AVisionary Roadmap for Advanced Driving Use Cases Connectivity Technologies andRadio Spectrum Needs, 5GAA, Sept 9th , 2020
3. 蜂窝车联网（C-V2X）技术与产业发展态势前沿报告（2020），中国通信学会白皮书（最终稿）, Nov，2020
4. https://5gaa.org/Feb 14th , 2022
5. AEC-Q 系列文档
6. www.3gpp.org

附：车规无线通讯行业常用名词解释和释义

英文简称	英文全称	中文释义
V2X	Vehicle to Everything	车联网
C-V2X	Cellular-Vehicle to Everything	基于蜂窝网络车联网（基于3GPP）
DSRC	Direct Short-Range Communication	直接短距离通讯（基于802.11p）
LTE-V2X	Long Term Evolution V2X	基于LTE的车联网
NR-V2X	New Radio V2X	基于新无线的车联网（5G NR）
V2I	Vehicle to Infrastructure	车到基础设施
V2N	Vehicle to Network	车到网络
V2I	Vehicle to Pedestrian	车到人
V2V	Vehicle to Vehicle	车到车
RSU	Roadside Unit	路侧单元（或者路侧设备）
OBU	On Board Unit	车载终端
5GAA	5G Automotive Association	5G汽车协会
ITS	Intelligent Transportation System	智能交通系统
C-ITS	China ITS Industry Alliance	中国智能交通产业联盟
CAICV	China Industry Innovation Alliance for the Intelligent and Connected Vehicles	中国智能网联汽车产业创新联盟
SAE	Society of Automotive Engineer	汽车工程学会
TIAA	Telematics Industry Application Alliance	车载信息服务产业应用联盟
PC5	Direct Communication C-V2X	直连通讯（3GPP上描述）
Uu	Network-based C-V2X	基于网络通讯（3GPP上描述）

图 5.284: 车规无线通讯行业常用名词解释和释义

在线阅读

本文摘自慧智微微信公众号，扫描以下二维码，在线阅读

5.17 车，怎么联网？

作者：彭洋洋 博士

智能汽车是当今科技领域的热门话题。相比于传统汽车，智能汽车不仅能提供更加舒适、安全、高效的驾乘体验，还能实现与其他车辆、基础设施、云端等的互联互通，从而打造一个智慧出行的生态系统。

汽车行业也因此正在经历一场前所未有的变革，这场变革被称为汽车的新四化，分别指的是：电气化、网联化、智能化、共享化。其中，**网联化**是新四化的基础和核心。网联化是指汽车通过各种通信技术与车内设备、车际设备、云端设备进行车联网，实现数据交换、信息共享和服务协同。



图 5.286: 汽车的新四化

那么，车联网是如何实现的呢？车联网与无线通信所使用的射频前端芯片又是什么关系？本文就尝试对以上技术做一个梳理。

5.17.1 车联网中的通信技术

车联网通信技术是指在交通环境中，实现车辆内部、路侧单元、行人、云端服务之间的信息交互和协同的技术。根据连接范围的不同，在此将车联网通信技术分为三类讨论，分别是：车内互联、车际互联和车云互联。这三种互联的应用范围和特点分别是：

- 车云互联：用于连接车辆与云端
- 车际互联：用于连接车辆与周边车辆、路侧单元、行人
- 车内互联：用于实现车内多种多样的灵活互联

以下就对这三种互联技术展开介绍。

5.17.1.1 车云互联

车云互联是指将汽车和云端服务器连接起来，这样就可以实现对车辆状态、位置、行驶数据等信息的采集和传输，并提供基于云端计算和大数据分析的各种信息服务。



图 5.287: 车联网中的部分通信技术

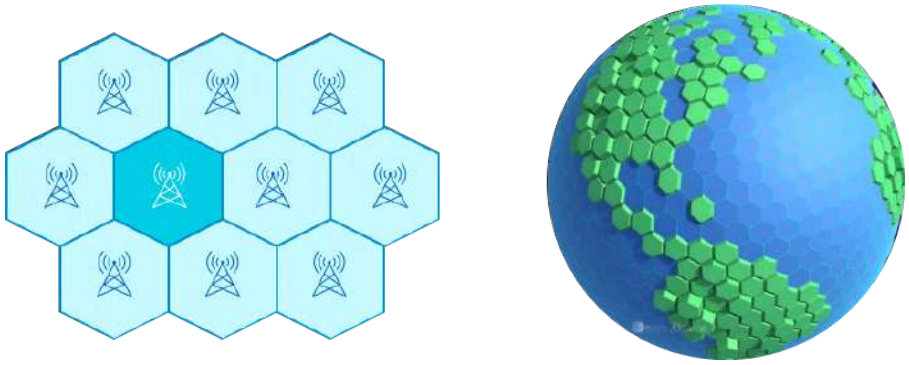


图 5.288: 蜂窝移动通信示意图

与云端互联之后，车辆就可以实现很多丰富的功能。比如：

- **车辆远程控制**：通过手机可以远程控制车辆启动、开门、空调、音响等功能，可查看车辆实时状态、故障诊断等功能
- **互联导航**：通过云端的地图数据和路况信息，可以为驾驶者提供最优的导航路线，也可以根据驾驶者的偏好和出行场景，推荐附近的停车场、加油站、餐厅等服务点
- **内容下载**：通过连接云端服务器，可以将音乐、视频、新闻、游戏等内容，下载到本地
- **车辆智能升级**：通过蜂窝网络，可以实现对车辆的软件和固件的远程升级，提高车辆的性能和安全性，也可以根据车主的需求，定制个性化的功能和设置
- **智能交通/智能城市管理**：通过将车辆连接至云端，就可以实现城市车辆的统一管理，实现更好的资源配置以及交通智能优化

与云端互联有诸多好处，也成为车联网必须实现的功能之一。目前车联网主要是**通过与蜂窝网络的连接**，实现与云端服务器的互联，实现数据传输。

蜂窝网络是一种利用蜂窝状的小区覆盖服务区域，实现移动通信的网络技术。由于网络覆盖的形状像蜂窝的一个个小格子，所以由此得名。英文称之为 Cellular Network（细胞网络），也是因为整个巨大的网络像由一个个小的“细胞”构成。

我们熟悉的 5G 手机就是典型的蜂窝技术代表，而经常提到的 2G/3G/4G/5G 的通信制式演进，指的也是全球的蜂窝通信技术。蜂窝网络的主要特点是：

- 蜂窝结构：通过将服务区域划分为多个相邻的正六边形小区，可以简化网络规划和管理，降低干扰和成本
- 移动性管理：通过在小区之间进行切换，可以保持移动用户与网络之间的通信连接，实现无缝漫游
- 频率复用：通过将同一组频率在不同的小区中重复使用，可以提高频谱利用率和系统容量

蜂窝通信所用到的基站一般只覆盖几百米或数公里，通过全球数千万个这样的基站，就可以实现整个地球上主要地区的全球通信。基站不动，手机只要连接到任何一个基站，都相当于连接到了整个世界。如图5.288所示，蜂窝通信是全球人类共同的大工程。

正是因为以上特点，自从 1983 年诞生以来，蜂窝通信快速迭代发展，经过 40 多年的发展，目前已经演进至第五代，也就是 5G 蜂窝通信网络。这张网络连接了全球人口，根据全球电信联盟的统计和预测，2022 年全球智能手机在网数已达到 64 亿，而到 2030 年，这一数字将达到 90 亿 [1]。

在 4G 之前，蜂窝网络主要给“人”用，其重要的通信载体是手机。但这一现象在 5G 到来之后迎来改变，5G 的网络定义将“物”的使用也考虑进来，力争实现万物互联的网络。蜂窝网络的三次重要变革如图5.289所示。

对于车云互联，最简单的方法就是让汽车可以接入蜂窝网络。根据 ABI 的统计，2020 年全球出售的新车中，41% 的汽车具备联网功能 [2]。这些车辆的联网功能基本均是由蜂窝网络接入车云互联来实现的。



图 5.289: 蜂窝网络的三次重要变革，力争实现万物互联



图 5.290: C-V2X 技术实现的车际互联

5.17.1.2 车际互联

尽管蜂窝网络覆盖范围广、产业链成熟，是汽车实现最初联网功能的首选通信技术。但蜂窝网络用作车际互联使用的时候也有一些弊端。比如：

- 连接蜂窝网络必须要基于基站建设，这就限制在车辆只能在有基站的地方使用
- 所有汽车都连接蜂窝网络（接入网-核心网）进行通信，导致更大的时延
- 车辆直接互联的创新场景，如车队编队、路侧感知等功能无法有效实现

于是，业界就开始开发专门给汽车使用通信技术，这其中的代表技术就是 3GPP 组织所推动的 C-V2X 技术。如图5.290，C-V2X 是一种为车辆设计的专门网络，C-V2X 的特点是：

- C-V2X 可以让车辆与其他车辆、路侧基础设施和行人直接通信，从而提高道路安全和交通效率
- C-V2X 不依赖于网络覆盖，可以在没有基站的情况下实现低时延、高可靠的直接通信
- C-V2X 可以利用 5G 技术提供更高的速度、更低的时延和更大的容量，从而支持更多的创新场景，如自动驾驶、车队编队、扩展传感器、远程驾驶等
- C-V2X 可以与现有的蜂窝网络和生态系统兼容，降低部署成本和复杂度

和传统蜂窝技术相比，C-V2X 实现自组网的原因是引入了短距离接口，一般称之为 PC5 接口，或者 Side-link 接口。称之为 Side-link 的原因是因为这个连接只负责和旁边（Side）物体连接（Link），又称为 PC5 的原因是因为短距离互联所需要的功率等级比较低，天线口为 20dBm，是 3GPP 所定义的第 5 功率等级（Power Class 5）。如图5.291所示。

Uu 接口与车云互联中提到的蜂窝网络互联无异，因为连接的是用户终端（User Equipment），这个连接在 LTE 时代的一些标准定义被称为 Uu，这个接口名称也在 C-V2X 中被沿用下来。这个接口的技术规格与 3GPP 所

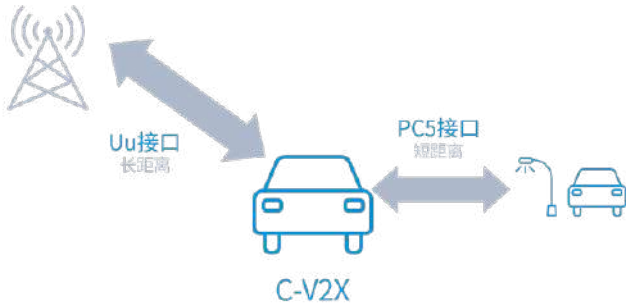


图 5.291: C-V2X 技术中的两种通信接口

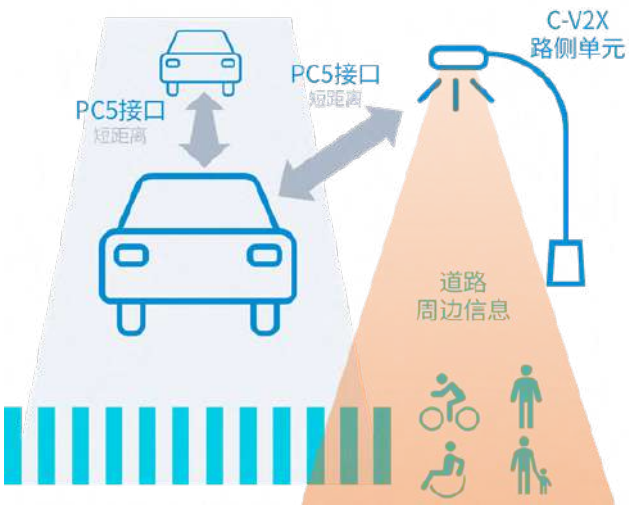


图 5.292: C-V2X 实现的对道路周边信息感知

定义的手机规格一致，并没有新的特性。

PC5 是为汽车等需要进行自组网互联的物体所定义的通信规格，这是 3GPP 首次开始定义自组网网络，通信规格对于 3GPP 来说是全新且充满挑战的。国内一些厂商在规格制定上也贡献出自己的力量。在 3GPP PC5 低功率终端协议的定义中，国内领先的射频前端厂商慧智微在 5G NR Rel-18 协议中参与 Side-link/PC5 相关的讨论，进行提案并立项，推动 PC5 标准在 3GPP 的完善（详见《5G 协议再演进：慧智微 3GPP 提案正式立项》）。

如 C-V2X 名称所示，C-V2X 期待实现基于蜂窝网络的汽车与任何物体的互联（Cellular- Vehicle-to-everything），这里大家可能会有疑问，如果车与人要互连，来感受车旁边的人的话，那是不是每个人都需要带一个支持 C-V2X 的手机？如果没有这部手机，是不是车辆就感受不到了呢？

其实在实现上，并不需要道路周边所有设备都加装 C-V2X 功能。汽车与道路周边万物的互联可以通过路侧的 C-V2X 路侧单元（RSU，Road Side Unit）实现。而路侧单元中可以加装人体识别、热感应、光学雷达等设备，用于感知道路周边的环境情况，再通过 C-V2X 中的 PC5 接口告诉车辆，这样就可以实现车辆对周边环境的感知 [4]。RSU 与 RSU 之间，RSU 与蜂窝基站之间，可以通过光纤或蜂窝网络连接。如图5.292所示。

以上汽车与路侧单元共同实现车辆对环境感知的方式叫“车路协同”。车路协同是 C-V2X 一大优势，通过车路协同，汽车不再是一个单独的个体，通过与路侧的协同、感知、控制，可以完成更智能、更安全、更可控的网联汽车。

如《车联网，需要什么样的射频芯片?》所述，目前 C-V2X 的设计规划及商业应用已经清晰，主要分为以下四个主要阶段：

- 阶段一：2020 年之前，市场起步
- 阶段二：2022 至 2022 年，辅助驾驶
- 阶段三：2022 至 2025 年，无人驾驶阶段 1（特定场景）
- 阶段四：2025 年以后，无人驾驶阶段 2（完全无人驾驶）



图 5.293: 部分车内设备使用到的通信技术

现阶段，中美两大市场都明确了发展方向为 C-V2X，专门用于短距离通信 PC5 接口的具体频谱资源暂定为[7][8]：

- 美国：5.850-5.925GHz
- 中国：5.905-5.925GHz
- 日本：5.850-5.925GHz

目前 C-V2X 产业正在快速发展中，根据佐思汽研发布的报告，2022 年已有 20 余款乘用车搭载 C-V2X，包括蔚来 ET7、别克 GL8 Avenir、广汽 Aion V、北汽 ARCFOX 等车型。2022 年 1-6 月，搭载 C-V2X 技术量产的乘用车约 4.6 万辆，这一数字在 2026 年将超过 200 万辆。未来五年实现数十倍的增长。

5.17.1.3 车内互联

车内互联技术主要实现汽车与车内物体的互联，比如：手机、车钥匙、临时接入的国内平板电脑、笔记本电脑等。所使用到的通信技术一般为近距离的无线通信技术。

目前车内所使用到的近距离无线通信技术主要有：

- Wi-Fi (Wireless Fidelity，**无线保真**) **技术**：Wi-Fi 技术是一种基于 IEEE 802.11 系列协议标准实现的无线通信技术，该通信协议于 1996 年由澳洲的研究机构 CSIRO 提出，Wi-Fi 凭借其独特的技术优势，是目前最为主流的无线局域网接入技术标准（蜂窝通信为广域网通信标准）。Wi-Fi 技术主要使用 2.4GHz 或 5GHz 的无线电频段，支持多种网络标准和安全协议，具有传输速度快、覆盖范围广、兼容性强等特点，所以被广泛应用于家庭、办公、公共场所等各种场景；
- **蓝牙** (Bluetooth)：蓝牙技术也是一种短距离无线通信技术，它可以在 2.4GHz 的 ISM 频段上实现设备之间的数据和语音传输。蓝牙技术由爱立信公司于 1998 年创立。相比于 Wi-Fi 技术，蓝牙技术的特点是低功耗、低成本、低复杂度、高安全性和高可靠性，同时蓝牙的数据传输速度也低于 Wi-Fi。蓝牙技术的应用领域非常广泛，包括汽车、工业、医疗、消费电子、智能家居等；
- NFC (Near Field Communication，**近场通信**)：NFC 技术是一种近距离无线通信技术，它可以实现电子设备之间的非接触式点对点数据传输。NFC 技术由 RFID（非接触式射频识别）演变而来，NFC 采用电磁耦合感应技术，电磁场频率是 13.56MHz。NFC 技术有很多应用场景，例如移动支付、公交卡、门禁卡等。NFC 技术的优点是方便快捷、安全可靠、低功耗、低成本等，其缺点是传输距离短、传输速率低、受干扰影响大等；
- UWB（Ultra-Wide Band，**超宽带**）：UWB 技术的基本原理是通过发送和接收极短的脉冲信号来进行无线通信。由于采用正弦载波，而是利用纳秒级的非正弦波窄脉冲传输数据，因此其所占的频谱范围很大，其频谱范围一般是在 3.1GHz ~10.6GHz 频率范围内的 500MHz 信号。UWB 技术目前的应用主要是在高精度定位中。由于 UWB 的高定位精度特性，未来可能应用于对安全有高需求的定向通信场景中。

在车内设备通信中，一般采用多种通信技术结合的方式。图5.293为部分车内设备使用到的通信技术。

在车内互联上，一般采用多种通信技术灵活互联的方式进行，可以根据场景的传输速率、成本、距离等，综合选取最为合适的传输技术。另外，也可以使用多种技术相结合。

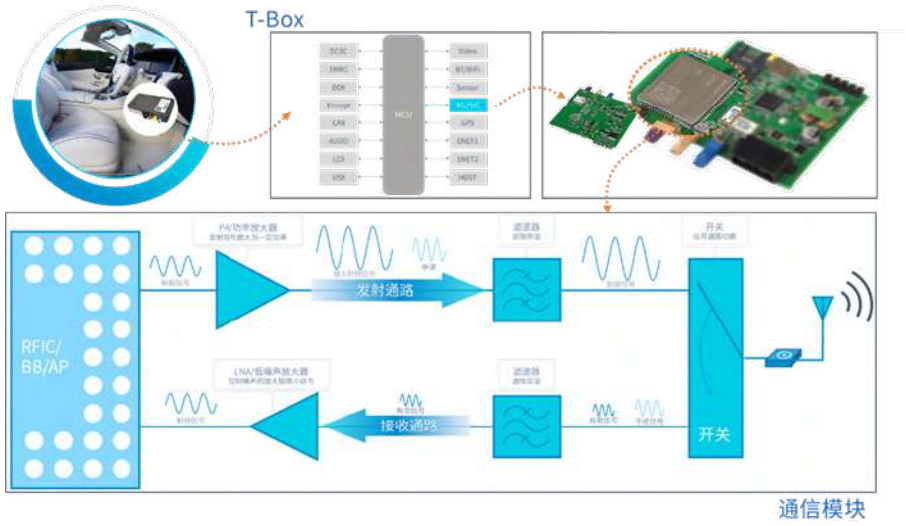


图 5.294: T-Box 及通信模块

5.17.2 车载通信的技术实现

汽车通信功能是依靠 T-Box 或智能网联模组来实现的。

T-Box 全称是 Telematics Box（远程通信盒子），T-Box 与智能网联模组中负责通信的主要功能模块是“通信模块”，而通信模块中实现无线互联的是射频前端芯片。

以下就针对 T-Box/智能网联模组、通信模块、射频前端进行介绍。

5.17.2.1 T-Box/智能网联模组的功能

T-Box 中的“T”是 Telematics 的缩写，Telematics 是 Tele-communications（电信）与 Informatics（信息科学）的合成词，是包含通信网络及信息处理的功能模块，T-box 一方面接入汽车总线，掌握车辆信息、整车控制信息，一方面又负责接入通信网络，实现网络通信功能。T-Box 的重要性在于 T-Box 汽车与外部设备之间的通信界面和功能实现，汽车的车云、车际、车内通信功能需要由 T-Box 实现。

T-Box 一般放置于中控台下方，由 MCU、4G/5G 通信模块、Wi-Fi 通信模块等多个电路模块构成，如图5.294所示。

智能网联模组一般设计成 SoC+ 通信模块集成的模组形式，直接与智能座舱系统集成，其功能与 T-Box 功能类似。

5.17.2.2 通信模块的功能

通信模块是 T-Box/智能网联模组中负责无线通信的单元。

随着协议的复杂，通信制式的增加，用于通信的射频系统也变的非常复杂。如《除了调匹配，射频人还要掌握系统知识》中提到，目前移动终端支持的频段数目达 30 个，并且每个频段都需要进行射频调试和适配，这就使得射频工作量极大。并且由于射频知识也较为晦涩难懂，射频专业人员招聘和培养相对困难，除非是年销量极大的头部手机公司，其他中小规模有无线互联需求的公司，很难建立起完整的射频开发和调试团队。

虽然射频技术复杂，工作量巨大，但其完成的功能却极其简单：射频就是把终端想要发射的信息，通过一定的形式，发射出去；再把需要接收的信息，想办法接收并把信息提取出来即可。

正是因为这种特点，物联网行业出现出一批专门提供专业通信解决方案的公司。这些专业的通信、互联解决方案公司将通信功能进行设计和包装，以实现完成通信射频复杂功能的同时，接口保持简化。使用户不再去关注射频电路的实现，而是把精力放到上层的方案设计中来。物联网模块公司的工作和提供的产品如图5.295所示。

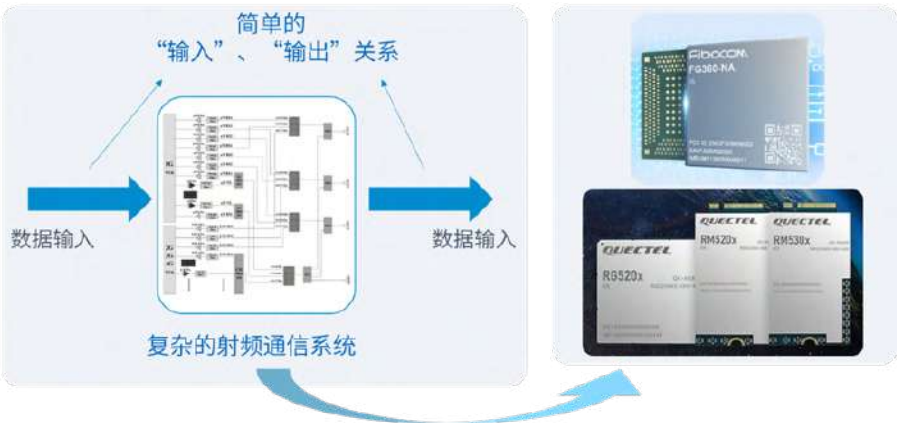


图 5.295: 物联网模块公司：将物联网无线通信功能实现方式专业化的公司

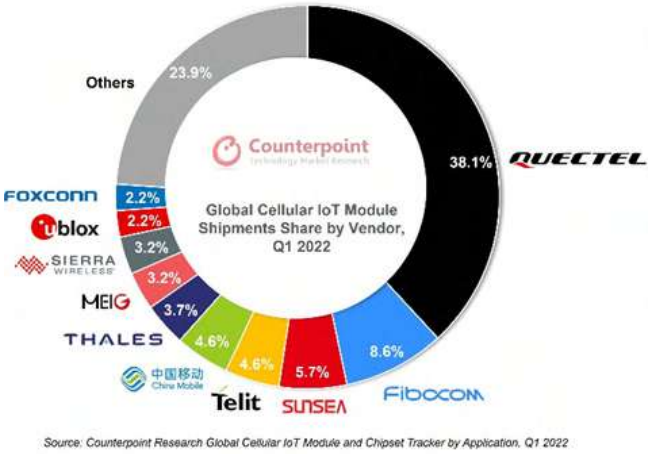


图 5.296

目前物联网模块行业已经发展较为成熟，如图5.296，根据 Counterpoint 的统计，2022 年全球物联网模块出货将超 6 亿片。其中根据 2022 年 Q1 的统计，移远、广和通、日海等厂商是此市场主要供应商 [10]。

在目前车载 T-Box 系统中，大部分方案也使用通信模块来实现通信功能。图5.297为广和通旗下全球车载通信模块厂商之一的 Rolling Wireless 的车载通信模块产品。该产品集成了蜂窝通信、C-V2X、GNSS 等功能 [11]。

5.17.2.3 射频前端的功能

通信模块中的射频连接功能是靠射频前端芯片来实现。

射频前端是无线通信模块的核心组件，它负责信号的发射和接收，由于位于整个系统的最前端，所以称为“射频前端”。射频前端的构成部分，及主要模块如图5.298所示。

射频前端包括功率放大器、低噪声放大器、滤波器、开关，它们共同构成了信号的放大、切换、滤波等功能。射频前端的性能直接影响到无线通信的速率、覆盖范围、功耗等指标。这些模块的主要功能是：

1. **功率放大器**（PA，Power Amplifier）：是指在给定失真率条件下，能产生最大功率输出的放大器。功率放大器的输出功率能力，决定了系统能输出的最大信号强度；
2. **低噪声放大器**（LNA，Low Noise Amplifier）：是指噪声系数很低的放大器，用于放大可能非常弱的信号。低噪声放大器的能力，决定了系统接接收到的最小信号强度；
3. **滤波器**（Filter）：是指能够根据信号频率的不同，选择性地通过或衰减信号的电路。主要用做滤除对其他频段的干扰，或者从众多干扰信号中抽取有用信号；
4. **开关**（Switch）：是指能够控制通路通断的器件。主要功能是铺路架桥，为射频信号提供必要的通路。



图 5.297: Rolling Wireless 提供的车载通信模块产品方案

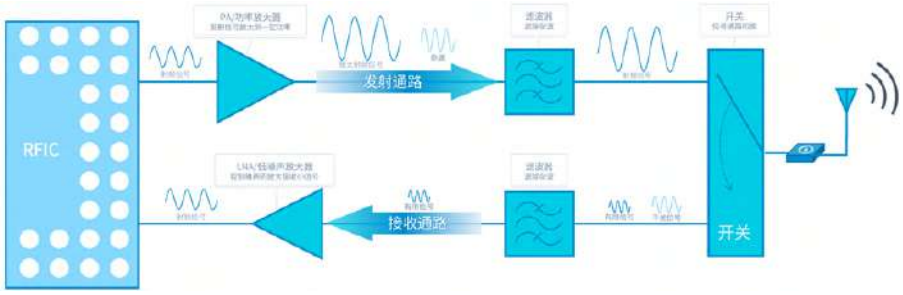


图 5.298: 射频前端的构成部分，及主要模块

5.17.3 车联网的芯片需求

随着汽车的智能化、网联化，根据英飞凌的统计，智能电动车中使用的芯片价值量在单车 834 美元。前述提到的各种复杂通信功能，最终都需要芯片承载实现。汽车成为芯片使用的巨大市场。

汽车对芯片的需求也和消费类有大的不同。为了保证汽车的安全、可靠和高效运行，同时也要适应恶劣的环境条件和多样的应用场景，针对车载应用的芯片，专门进行了车规级芯片需求。

车规级芯片和商用芯片在规格上有很大的不同，主要体现在以下几个方面：

1. **工作温度范围**：车规级芯片需要能够在极端的高温和低温下正常工作，一般要求在-40℃ 到 85℃ 之间，甚至到 125℃ 之间工作。而商用芯片一般只能在 0℃ 到 70℃ 之间工作；
2. **工作稳定性**：车规级芯片需要能够抵抗汽车运行过程中产生的各种干扰，如电磁干扰、电压波动、振动冲击等，而商用芯片一般只能在相对稳定的环境下工作；
3. **不良率**：车规级芯片需要有很低的不良率，目标达到 0 DPPM（Defect Part Per Million，每百万失效数），而商用芯片一般要求为 100 DPPM 左右；
4. **功能安全**：车规级芯片需要遵循 ISO 26262 等功能安全标准，保证在出现故障时能够及时检测、隔离和恢复，避免造成严重后果，而商用芯片一般没有这样的要求；
5. **认证流程**：车规级芯片需要通过 ITAF16949 等质量管理体系认证，以及 AEC-Q 系列产品质量认证，证明其符合汽车行业的要求，而商用芯片一般只需要通过普通的质量检测。

正是因为以上原因，车规芯片的价格也会高于商规芯片。成本提升的主要原因来自于应用于车载时，所做的必要的质量控制与标准认证。

5.17.4 车联网芯片的未来

车联网后，智能汽车可以利用先进的信息通信技术，实现车与车、车与路、车与人、车与云等多方面的网络连接，提高交通安全、效率和智能化水平。车联网将是未来智能汽车发展的大势所趋。

在车联网发展中，车联网芯片也会随之演进、优化，在未来车联网芯片的发展中，我们看到有以下趋势：

- **车联网芯片的市场需求将持续增长，**随着 5G、C-V2X 等技术的推广和应用，车联网芯片需要支持多模双通、高速率、低时延、高可靠性等特性，同时还需要满足车规级的可靠性和安全性要求；
- **芯片厂与整车厂将有更多协同，**随着车联网技术的发展，以及对车联网技术的提升，目前“整车-T-Box-物联网模块-射频前端芯片”的车联网产业链条太过冗长，未来可能出现整车厂与射频前端芯片厂商直接协同，推进车载射频连接技术快速发展的情况；
- **标准化将逐步完善，**车联网芯片需要遵循统一的标准和规范，以保证兼容性和互操作性，同时也需要支持标准的前向演进能力，以适应未来的技术变化和业务需求。目前车联网芯片的标准在伴随行业发展逐步演进中，车联网芯片的标准也将逐步完善。

5.17.5 总结

百年以来，通信技术、电子技术的发展，将人类带入了一下丰富多彩的智能世界。在这个世界里，电话、电脑、家用电器，这一切的物体，都在走向“移动”且“互联”。

汽车也不例外，自发明之日起，汽车就是一个已经实现“移动”的大家伙，但如何让汽车也实现“互联”却是通信从业者过去几十年不断努力的方向。

5G 的到来，使移动通信的关注点不再局限于手机；汽车的电气化也是智能化、网联化有了基础。“车联网”终于得以完整实现。

因为应用环境、功能需求、可靠性需求的不同，车联网对用于通信的射频芯片也提出了更多不一样的要求，射频芯片厂商也在不断地学习与提升，为车载应用提供更可靠、更便捷、更优性能的通信芯片。慧智微期待和您一起，共同学习更适合车联网的芯片设计。

您觉得车联网对芯片有哪些特殊需要？欢迎扫码在线留言讨论。

参考文献

1. GSMA，The Mobile Economy 2023，2023
2. How many connected cars are there worldwide , https://smartcar.com/blog/connected-cars-worldwide/
3. RWS-210135, Support lower power class for NR, SmarterMicro Inc. https://www.3gpp.org
4. ZTE promotes autonomous driving in the 5G era with commitment to the integrated C-V2X solution
5. 慧智微，5G 协议再演进：慧智微 3GPP 提案正式立项，2022
6. 慧智微，车联网，需要什么样的射频芯片?，2022
7. 《蜂窝车联网（C-V2X）技术与产业发展态势前沿报告（2020）》，中国通信学会白皮书, Nov，2020.
8. https://5gaa.org/ Feb 14th , 2022.
9. 佐思汽研，2022 年 C-V2X 和车路协同行业研究报告
10. https://www.communicationstoday.co.in/global-cellular-iot-module-shipments-grew-35-yoy-in-q1-2022/
11. https://www.rollingwireless.com/solutions

在线阅读

本文摘自慧智微微信公众号，扫描以下二维码，在线阅读



第六篇 射频工具及学习

本篇内容

- AI 会取代射频工程师吗？
- 射频最推荐书目榜
- 《射频技术百问百答》发布
- 《无线重构世界》正式全面开售
- 射频人专属鼠标垫更新
- 慧智微射频小程序

射频技术，以其独特的复杂性和抽象性，常常让人感到既神秘又充满挑战。然而，在这看似难以捉摸的技术背后，站立着的是一群鲜活的、充满创造力的射频工程师。正是我们这些工程师用智慧和汗水，一点一滴地推动着射频技术的日新月异，让无线通信的奇迹得以实现，让世界因连接而更加精彩。这一章，我们将把目光的焦点转向这些默默无闻的英雄——射频工程师。射频技术背后的工程师，不仅是技术的践行者，更是创新的推动者。

随着人工智能（AI）技术的蓬勃发展，射频领域也迎来了新的变革。AI 的引入，不仅改变了射频工程师的工作方式，还为他们提供了更强大的工具和更广阔的视野。我们将详细探讨 AI 如何影响射频工程师的工作，如何帮助他们更高效地解决问题，如何提升射频系统的性能和稳定性。

当然，学习射频技术离不开优质的资源和工具。因此，在这一章中，我们还将为大家推荐一些射频领域的经典书籍。这些书籍涵盖了从基础理论到实践应用的各个方面，是射频工程师成长的必备读物。同时，我们还将介绍一些实用的射频设计工具和仿真软件，帮助大家更好地理解和掌握射频技术。

向每一位射频工程师致敬！

6.1 AI 会取代射频工程师吗？

作者：彭洋洋 博士

2025 年春节假期期间，来自杭州的 AI 创业公司 DeepSeek 以惊人的速度，再次掀起了全球对 AI 技术的热潮。据报道，DeepSeek 的性能可与 ChatGPT 相媲美，但其训练和运行成本却显著降低。这些技术优势将有助于加速 AI 技术的广泛应用和普及。

在人们为低成本、高性能 AI 技术带来的“AI 自由”欢呼之际，也有声音开始担忧，AI 是否会取代现有的人类工作岗位。

射频领域同样受到了这一波浪潮的影响。2024 年 12 月 30 日，《自然》（Nature）杂志网站刊登了一篇关于 AI 在射频电路设计应用的文章：《基于深度学习的多端口射频及亚太赫兹无源器件和集成电路的广义逆设计》[1]。该文章展示了利用 AI 技术设计的射频电路，一时间在射频工程师群体中引起了广泛传播和热烈讨论。

文章中，普林斯顿大学的研究团队利用 AI 技术设计了多种射频微波电路结构。这些电路结构在人类看来相当抽象，如同一堆难以理解的二维码，但其射频性能却可能与射频工程师精心设计的电路不相上下。

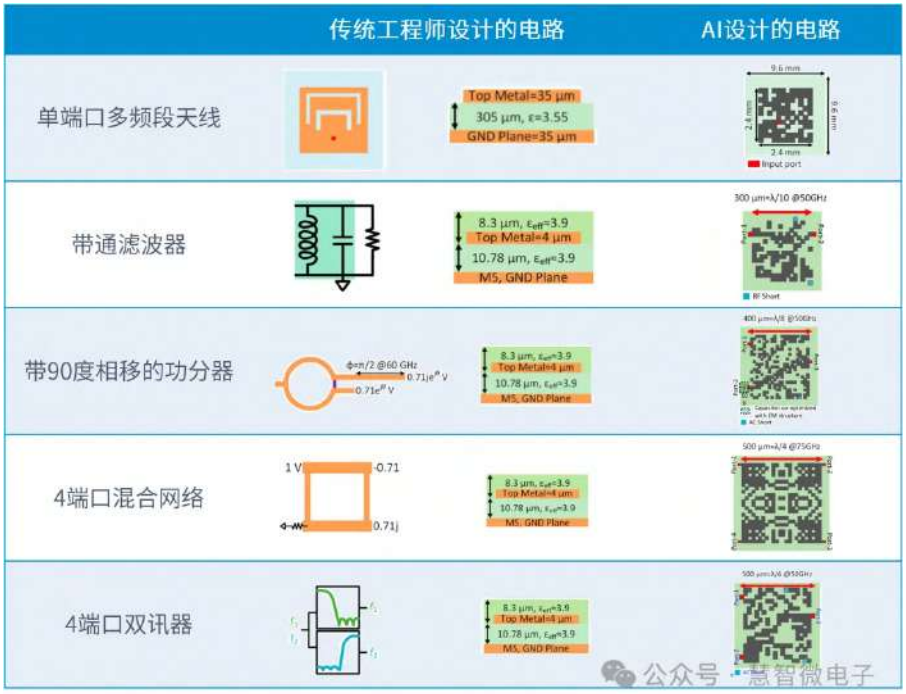


图 6.1: 人类理解的射频电路和 AI 设计的射频电路对比

此研究团队还运用这一设计方法，成功实现了基于 SiGe 工艺的毫米波频段放大器。要知道，在 20 年前，关于硅基毫米波放大器的研究足以在集成电路顶级会议 ISSCC 上发表论文；甚至在 10 多年前，博士生或研究生还可以凭借毫米波放大器的研究成果完成学位论文并顺利毕业。然而，如今这些曾经让射频工程师引以为豪的工作，却被 AI 以一种看似难以捉摸的方式轻松实现。

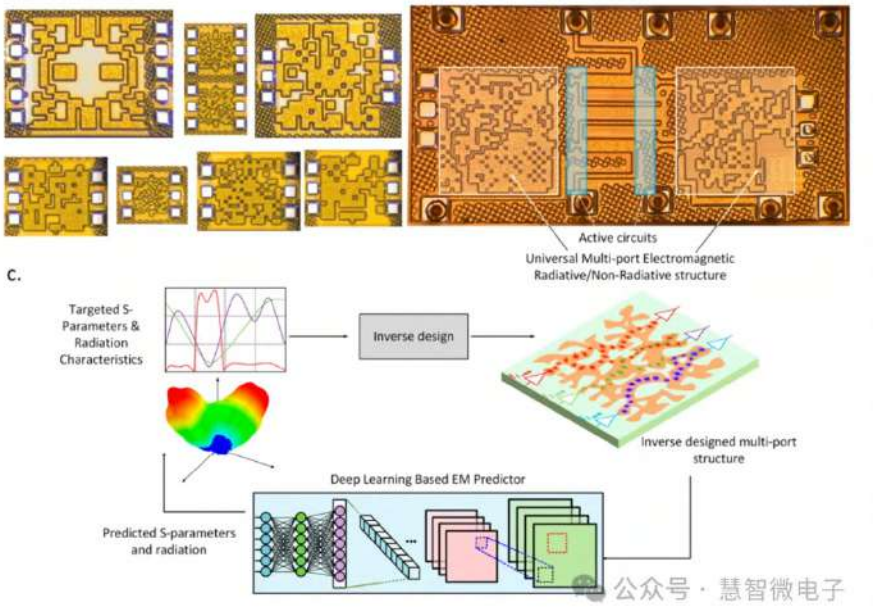


图 6.2: 文章中利用 AI 技术设计实现的毫米波放大器

随着 AI 技术的迅猛发展，及其在射频领域的初露锋芒，也让射频从业者不禁开始思考：**射频工程师会被 AI 替代吗？**

本文将从 AI 技术的特点和射频电路的特点出发，初探 AI 技术在射频领域的应用潜力，并抛砖引玉，探讨未来 AI 与射频技术可能的发展关系。

6.1.1 AI 的一些基础概念

尽管 AI 技术如潮水般涌来，似乎要席卷各行各业，但作为一项技术，它同样具有自身的特性和局限。要深入探讨 AI 能为射频技术带来何种变革，首先我们需要对 AI 的技术特点有所了解。

AI 领域涉及众多术语，如人工智能、机器学习、神经网络、深度学习等，这些名词听起来既抽象又科幻，让人难以捉摸。接下来，让我们试着揭开这些名词的神秘面纱，探寻它们的本质。

6.1.1.1 人工智能

在 AI 相关的所有名词中，人工智能无疑是最具代表性的。

AI，即 Artificial Intelligence 的缩写，这一术语有时显得过于抽象和魔幻，甚至带有一些营销的色彩，没有一个明确的定义，因此人们经常陷入关于什么是人工智能、什么不是人工智能的讨论中。

尽管人工智能没有一个清晰的定义，但这个名词的出现的时间点却有共识。大家普遍认为 1956 年的达特茅斯研讨会标志着人工智能的诞生。这次会议明确提出了人工智能的概念，并将其作为独立的学科方向进行研究。



图 6.3: 1956 年达特茅斯会议中的青年参会者

然而，当时对人工智能的想象远不如现在如此丰富。在 20 世纪 50 年代，将数学公式转化为计算机执行的指令就已经被视为人工智能了。但现在看来，这不过是简单的编译器功能而已。

2018 年图灵奖获得者、深度学习创始人杨立昆（Yann LeCun）在其专著中曾指出：“人工智能有个特点，就是一旦一个技术实现了，它立刻就会被认为不属于智能。”因此，真正的人工智能总是似乎遥不可及，就像吊在我们面前的一根胡萝卜，永远无法实现。这种现象也被称为“AI 效应”。



图 6.4: Yann LeCun 教授和其专著（中文译为：《科学之路——人、机器与未来》）

此外值得一提的是，虽然现在人工智能被视为技术发展中当之无愧的正确路径，但在当时的技术环境下，它并不是行业统一的技术共识。即使在 1956 年人工智能这一概念确立之后，仍然存在着较多的争论和讨论。

有人认为，这不过是让计算机处理更复杂的信息而已，那么我们为什么不称之为“复杂信息处理”呢？这样既能更准确地描述这项技术，又能避免使用过于科幻的名词。这一观点在当时也得到了很多专家的支持。直到 1961 年，IEEE 才正式将“人工智能”这一名词确定下来。

从技术角度来看，无论是称为“人工智能”还是“复杂信息处理”，或者其他什么名词，都不会对这项技术本身产生任何影响。但“人工智能”这一术语确实表达了人类对于计算机所能实现的智能疆界的不断追求，反

映了人类对于智能边界的不断探索。虽然这个名词抽象、不易定义，但它却成为了过去 60 多年来人类计算机信息技术发展的强大动力。

在 2006 年达特茅斯会议 50 周年纪念会上，当时开创人工智能概念的先驱们提到：如果当时用“计算智能”或者其他名称，这 50 年来的研究会不会有所不同？我个人认为，肯定会有所不同。因为名词的选择不仅影响着我们对技术的认知和理解，更在一定程度上引导着技术的发展方向和研究重点。甚至，这个名词还代表着人类对于创造智能不断探索的梦想。



图 6.5: 2006 年，达特茅斯会议 50 周年纪念会

6.1.1.2 如何建立复杂逻辑

无论学者们选择用“人工智能”、“复杂信息处理”还是“计算智能”来称呼这项技术，其本质都是期望计算机能够处理越来越复杂的信息和逻辑。

在教授计算机处理复杂信息和逻辑的技术实现中，主要有三个流派，分别是：符号主义、连接主义和行为主义。

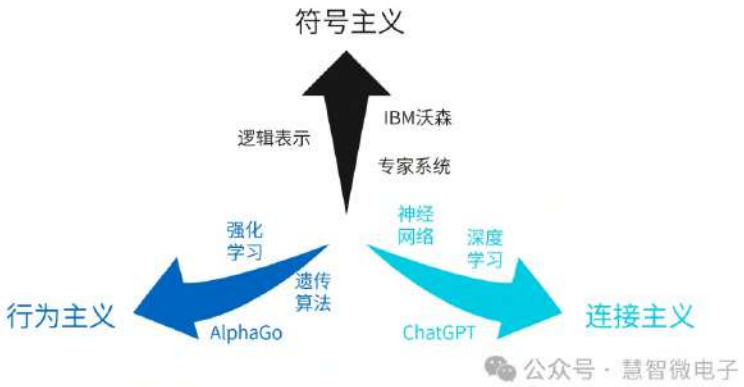


图 6.6: 建立复杂逻辑的三个学派

- **符号主义**

  符号主义（Symbolism）是人工智能发展的早期阶段，也是该领域中最早且最重要的学派之一。它起源于 20 世纪 50 年代，其理论基础源自数理逻辑和符号处理。

符号主义学派强调知识的表示和推理机制的设计，通过高层次的符号表示，如逻辑公式、规则等，来描绘问题和求解过程。这种基于符号和规则的方法为人工智能研究提供了重要的思路，奠定了该领域的研究基础。

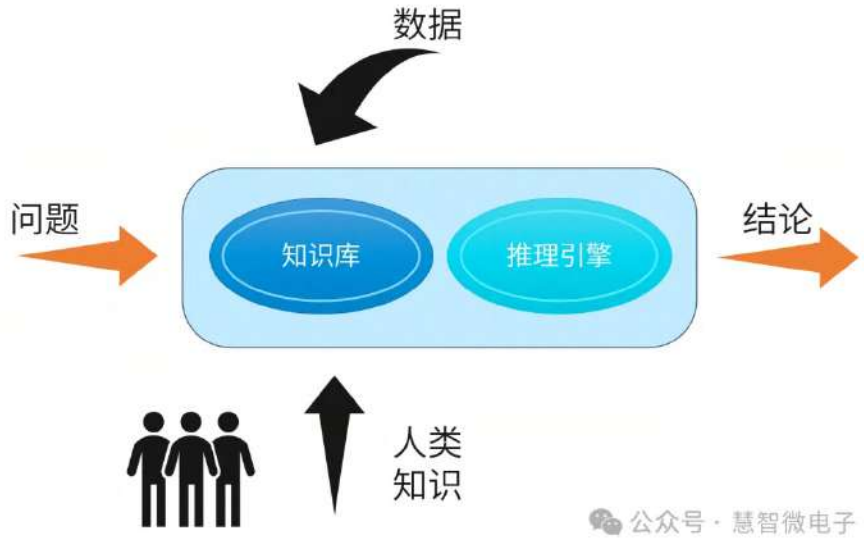


图 6.7: 符号主义的运行方式

符号主义学派的核心在于其明确的规则和可解释性。每一个推理步骤都基于预先设定的逻辑规则，计算过程是透明的。该学派的主要特点是将知识表示为符号，通过逻辑推理和规则操作来实现智能行为。

然而，符号主义也面临一些局限性。首先，它依赖于严格定义的规则和逻辑推理，这需要对知识进行精确编码。但现实世界中的知识和信息往往是模糊、不精确和动态的，这使得符号主义在处理不确定性和模糊性问题时表现不佳。

其次，由于符号主义人工智能系统需要预先定义的规则和知识，其可扩展性受到限制。在面对大规模或持续变化的数据时，由于无法预知新的规则和知识，系统很难适应和扩展。

此外，符号主义也难以有效地表示和利用常识知识。常识知识通常具有高度的复杂性和不确定性，难以用符号和规则来精确描述。例如，在自动驾驶中，“保持安全距离”是一个基本的驾驶原则，但“安全距离”具体是多少米，却取决于车速、天气、路况等多种因素，很难用规则来事先定义清楚。

当然，对于复杂的逻辑问题，可以尝试用更复杂的公式来表征。但随着命题逻辑公式的复杂性增加，判断其真假的计算量将呈指数级增长。这意味着使用符号主义的逻辑推理和规则操作来解决某些复杂问题时，可能会变得非常耗时和困难。

对于符号主义来说，建立人工智能就像构建一套专家系统。这位“专家”期望了解世间万物背后的原理，并用精巧的公式来描述世界的运行。这种认知方式在拓展人类理解那些规律明显、清晰且被准确定义的问题时非常有效，但在解决日常、复杂的问题时却显得束手无策。

这也像人类对世界的认识一样。早在 500 年前的伽利略时代，人类就理解了天体运动的规律，并总结成了精确的公式。但直到现在，人类也无法用定义的方式来清楚界定什么样的图像可以被判定为一只猫。

- **连接主义**

人工智能中的连接主义（Connectionism），也称为神经网络学派或仿生学派，是人工智能领域的一个重要流派。其核心思想是通过模拟人脑神经网络的结构和功能来实现人工智能。

连接主义的起源可以追溯到 20 世纪初期对大脑工作原理的研究。到了 20 世纪 80 年代，随着对大脑认知的深入以及计算技术的飞速发展，连接主义逐渐在人工智能领域中崭露头角，成为一个显著的流派。

连接主义的代表性成果是 1943 年由生理学家麦卡洛克（McCulloch）和数理逻辑学家皮茨（Pitts）创立的 MP 模型。然而，这种超前的概念在当时的技术和理论背景下，并没有得到学术界和业界的广泛接受。此外，在连接主义兴起的初期，计算机硬件和算法技术都相对落后，使得连接主义模型在实际操作中面临着

巨大的挑战。



图 6.8: 神经网络 MP 模型创始人：麦克洛克与皮茨（1949 年）

与符号主义相比，连接主义具有更强的数据处理能力和自适应能力。符号主义强调通过符号和逻辑推理来实现智能，而连接主义则将信息处理过程理解为不同信息元之间的连接和交互。因此，连接主义更擅长处理复杂、非线性问题和模式识别任务。

为了更直观地理解连接主义，我们可以以识别手写 0-9 数字的过程为例。

在识别手写的 0-9 数字时，我们并没有一个固定的公式来告诉我们哪个数字是什么。相反，我们会根据数字的某些特征，如闭合的圆圈数量、直线的数量、拐点的位置等，来判断这些数字最可能是哪个。这些特征可以理解为不同的“神经元”，它们与不同的数字结果相对应。

例如，如果看到一个闭合的圆圈，我们可能会认为这个数字是 0、6 或 9；如果看到一段直线，我们可能会认为这个数字是 1 或 7。通过这些“神经元”对手写数字的感知和判断，我们可以最大概率地确定这个数字是什么。

在这个过程中，我们并没有预先设定好的公式或规则，而是通过多个信息元的判断和它们与结果的连接来得出结果。这正是连接主义的基础思想。

基于这一思路，一些早期的科学家设计制造了连接主义的人工智能计算机。其中最具代表性的就是罗森布拉特发明的，可以识别手写数字的感知机 [11]。这项技术在当时看来颇具科幻感，它是通过采用大量连接线相连的网络实现的。

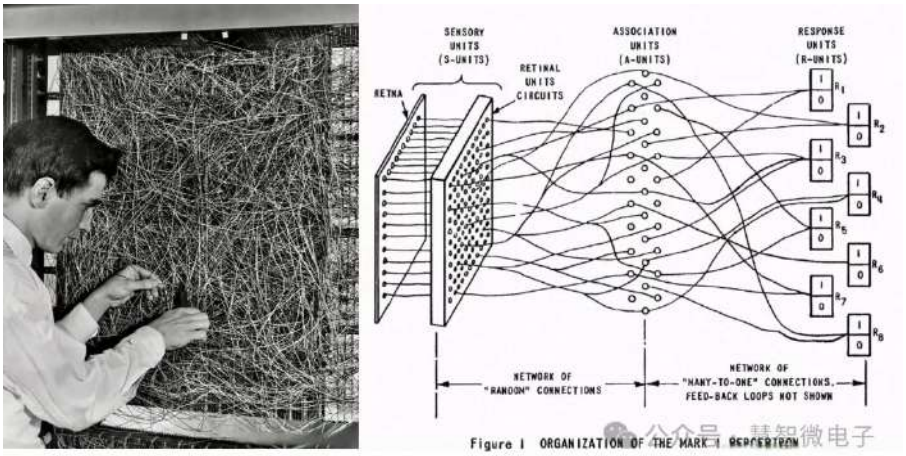


图 6.9: 罗森布拉特发明的感知机 Mark-I

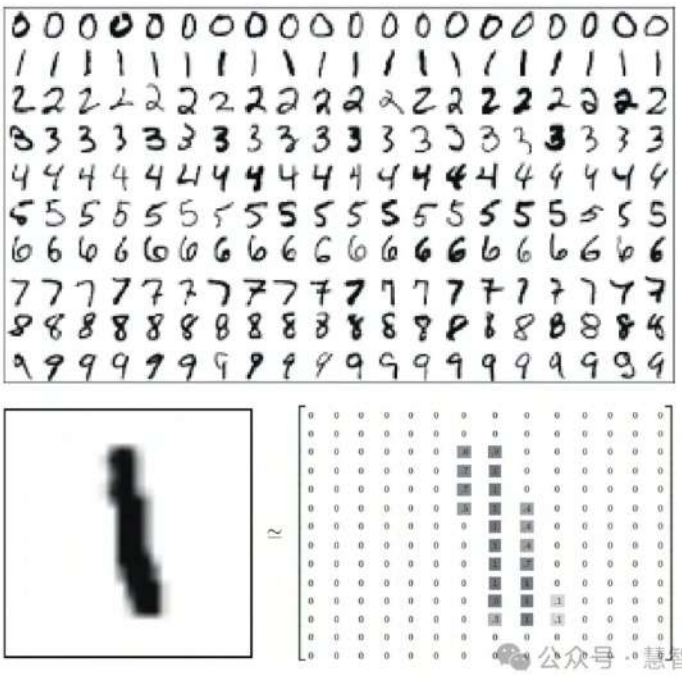


图 6.10: 早期用于数字识别训练的手写数字，以及字符“1”的识别

尽管现在连接主义已经成为人工智能最为重要的技术流派之一，但在当时，这一理论还是太过超前了。

当时，符号主义在人工智能领域占据主导地位。符号主义基于逻辑推理方法，强调用计算机的符号操作来模拟人的认知过程。这种方法在当时的技术背景下具有更强的可操作性，因此得到了学术界和业界的广泛认可。相比之下，连接主义由于理论超前和技术条件限制，难以与符号主义竞争。

然而，近年来随着深度学习、大数据以及计算资源的丰富和发展，连接主义得到了迅速的发展，并已经成为人工智能重要的发展方向之一。

- **行为主义**

  行为主义（Behaviorism）是人工智能领域的三大主要流派之一，它强调智能体（如机器人、软件系统等）与环境的交互作用，认为智能的本质在于能够根据环境的反馈来灵活调整行为。行为主义主张，智能不仅仅依赖于内部的逻辑推理或神经网络计算，更重要的是通过与环境的持续互动和适应来展现智能。

  在人工智能领域，行为主义有着广泛的应用前景，特别是在那些需要智能体自主适应环境、进行决策和行动的场合，如机器人技术、自动驾驶等领域。

  相较于连接主义，行为主义在智能实现上有着不同的侧重点。连接主义通过模拟神经元之间的连接机制来实现智能，而行为主义则更侧重于智能体与环境之间的动态交互作用。然而，近年来随着深度学习和强化学习技术的迅猛发展，联结主义和行为主义之间的界限逐渐变得模糊，两者开始相互融合，共同推动人工智能技术的进步。

6.1.1.3 神经网络

在人工智能技术领域，近期最为频繁提及的名词之一便是“神经网络”。“人工智能”这一概念已颇具抽象性，而“神经网络”则更添了几分科幻色彩。

相较于“人工智能”这一由科学家们为早期复杂信息处理技术所赋予的新名称，“神经网络”确实是由神经学和逻辑学两个交叉学科的杰出科学家共同发展起来的一门技术。

沃伦·麦卡洛克是一位神经生理学家，他曾在耶鲁大学学习哲学和心理学，后来获得医学博士学位，并在精神病学领域进行了深入研究。而沃尔特·皮茨则是一位年轻的数理逻辑学家，对数学和哲学怀有浓厚兴趣。他早年生活颇为艰辛，甚至有过流浪的经历，但最终凭借自己的才华赢得了芝加哥大学等学术界的广泛认可。

20 世纪 40 年代，神经生理学家沃伦·麦卡洛克与数理逻辑学家沃尔特·皮茨在芝加哥大学附近相遇，并因

对神经科学与逻辑学的共同兴趣而携手合作。1943 年，他们在《数学生物物理期刊》上发表了一篇开创性的论文，题为《神经活动中内在思想的逻辑演算》（A Logical Calculus of Ideas Immanent in Nervous Activity）。在这篇论文中，麦卡洛克和皮茨首次提出了人工神经元的数学模型，即 MP 模型（McCulloch-Pitts Model）。该模型基于生物神经元的结构和功能，通过数学方法抽象和简化了神经元的处理过程，具有以下特点：

- 每个神经元被视为一个多输入单输出的信息处理单元。
- 输入信号通过加权求和，并与设定的阈值进行比较，以决定是否产生输出。
- 输出为二进制形式（0 或 1），模拟了神经元的兴奋与抑制两种状态。

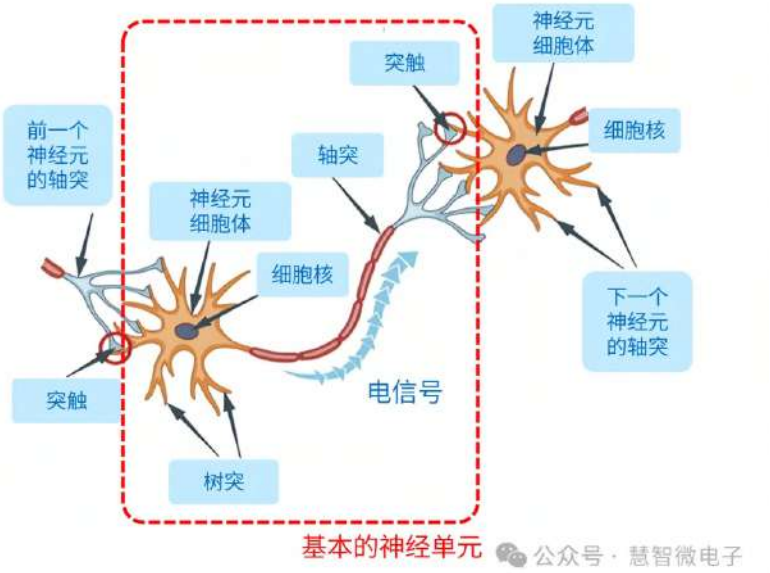


图 6.11: 脑神经元细胞感知信号示意图

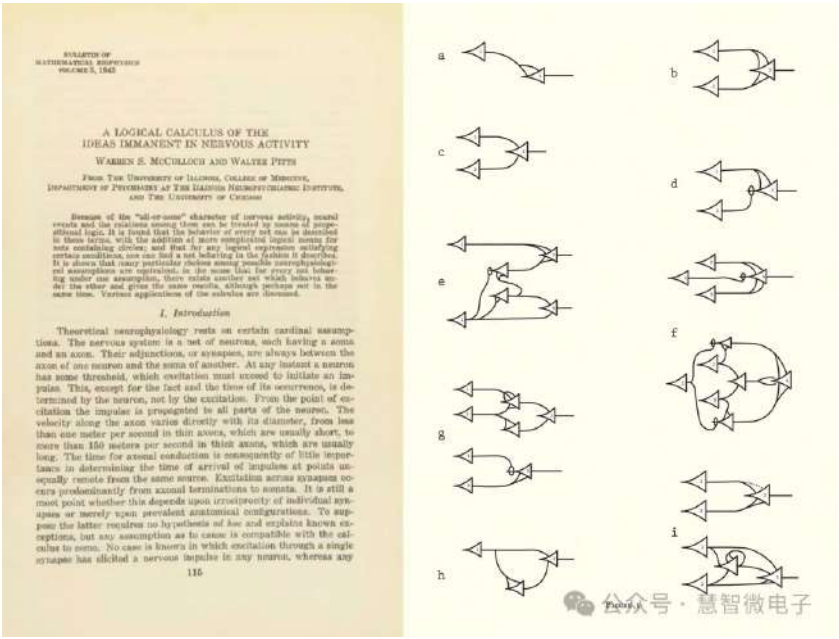


图 6.12: 麦卡洛克与皮茨在 1943 年发表的神经网络 MP 模型论文

麦卡洛克和皮茨的初衷并非建立复杂的计算机网络，而是鉴于当时神经科学和计算机科学均处于起步阶段，他们尝试用数学模型来理解和模拟大脑的工作方式，并希望通过这一数学模型揭示神经元处理信息的逻辑基础。

尽管 MP 模型相对简化，但它的提出为人工神经网络的研究奠定了坚实基础。该模型展示了如何通过简单

的数学规则来模拟神经元的逻辑运算，也激发了后来研究者们对神经网络复杂性和潜力的进一步探索。

MP 模型的提出激发了人们对神经网络的兴趣，但随后的研究也面临了诸多挑战。例如，单层感知器无法解决非线性可分问题，多层网络的训练算法也尚未成熟。然而，这些挑战并未阻挡研究者们的前进脚步。随着计算能力的提升和算法的优化，神经网络技术在 20 世纪 80 年代迎来了复兴，并在 21 世纪取得了巨大成功。

目前的神经网络通常由以下几个基本部分构成：

1. **神经元（Node/Neuron）：**
   神经元是神经网络的基本单元，负责接收输入信号，进行加权求和，并应用激活函数产生输出。
2. **层（Layer）：**
   层由多个神经元组成，神经网络通常包含输入层、隐藏层和输出层。输入层负责接收输入数据，隐藏层位于输入层和输出层之间，负责数据的加工和转换，输出层则输出最终的计算结果。
3. **权重（Weight）和偏置（Bias）：**
   权重和偏置是连接不同神经元的参数。权重代表一个神经元输出对另一个神经元输出的影响力，而偏置是加到加权和上的一个常数，可以看作是每个神经元的一个额外输入。
4. **激活函数（Activation Function）：**
   激活函数决定神经元是否应该被激活（即输出信号），并引入非线性，使神经网络能够拟合复杂的非线性关系。

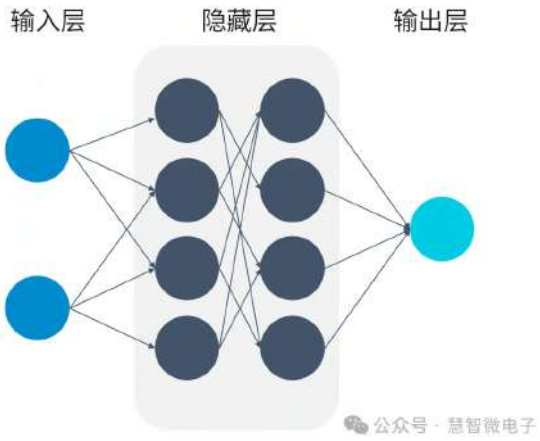


图 6.13: 人工神经网络的基本架构

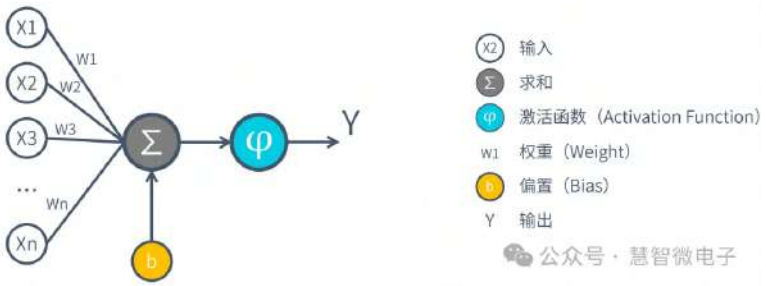


图 6.14: 人工神经网络中参数示意

这些名词初看起来可能有些复杂，但拆解开来并不难理解。每个神经元中的连接计算其实都较为简单：

- **神经元：**
  相当于被计算感知的信息元素，例如图像处理中的一个像素。
- **层：**
  将信息分层处理，便于理解。

最重要的三种层是：输入层（负责信息的输入）、输出层（负责结果的输出），以及隐藏层（其他层是计算机自己的运算过程，我们不需要关心，因此归为隐藏层）。

- **权重和偏置：**

  神经元获取到的信息需要经过处理。现在的神经元不再像最早的 MP 模型那样只有 0/1 的简单输出，而是要对输入的信息进行运算。

  例如，如果对输入的信息 x 进行线性运算，则输出可以表示为 $y = w * x + b$。其中，w 是信息处理的权重（表示信息的重要程度），b 是偏置（对输出结果进行修正）。

- **激活函数：**

  激活函数决定神经元是否被激活，并引入非线性。

  例如，可以使用 $y = w * x + b$ 这样的线性函数来表示 y 与 x 之间的线性关系，也可以使用如 $y = x^3$ 这样的非线性函数来拟合 y 与 x 之间的复杂关系。

通过以上步骤，一个神经网络就建立起来了。接下来，通过训练和优化，神经网络可以确定模型和参数。一旦神经网络有了准确的模型和参数，它就具备了处理复杂信息的能力。

6.1.1.4 机器学习

机器学习（Machine Learning，ML）技术是现代人工智能中的一项重要技术。它指的是一种通过算法和统计模型，使计算机系统能够自动地从数据中学习和改进的技术。这种技术不需要进行明确的编程，而是依赖于数据来训练模型，从而使其能够完成特定的任务。

机器学习之所以重要，是因为在信息化时代，数据量呈爆炸式增长。依靠传统的符号主义方法，如逻辑推理、理论抽取等，已经无法满足信息处理的需求。人们需要计算机能够自主学习和判断，以处理越来越复杂的信息。

机器学习能够高效地处理和分析海量数据，适应复杂多变的环境，提高决策效率和准确性，因此在当今信息时代有着强烈的需求。

机器学习技术的基本原理是利用大量数据训练模型，通过模型对未知数据进行预测和分类。具体来说，机器学习算法会从数据中提取特征，学习输入与输出之间的映射关系，然后利用这种关系对新样本进行预测。

根据学习方式的不同，机器学习技术可以划分为以下几种主要类型：

- **监督学习：**

  监督学习是指训练数据中有标签，计算机需要学习如何根据特征来预测标签。例如，通过给计算机展示大量的猫和狗的图片，并标注每张图片是猫还是狗，计算机就可以通过学习图片中的特征来识别未知的猫或狗的图片。

- **无监督学习：**

  无监督学习是指训练数据中没有标签，计算机需要自己发现数据中的模式和规律。例如，在大量的新闻文章中，计算机可以通过无监督学习来自动发现文章中的主题和关键词。

- **半监督学习：**

  半监督学习是指训练数据中只有部分标签，计算机需要利用已有的标签来预测未知数据的标签。例如，在一组商品评论中，只有一部分评论被标注为正面或负面，计算机可以利用这些已知的标签来预测未知评论的情感倾向。

- **强化学习：**

  强化学习是一种通过与环境的交互来学习的方式。计算机通过设定和获取奖励，不断尝试不同的动作，并根据环境的反馈（奖励或惩罚）来调整自己的策略，以最大化累积奖励。

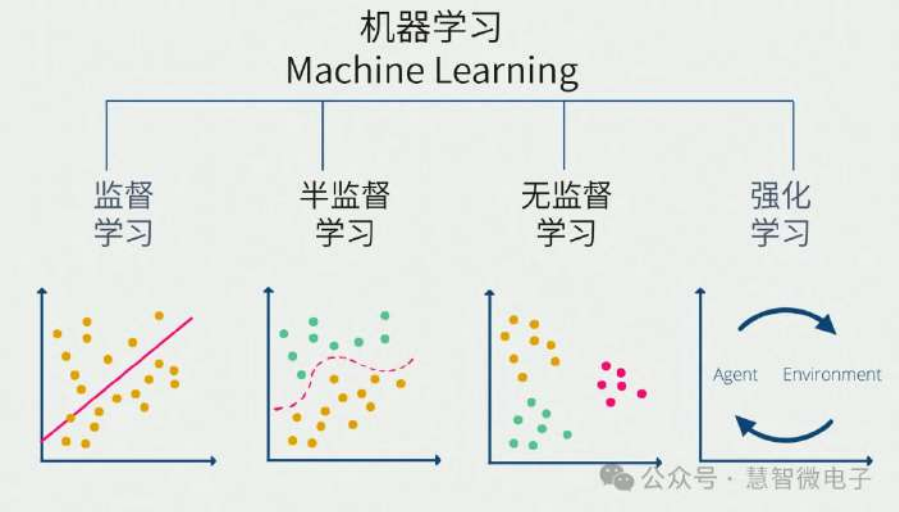


图 6.15: 机器学习的三种学习方式分类

虽然机器学习可以基于多种技术展开，但神经网络无疑为机器学习提供了一条极为重要的实现途径。

在早期，机器学习算法主要依赖于手工设计的特征，这不仅耗时费力，而且难以适应复杂的数据分布。神经网络的出现，使得计算机可以从数据中自动学习和提取特征，极大地简化了模型的设计过程。

如今，机器学习技术和神经网络技术已经密不可分。神经网络是机器学习中的一种重要技术，它推动了机器学习的发展，并在多个领域取得了显著成就。同时，机器学习也为神经网络提供了理论支持和技术手段，使得神经网络的结构和训练方法得到了极大的丰富和优化。两者相互增强，共同推动着人工智能技术的发展和进步。

6.1.1.5 深度学习

最近，人工智能领域中的另一个热门词汇是“深度学习”（Deep Learning）。以“Deep”开头的人工智能产品也屡见不鲜，比如我们熟知的杭州幻方旗下的 DeepSeek，以及 Google 旗下的人工智能公司 DeepMind。

那么，究竟为什么要进行“深度”学习呢？什么样的学习才算是有“深度”的呢？

深度学习的概念可以追溯到 2006 年。当时，加拿大多伦多大学的杰弗里 • 辛顿（Geoffrey Hinton）教授及其同事在《科学》杂志上发表了一篇关于深度信念网络（Deep Belief Networks）的研究论文 [2]。这篇论文的发表标志着深度学习概念的正式提出，并使得深度学习开始受到学术界的广泛关注。

辛顿教授所提出的“深度信念网络”其实是一种多层次的神经网络。然而，在当时，神经网络并未被学术界广泛接受，甚至受到符号主义学派的打压，导致关于神经网络的论文经常难以发表。于是，辛顿教授将其对多层次神经网络的研究更名为“深度信念网络”。没想到，这一略显被动的更名，却开创了人工智能领域一个重要的研究方向。

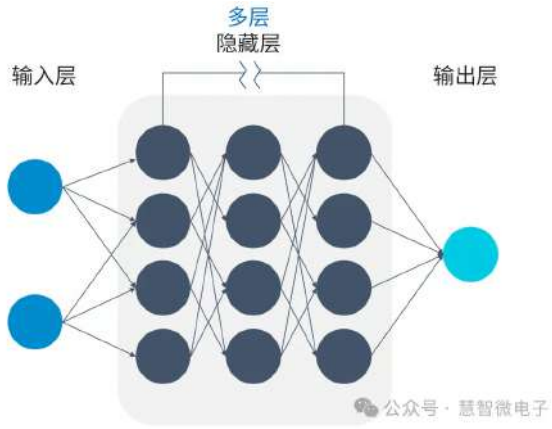


图 6.16: 深度学习中所用到的多层神经网络

在深度学习研究之前，虽然科学家已经实现了神经网络，但由于计算能力的限制和理论上的瓶颈，神经网络一直无法实现较多的层次，只能处理输出和输入之间较为简单的逻辑关系，这极大地限制了神经网络的效果。

多层神经网络通过增加隐藏层的数量，能够学习到更复杂、更抽象的特征表示。这种层级的特征学习使得网络能够处理更复杂的数据和任务，如图像识别、语音识别、自然语言处理等。

然而，早期神经网络的研究受到计算机性能限制、算法和技术局限性、数据集和数据限制、模型复杂度和表达能力不足以及理论研究不足等多方面的制约，多层神经网络一直难以实现。这也一度让神经网络的研究陷入了低谷，人们甚至开始怀疑神经网络的研究方向是否正确，还有不少论文证明为何多层神经网络是不可实现的。

2006 年，辛顿等人的工作为深度学习注入了新的活力。他们提出的深度信念网络通过无监督学习逐层初始化网络参数，然后使用有监督学习进行微调，有效地缓解了深层网络训练中的梯度消失问题。这一成果为深度学习的发展奠定了坚实的基础。

此后，深度学习在理论、算法和应用等方面都取得了显著的进展。各种新的模型和架构不断涌现，如卷积神经网络（CNN）、循环神经网络（RNN）、生成对抗网络（GAN）等。这些模型在图像识别、语音识别、自然语言处理等领域取得了突破性的成果。

如今，深度学习已经成为人工智能领域的重要分支，广泛应用于计算机视觉、自然语言处理、智能推荐、自动驾驶、医疗诊断等多个领域，为人类社会带来了前所未有的便利和创新。

值得一提的是，为表彰辛顿教授“为推动利用人工神经网络进行机器学习做出的基础性发现和发明”，2024 年 10 月 8 日，瑞典皇家科学院将 2024 年诺贝尔奖颁发给辛顿教授。

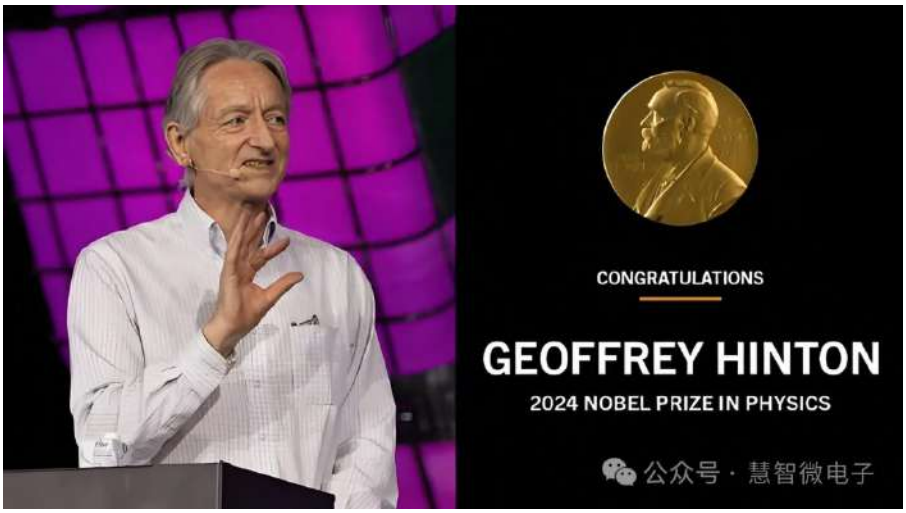


图 6.17: 辛顿教授于 2024 年 10 月 8 日获得诺贝尔物理学奖

6.1.1.6 ChatGPT/DeepSeek/Kimi 等的涌现

近年来，人工智能领域的发展速度迅猛，影响范围广泛，堪称科技史上的奇迹。其中，ChatGPT、DeepSeek、Kimi 等人工智能软件的涌现，更是将这一领域推向了新的高度，让人类深刻领略到了人工智能的魅力与潜力。

这些软件充分运用了神经网络技术，通过构建复杂而精细的多层神经网络，实现了深度学习。它们依托海量的数据和强大的算力，进行了多次的训练和优化，使得计算机在理解和生成自然语言方面取得了突破性的成果。如今，这些人工智能软件已经能够预测并模拟人类的言语交流，宛如被赋予了真正的智能，能够与人类展开自然、流畅的对话。

更为令人瞩目的是，与传统的符号主义人工智能相比，这些基于神经网络的人工智能软件其运行逻辑并非显式可见。在神经网络的运行过程中，我们只能观察到输入与输出的结果，而中间层均为隐藏层，我们无法直接窥探其内部的运作机制。这种“黑盒”特性，使得人工智能技术显得更加神秘且充满未知，也激发了人们对它更深入探索的兴趣。

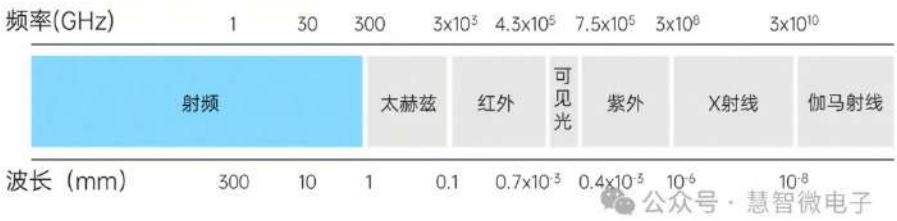


图 6.19: 射频的频率范围

6.1.2 射频电路的特点是什么

6.1.2.1 射频电路的功能

射频电路的主要功能是生成、发射及接收射频信号。

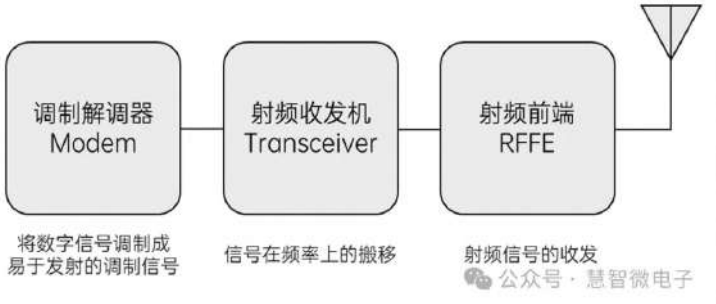


图 6.18: 射频电路在系统中的位置

射频信号位于电磁波频谱的 3kHz 至 300GHz 频率范围内，这段频率的电磁波波长适中，便于通过天线进行发射和接收，且具备较高的数据传输速率和较远的传输距离，因此非常适用于通信领域。如今，手机通信、无线网络、卫星广播等无线连接方式，都是基于射频技术实现的。射频技术已成为我们生活中不可或缺的一部分。

要实现射频连接，需要将数字和模拟信号转换为射频信号，并对这些信号进行频率转换、放大、滤波、切换等处理。完成这些信号处理的电路就是射频电路。

射频电路主要包括信号源、上下变频器、功率放大器、滤波器、开关以及低噪声放大器等模块。各电路模块的主要功能及其在系统中的位置如下图所示。

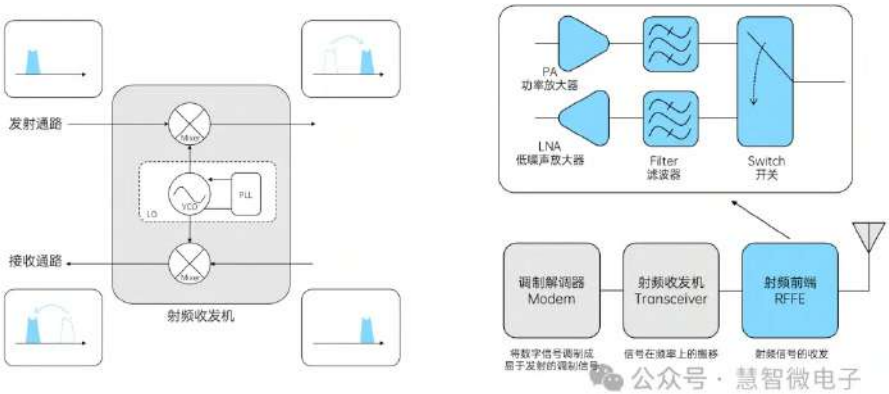


图 6.20: 射频系统中各电路的位置

具体来说，

- 信号源负责产生所需的射频信号；
- 上下变频器用于实现信号频率的转换，以适应不同的通信频段和处理需求；

- 功率放大器将信号放大至足够的功率水平，以便进行发射；
- 滤波器对信号进行频率选择性处理，滤除无用信号和噪声；
- 开关用于在不同信号路径或功能之间进行切换；
- 低噪声放大器则负责放大微弱的接收信号，同时尽量减少噪声的引入，提高信号的信噪比。

这些模块相互协作，确保射频信号能够在复杂的电磁环境中高效、准确地传输和处理。

6.1.2.2 射频电路的特点

作为射频电路，其与传统的模拟、数字电路相比，具有以下特殊之处：

1. **处理的是不可见的射频信号**

   射频电路所处理的是射频信号，这些信号以光速传播，肉眼无法直接观测，也无法通过简单的电压、电流方式直接表征。

   由于射频信号的不可见性，射频电路的设计和调试必须依赖于深入的理论分析、缜密的逻辑推理以及专业的指标测试，这大大增加了设计的难度和复杂性。

2. **工作频率高，传统电路分析方法不再适用**

   在低频电路中，通常可以忽略电磁场的影响，主要通过电压、电流等电路参数进行分析和设计。然而，在射频电路中，由于工作频率极高，电磁场的作用变得至关重要，传统的电路电压分析方法不再有效。

   高频工作条件下，会带来一系列非理想特性，如趋肤效应、辐射干扰、阻抗匹配问题等。这些高频特性给设计带来了诸多挑战，每个问题都需要仔细斟酌和解决。同时，高频信号的处理也要求射频电路具有更高的精度和稳定性，以确保信号的准确传输和处理。

3. **射频电路设计复杂，涉及多学科知识**

   射频信号在处理过程中环境复杂，因此射频电路的设计需要掌握多学科的知识，包括电磁学、电路理论，甚至材料科学等。这使得射频电路的设计过程相当复杂，要求工程师具备深厚的专业知识和实践经验。此外，射频电路的性能往往受到多种因素的影响，如电路布局、材料选择、制造工艺等，这都进一步增加了设计的难度和复杂性。

   有经验的工程师需要在各个指标之间进行权衡和折中，才能找到最佳的设计方案。著名的射频专家 Razavi 教授提出的射频电路设计“六边形”理论，就形象地展示了这种折中考虑的重要性 [13]。

4. **射频电路对参数极为敏感**

   由于射频信号在空中自由传输，对每个射频终端所发射信号的精准控制就显得尤为重要。同时，射频信号在传输过程中容易受到干扰、损耗等影响。因此，射频电路的性能对各个参数的精度极为敏感。

   微小的设计缺陷或参数变化都可能导致电路性能的显著下降。因此，射频工程师在设计过程中必须非常谨慎，对每一个细节都要进行仔细的考虑和优化，以确保电路的性能达到最佳状态。

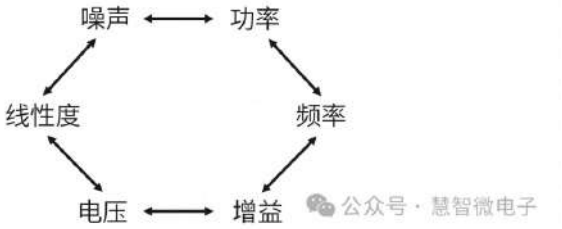


图 6.21: Razavi 教授的射频电路六边形设计理论图

6.1.2.3 为什么常说：“射频工程师越老越吃香？”

射频电路因其特殊性，使得射频行业成为了一个非常依赖经验的行业，这也正是“射频工程师，越老越吃香”这一说法的由来。

这一观点主要基于以下几个方面的原因：

首先，射频电路设计高度依赖实践经验。随着工作年限的积累，射频工程师会不断累积实际设计案例和经验，这些宝贵的经验是书本上无法获取的。当面对复杂的设计问题时，老工程师能够凭借丰富的实践经验迅速定位问题所在，并提出行之有效的解决方案。这种经验的积累，是射频工程师越老越吃香的重要基石。

其次，射频电路设计需要工程师具备敏锐的直觉和准确的判断力。在设计的初期阶段，很多问题可能无法通过理论计算或仿真来准确预测，这时就需要工程师依靠自己的经验和直觉来做出判断。这种直觉和判断力是长期实践和经验积累的结果，也是老工程师在设计中独具的优势。

再者，尽管射频技术在不断更新和发展，但其基础理论和设计方法并未发生根本性变化。老工程师通常具备扎实的基础理论知识和丰富的实践经验，**这使得老工程师在学习新技术时能够更快地理解和掌握，**并将其灵活应用到实际设计中。因此，即使面对新技术的挑战，老工程师也能够保持强大的竞争力。

正是由于以上原因，业界普遍认为射频行业门槛较高，难以一蹴而就。而工程师们一旦跨越了射频这个门槛，就可以通过长期积累和实践，步步高升。在电子类专业中，射频工程师宛如“铁饭碗”般的存在。因此，当 AI 浪潮席卷而来时，相较于软件工程师、数字电路工程师而言，射频工程师所感受到的冲击与震撼显得更为强烈。

6.1.3 AI 在射频设计中的应用

无论我们是积极拥抱还是有所抵触，AI 都已经毫无疑问地给我们的生活带来了重大的变革。在射频设计领域，AI 也在悄然渗透，并展现出了一系列应用潜力。

回顾过去，AI 在射频电路设计中的应用主要可以划分为以下三个阶段：

1. 阶段一：AI 辅助优化阶段
2. 阶段二：AI 自动化设计阶段
3. 阶段三：AI 自主创新阶段

这三个阶段展示了 AI 在射频设计领域逐步深入和拓展的应用过程，预示着 AI 将在未来射频技术发展中发挥更加重要的作用。

三个阶段的核心差异如下：

维度	AI辅助优化	AI自动化设计	AI自主创新
人类参与度	高（需预设框架）	中（定义目标和约束）	低（仅需设计需求）
AI创造性	无（局部优化）	有限（基于规则）	高（突破传统设计）
技术成熟度	成熟（工业界广泛应用）	发展中（部分场景落地）	前沿（实验室阶段）
典型算法	遗传算法	图神经网络、规则引擎	机器学习、生成式AI

公众号 · 慧智微电子

图 6.22: 三个阶段的核心差异

6.1.3.1 阶段一：AI 辅助优化阶段

在这个阶段，AI 作为辅助工具，对人类设计的初始方案进行局部参数优化，以提升射频电路的性能。

射频电路中的参数优化，是指通过调整电路中的关键参数，如晶体管的参数、电感、电容、阻抗等元件的值，以及电路的布局、布线、传输线长度等物理设计因素，来改善电路性能的过程。其主要目的包括提高信号质量、减少信号失真、增强电路稳定性以及提高能量传输效率等。

以射频电路设计中的阻抗匹配参数优化为例，这是一个耗时且依赖经验的过程。传统方法需要工程师不断调整短截线的长度和位置，并反复测量调整后的驻波比（SWR），效率较低。而在 1993 年的文献 [4] 中，提出了一种基于神经网络的自动调谐方法。该方法通过神经网络实时分析 SWR 的变化，自动调整短截线参数，实现了阻抗匹配的快速优化。实验结果表明，神经网络设计的双短截线网络与传统方法的结果一致，且调谐过程无需

人工干预。这一工作充分展示了 AI 在射频参数优化中的潜力，能够显著减少设计时间，为复杂电路的快速优化提供了新思路。

此后，陆续有文献对射频参数优化展开深入研究，并将其应用在天线参数设计、功放设计等领域。近年来，AI 在功率放大器（PA）设计方面的参数优化也成为研究热点之一 [6]。对此感兴趣的读者可以搜索相关文献 [7] 进行深入了解。

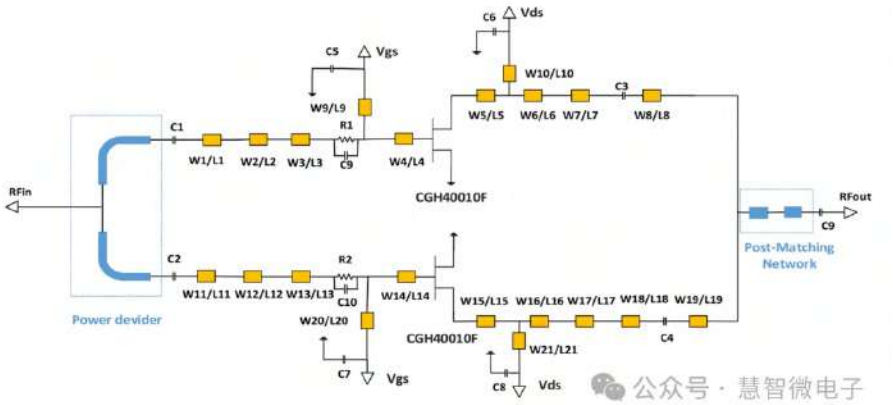


图 6.23: 文献 [6] 中 Doherty PA 架构及所优化参数

此外，AI 还可以用来辅助射频电路中耗时的仿真过程进行加速。例如，在文献 [5] 中，Horng 等人提出了一种基于神经网络的微带电路设计方法。该方法通过神经网络学习全波仿真生成的输入-输出数据，实现了对微带电路性能的快速预测。利用神经网络的强大映射能力，该方法显著减少了仿真时间，提高了设计效率。

虽然在这一阶段，AI 已经能够明显提升射频电路的设计效率，但设计的主导权仍然掌握在工程师手中。AI 只是在人类定义的范围内，做一些提升效率、减少仿真时间的工作，扮演的是辅助角色。因此，人们并没有将其视为跨时代的变化。

6.1.3.2 阶段二：AI 自动化设计阶段

随着射频技术的飞速发展，射频电路设计的复杂度也随之大幅提升。传统的人工设计方法逐渐暴露出两大瓶颈：

1. 一是设计参数的爆炸式增长；
2. 二是规则约束的复杂性，设计需满足多重限制条件。

在阶段一中，AI 虽然能辅助优化参数，但设计框架仍由人类工程师主导，需要大量人工介入。为了减少人工的长时间参与，并充分考虑设计中的多种规则和约束，AI 自动化设计阶段应运而生 [8]。这一阶段的目标是让 AI 基于规则和约束，自动生成完整的设计方案，实现“设计半自动化”。

这一阶段的核心特点是 AI 从辅助工具升级为“设计助手”，具体表现为：

- 规则驱动设计：AI 严格遵循物理规律（如阻抗匹配原理）和行业标准（如 PCB 布局规则），确保设计方案的可行性。
- 端到端流程覆盖：从输入设计目标到输出完整的电路图、版图，整个流程无需人工分步干预。
- 多目标协同优化：AI 能够同时优化增益、带宽、功耗等多个指标，找到全局最优解。

在数字电路领域，AI 的自动化设计已经过多年的研究，如数字电路的自动布局布线等，已成为大规模数字集成电路设计中的必备流程。近年来，AI 在射频电路方面的自动化设计工具也开始崭露头角。

RFIC-GPT 是冉谱微电子（上海）有限公司研发的一款射频芯片智能设计工具。它基于高级算法和数据分析技术，旨在帮助工程师更高效、更准确地完成 RFIC 设计工作。RFIC-GPT 利用 AI 技术，根据设计者提供的电学指标，自动生成满足要求的射频器件和电路的 GDSII 或原理图。据介绍，RFIC-GPT 生成的设计方案准确度高达 95%，在大多数情况下可以直接应用于实际生产中，无需过多修改。目前，该工具已支持无源器件和电路（如电感、变压器、匹配电路、宽频多值匹配）以及有源电路（如功率放大器 PA）的设计 [9]。



图 6.24: 冉谱微 RFIC-GPT 自动生成变压器版图的界面

RFI Pro 是广州十方尺科技有限公司推出的一款旨在提升射频 IC 设计效率的自动化工具。它通过 AI 技术实现了射频版图与原理图的联合优化，显著减轻了工程师在设计过程中的负担，提高了设计的准确性和效率。针对射频 IC 设计中的版图迭代耗时、电路与无源网络优化脱节、复杂无源结构设计难度高以及射频 IP 复用性差等痛点，RFI Pro 实现了从有源电路到无源版图的协同优化。该工具具有 AI 驱动的版图与原理图协同优化、全面的版图阅读和修改能力以及与 Virtuoso 和 EMX 的无缝衔接等亮点，提高了设计效率，支持灵活多样的优化目标，并让用户能够轻松上手 [10]。

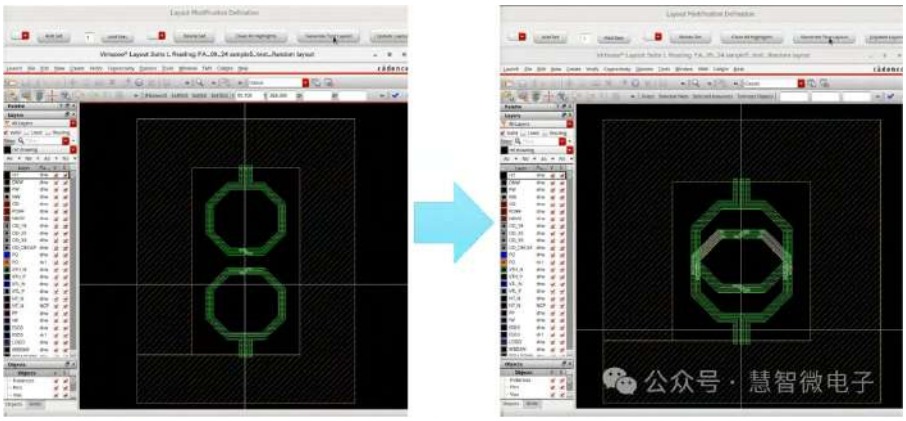


图 6.25: 广州十方尺科技 RFI Pro 展示的变压器版图自动跳线功能

6.1.3.3 阶段三：AI 自主创新阶段

近年来，AI 技术迅猛发展，其强大的数据处理能力和学习算法在各个领域都展现出了巨大的潜力。在射频设计领域，AI 的应用也不再局限于辅助人类经验，而是开始引领新的设计潮流。

射频行业虽然已历经百年发展，但由于射频概念的抽象性和复杂性，人类对射频的掌握仍然有限。传统的射频设计往往依赖于工程师的经验和直觉，这在一定程度上限制了设计的创新性和效率。然而，随着 AI 技术的引入，我们有望突破人类经验的局限，探索更多未知的设计空间。

这一阶段的目标是让 AI 完全主导设计过程，从零开始生成创新性方案，突破人类经验的限制。例如，可以让 AI 从一个抽象的设计目标出发，直接生成电路方案，完成端到端的设计，而不对 AI 进行预设规则的限制。

2024 年，Karahan 等人在《Nature》（自然）杂志上发表了一篇论文 [1]，他们通过引入深度学习模型，成功

构建了一个强大的电磁仿真器。这个仿真器能够准确预测任意平面多端口电磁结构的散射和辐射特性，极大地减少了传统电磁仿真所需的时间和资源，使得设计过程更加高效。

传统射频电路设计受限于预定义的模板和拓扑结构，设计空间相对有限。而文献中提出的逆设计方法能够处理任意形状和结构，极大地扩展了设计空间。这使得设计师（尤其是 AI 设计师）能够探索更多未知的设计方案，推动射频设计的创新和突破。

在这种方法下，AI 已经设计出了多个设计实例，包括多频带天线、带通滤波器、功率分配器、正交混合器等。这些设计方案均展现出了优越的性能指标，如宽带响应、低插入损耗、高幅度平衡等。

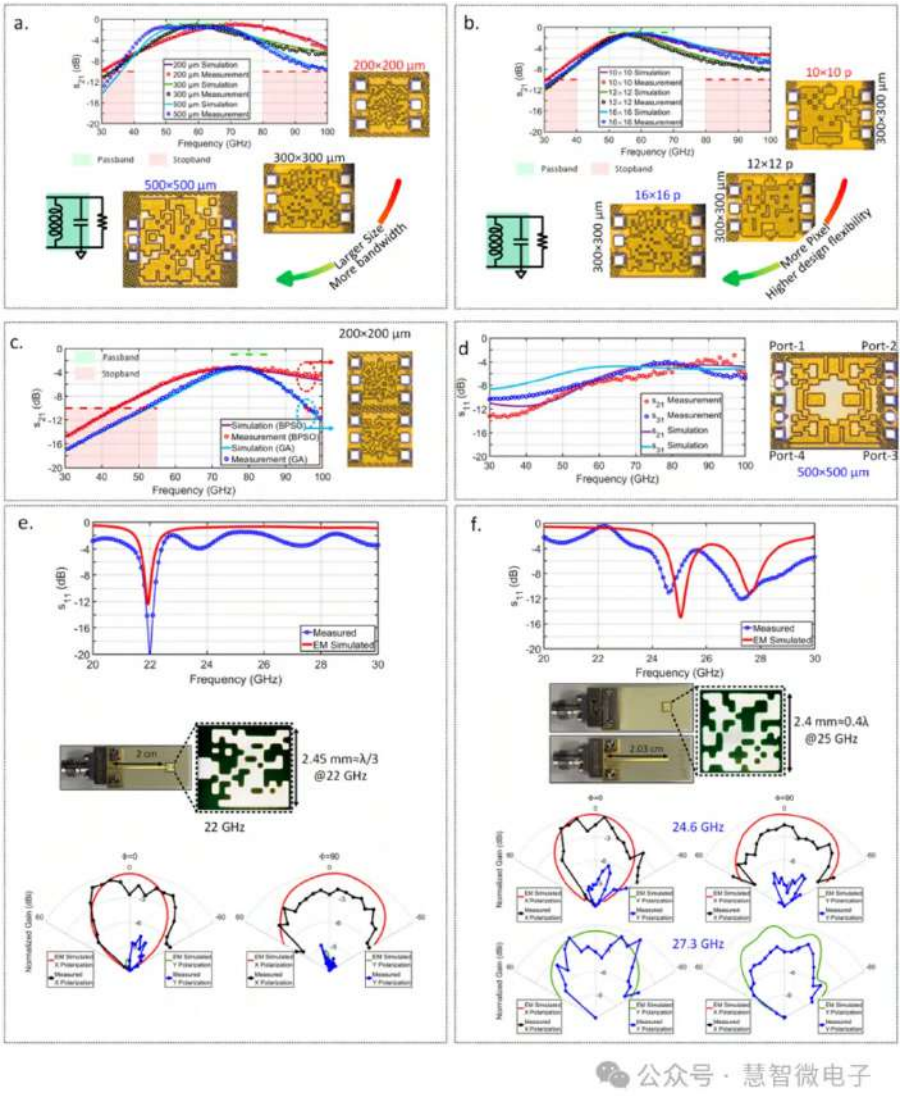


图 6.26: AI 设计的天线图案及测试结果对比

这些设计实例与传统射频电路在外观上存在巨大差异，但它们的性能却无一例外地优越。这为未来 AI 自主创新设计的射频电路提出了新的发展思路，也预示着射频设计领域将迎来一场革命性的变革。

6.1.4 射频工程师的未来塌了吗？

在 AI 技术迅猛发展的今天，以 ChatGPT/DeepSeek 等为代表的通用 AI 工具的出现，无疑展现了其巨大的功能与潜力，短短两年时间便引发了社会各界的广泛关注与讨论。而最近 AI 技术向射频领域的渗透，也让不少射频从业者感到惊讶，甚至有人提出了“射频工程师的未来塌了吗？”的疑问。

我认为并不会，而且射频工程师的未来反而因为 AI 而变得更加值得期待。

6.1.4.1 AI 浪潮不可逆：射频工程变革已成定局

ChatGPT 仅用两年便验证了通用 AI 的工程化能力，而 DeepSeek 等专用工具的出现，使 AI 开发成本下降 90% 以上，性能提升 300%。这意味着 AI 渗透速度正在加速，也预示着 AI 技术即将全面渗透到千行百业，这仅仅是 AI 时代的序章，一切都刚刚开始。

对于射频工程师而言，这一趋势同样不容忽视。尽管射频领域长期以来强调工程师的丰富经验与深厚积累，但 AI 的到来，无疑将像在其他行业一样，为射频工程带来本质性的变革。

6.1.4.2 AI 可以帮助射频工程师完成能力跃迁

面对这一变革，射频工程师无需恐慌或担忧被 AI 取代。相反，可以将 AI 视为强大的能力放大器。

AI 不仅能够助力工程师快速掌握电路知识，突破以往职业生涯中仅能精通两三个电路模块的局限，使掌握十种甚至更多电路模块成为可能。更重要的是，AI 能够协助工程师全面掌握射频系统，克服射频技术因系统知识复杂而难以驾驭的难题，让工程师们不再如盲人摸象般片面理解射频系统，而是能够全局把控。

在工作效率方面，AI 的引入更是带来了革命性的提升。原本需要数天乃至数月精心优化的电路，在 AI 的辅助下可能仅需数小时便能完成，极大地提高了工作效率，让工程师能够将更多时间投入到更具创意与价值的工作中。

正如“加油射频工程师”公众号中的作者袁月 Yue 所言：“作为脑力劳动者的我，还是要努力拥抱 AI 啊！它脑子这么好使，可以使劲问问题啊！不满意，还可以连续追问；不信任它的答案，还可以让它把来源说清楚。”这或许才是 AI 到来之时，射频工程师应有的态度。

6.1.4.3 与工程师竞争的始终是人类，而非 AI

与射频工程师竞争的，始终是人类自身，而非 AI。

正如在棋类领域所展现的，AI 虽曾让人类高手一时难以适应，但如今已成为人类提升棋力的得力助手，不断推动人类对弈的新高度。人类不会去纠结人类和 AI 谁下棋更强，而是会利用 AI 工具，帮助人类探索未知的棋类世界。

中国国际特级大师，历史上第十七位国际象棋世界冠军在采访中就表示：“AI 的下棋质量远超人类，和 AI 对弈是自讨苦吃”。但丁立人会经常使用 AI 来帮助自己探索不同的棋局，不断让自己的棋力得到提升。



图 6.27: 国际象棋世界冠军丁立人的采访

在射频技术领域，这一幕也将重演。随着 AI 技术的发展，人类可能在很多方面做的并不如 AI。但人类将借助 AI 工具，不断拓宽研究的边界，提升技术的水平。最终的对弈与竞争，始终是人类与人类之间的智慧碰撞。

另外，虽然原本需要数天乃至数月精心优化的电路，在 AI 的辅助下可能仅需数小时便能完成，但于 AI 是否会取代人类工作，实则无需过多忧虑。当前 AI 的操作仍需在人类指导下进行，即便是未来进入自主创新阶段，AI 设计出的电路也需人类进行判定与验证，确保其实用性与正确性。

判断与决策，始终是人类的独特优势，正如那句玩笑话所言，“AI 取代不了人类，因为只有人才能背锅”，实则强调了人类在决策与责任承担上的不可替代性。

6.1.5 总结

近年来，AI 技术迅猛发展，特别是以 DeepSeek 为代表的先进 AI 工具，给各行业工程师带来了前所未有的冲击。随着这些工具功能的日益强大和成本的显著降低，射频从业者不禁开始思考：未来射频工程师是否会被 AI 取代？

然而，我认为 AI 技术的到来并非射频工程师的天塌了，反而为工程师们打开了一扇全新的大门。AI 与人类并非简单的取代关系，而是协同共进的伙伴。AI 擅长处理复杂逻辑关系、分析海量数据、进行高效计算，而这些正是射频工程师在学习和工作中所亟需的能力。AI 技术结合射频工程师的创意、审美以及对射频物理本质的深刻理解，将助力射频工程师实现能力跃迁。

射频工程师可以积极拥抱 AI 技术，将其作为提升自身能力的工具。在 AI 的辅助下，工程师们可以更加专注于创新设计，探索射频技术的未知领域。在 AI 的强大加持下，AI 可以与人类共同推动射频技术的进步，开拓射频技术的边界限制，开创射频技术的辉煌未来。AI 的到来，不是射频工程师的末日，而是可以使我们飞翔的翅膀更加坚实有力。

您觉得 AI 到来会给射频电路设计带来什么变化？欢迎查看在线文章，留言讨论。

参考文献：

1. Karahan, E. A., Liu, Z., Gupta, A., Shao, Z., Zhou, J., Khankhoje, U., & Sengupta, K. (2024). Deep-learning enabled generalized inverse design of multi-port radio-frequency and sub-terahertz passives and integrated circuits. Nature Communications, 15(1), 10734. https://doi.org/10.1038/s41467-024-54778-1
2. Hinton, G. E., and Salakhutdinov, R. R. ”Reducing the dimensionality of data with neural networks.” Science, vol. 313, no. 5786, 2006, pp. 504-507. DOI: 10.1126/science.1127647.
3. Razavi, B. RF Microelectronics. 2nd ed., Prentice Hall, 2011.
4. Vai, M., and Prasad, S. ”Automatic impedance matching with a neural network.” IEEE Microwave and Guided Wave Letters, vol. 3, no. 10, 1993, pp. 353-354. DOI: 10.1109/75.244867.
5. Horng, T. S., Wang, C. C., and Alexopoulos, N. G. ”Microstrip circuit design using neural networks.” IEEE MTT-S International Microwave Symposium Digest, 1993, pp. 413-416. DOI: 10.1109/MWSYM.1993.276947.
6. Guo, J., Crupi, G., and Cai, J. ”A Broadband Asymmetric Doherty Power Amplifier Design Based on Multiobjective Bayesian Optimization: Theoretical and Experimental Validation.” IEEE Access, vol. 10, 2022, pp. 89823-89834. DOI: 10.1109/ACCESS.2022.3200421.
7. 李博士. ”AI 射频功放设计.” 知乎专栏, 2023. [Online] Available: https://zhuanlan.zhihu.com/p/5891287585.
8. Smith, J., and Johnson, K. ”Automated RF Circuit Design Using a Hybrid Neural Network and Genetic Algorithm Approach.” IEEE Transactions on Microwave Theory and Techniques, vol. 68, no. 5, 2020, pp. 2100-2112. DOI: 10.1109/TMTT.2020.2975812.
9. 冉普微. IC Prophet. [Online] Available: https://www.icprophet.com/.
10. 广州十方尺. Sephonchi Technology. [Online] Available: https://www.sephonchitech.com/.

11. 周志明. 智慧的疆界 [M]. 机械工业出版社, 2018.
12. 彭洋洋. 无线重构世界: 射频技术的过去、现在与未来 [M]. 电子工业出版社, 2024.
13. B. Razavi. RF Microelectronics[M]. 北京：清华大学出版社, 2003.

在线阅读

本文摘自慧智微微信公众号，扫描以下二维码，在线阅读



6.2 射频最推荐书目榜-慧智微射频问答群

作者：小慧，发布时间：2022 年 6 月

在射频技术问答群中，热心的群友推荐了许多射频领域优秀书籍，其中包括：

1. 无线通信系统类：7 本
2. 射频微波电路设计类：16 本
3. 射频 PA 设计类：7 本
4. 其他：2 本

问答群的读者们也对以上书籍进行了投票，评选出“射频最推荐书目榜”，现将榜单公布如下：

6.2.1 无线通信系统类：



图 6.29: 无线通信系统类图书

榜 1.《无线通信中的射频收发系统设计》

作者：Qizheng Gu

推荐理由：无线通信系统介绍比较全面的书，实战性强，适合射频系统设计人员。

榜 2.《射频微电子》（第一版）

作者：拉扎维

推荐理由：系统类的好书，比第二版薄，适合轻松读。

榜 3.《Microwave and RF Design of Wireless Systems》

作者：Pozar

推荐理由：微波届的大神 Pozar 的另外一本神书，和微波工程不同，这本书主要讲系统的。

6.2.2 射频微波电路设计类：

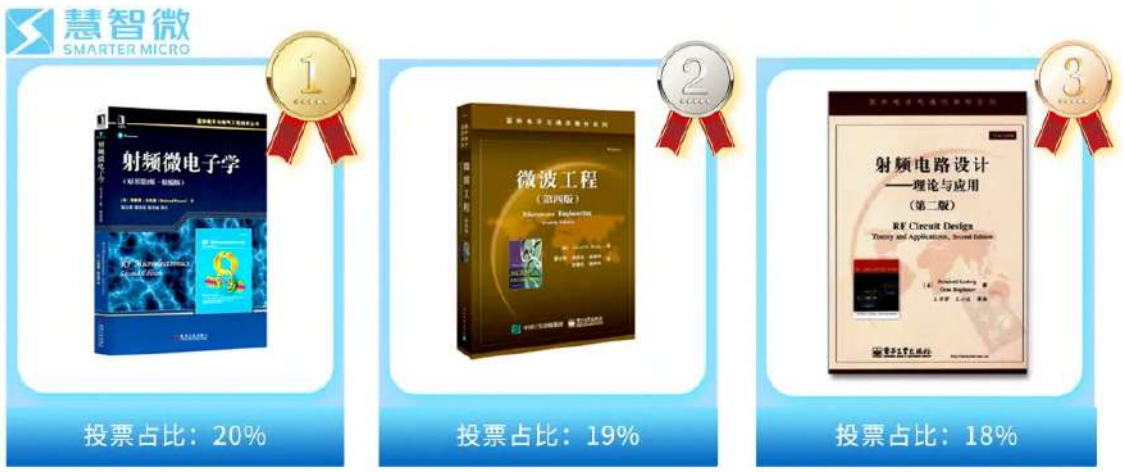


图 6.30: 射频微波电路设计类图书

榜 1.《射频微电子》（第二版）
作者：拉扎维
推荐理由：拉神的书，写的通俗又不失严谨，适合入门看。

榜 2.《微波工程》
作者：Pozar
推荐理由：微波届的教科书，介绍全面、基础，必读书目。

榜 3.《射频电路设计射频电路设计-理论与应用》
作者：Reinhold Ludwig
推荐理由：经典书目，介绍全面，从基本理论到电路设计都有涉及。

6.2.3 射频 PA 设计类：



图 6.31: 射频 PA 设计类图书

榜 1.《RF Power Amplifiers for Wireless Communications》

作者：Cripps

推荐理由：Cripps 的神书，做 PA 必看。

榜 2.《射频与微波晶体管功率放大器工程》作者：张玉兴

推荐理由：本书内容丰富，几乎绝版。是一部很经典的功放设计指导书籍，它涉及到功放设计的方方面面，包括晶体管结构，功分合成网络，预失真，线性化，偏置电路，效率提升等，偏向于解决实际工程应用问题，内容全面，墙裂推荐！

榜 3.《Microwave Transistor Amplifiers, Analysis and Design》

作者：Gonzalez

推荐理由：射频微波放大器的教科书，讲的扎实且精准，值得将书中每一章节细细读，每一道题仔细做。

6.2.4 其他类：



图 6.32: 其他类图书

榜 1.《半导体物理与器件》

作者：Donald A. Neamen

推荐理由：半导体物理器件的基础书目，不过公式和理论比较多，适合有一些半导体物理背景的人看。

榜 2.《The Unwritten Law of Bussiness》

作者：King, W, J

推荐理由：名字里有“bussiness”，但书里确全是工程师工作中的那些“不能说的准则”，新人工程师一定要看的训练书目。国内有翻译版，名字取成了“第一商规”，不过还是建议看英文版。

以上书单来源于慧智微射频技术问答群，感谢群友们无私分享，同时欢迎更多粉丝参与优秀书单征集活动，如果您有更多书籍想要推荐给大家，可在加入问答群后留言。为支持大家参与互动，我们将为精彩留言赠送纪念礼品。

在线阅读

本文摘自慧智微微信公众号，扫描以下二维码，在线阅读



6.3 慧智微射频技术百问百答发布

作者：小慧，发布时间：2024 年 4 月

射频技术，听起来有点高大上，它其实是改变我们生活的重要力量。从手机通信、无线网络，到卫星导航、远程遥控，都离不开射频技术的支持。它是连接世界的桥梁，让信息传递变得无比迅速和便捷。

射频技术重要，但并不好学。射频的教材普遍老旧，但射频技术却日新月异，让人应接不暇。那些复杂的公式、繁琐的参数，更是让人头疼不已。每个射频人对射频的学习都是一个漫长的过程。我个人在这个过程中，发现与专家们的问答交流是学习最快、最有效的方式。他们的解答往往能一语惊醒梦中人，让我豁然开朗。

两年前，2022 年 4 月 7 日，为了方便大家对射频技术的交流讨论，我们建立了“**慧智微射频问答群**”。在这个群里，没有广告，没有闲聊；不问来处，只谈技术。我们最初的想法只是有一个地方可以交流一些射频学习的心得，没想到，短短两年时间，问答群已经汇聚了 2000 多位射频的朋友，积累了近 500 条精彩的问答记录。在问答群里，我们发现，有时候问题的答案并不是最重要的，解答和讨论的过程同样珍贵，甚至更加精彩。大家各抒己见，相互启发，往往能碰撞出意想不到的火花。这些讨论不仅解决了我们对射频技术的疑惑，还拓宽了我们的视野和思路。于是每隔一两周，我们都会将这些问答过程整理成集，方便大家查阅和学习。现在问答集已经出到第 40 多期了，每一期都凝聚着大家的智慧和心血。

“君子之学必好问，问与学，相辅而行者也。”群友们不断提问、热心回答，大家在射频的知识海洋里一起遨游，逐步掌握射频的基本知识，共同探讨射频的最新技术。这种互帮互助的学习氛围，让我们在射频学习的路上不再孤单，也让我们对射频技术有了更深入的理解。

这些宝贵的交流和讨论记录，我们不敢独享。在问答群设立两周年的日子里，**我们将过去两年的问答集结成册，作为第一部《射频技术百问百答》分享给每一位热爱射频技术的朋友**。这本册子不仅仅是一本射频技术问答集，更是射频问答群群友共同成长的见证。它记录了我们探索的历程、思考的过程，还有收获的喜悦。

另外需要说明的是，本册并不是严肃的教科书，整理过程中我们也没有过分苛求讨论的正确性。本册只是如实记录射频学习者们的讨论和思考过程，希望这个记录能给大家的学习带来启发。

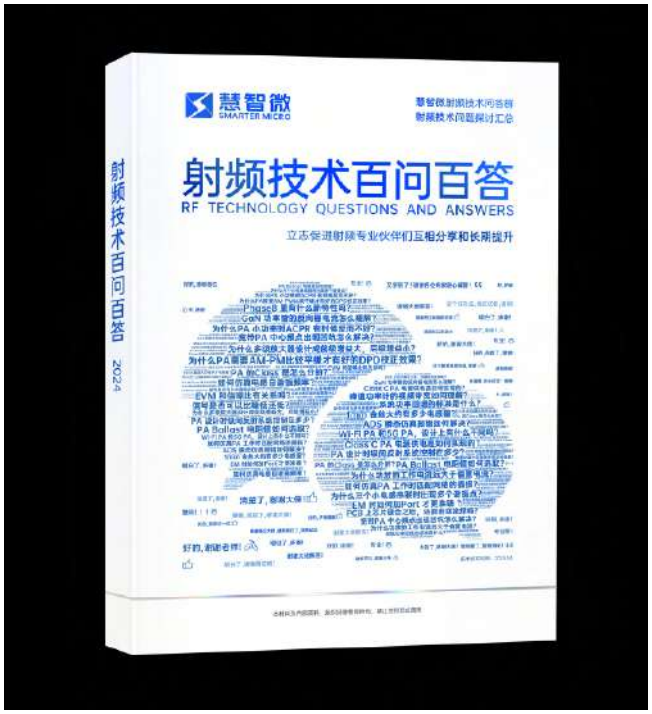
感谢每一位为问答群做出贡献的朋友。你们的热情参与、无私奉献和智慧分享，让问答群充满了活力和温暖。展望未来，我们希望这个射频技术问答社区能够继续发展壮大。我们期待更多的朋友加入进来，共同分享知识、交流经验、探讨问题。无论你是初学者还是资深专家，都期待这里是你学习射频的乐园。

这本册子是我们共同努力的成果，也是我们未来共同继续前行的动力。让我们一起在射频技术的道路上携手前行，创造更加美好的未来！

欢迎关注慧智微公众号，回复“射频技术百问百答”，下载本册。



图 6.34: 扫描关注慧智微公众号



目录

第一章 PA 设计技术 14
1.1 PA 的小信号特性 15
1.1.1 什么是“FET 晶体管等本身具有 6dB/倍频程的滑盖滤除” 15
1.1.2 为什么 $G_{s,max}$ 是 6dB/8 倍频程的滤除 15
1.1.3 为什么 G_{max} 会自动折点 16
1.1.4 为什么多级放大器设计成功增益大，后级增益小 17
1.1.5 PA 仿真时为何选择频率完成 17
1.1.6 宽带 PA 中心频点出现增益怎么解决 18
1.1.7 有关于稳定因子 K 和 B 具体推导过程的文献吗 20
1.1.8 功放在设计时要保证每一级都在全频段稳定吗 20
1.1.9 有关 PA S11 的讨论 20
1.1.10 PA 设计时，什么时候需要考虑噪声系数 21
1.2 PA 的匹配设计 22
1.2.1 哪里有非功放的短路匹配如何设计 22
1.2.2 PA 漏端为什么用 1/4 传输线经电源 22
1.2.3 5-6GHz PA 输出匹配如何设计 24
1.2.4 PA 设计时，级间反馈系统控制在多少 24
1.2.5 多级 PA 级间匹配如何设计 24
1.2.6 多级 PA 级间匹配如何设计再讨论 25
1.2.7 PA 输出网络 Q 值的影响 26
1.2.8 毫米波宽带功放的定义，是以功率还是 S21 定义的? 26
1.2.9 宽带 PA 设计中输入阻抗的位置 27
1.2.10 宽带功放各频点最佳阻抗有差距时，如何匹配 27
1.2.11 功放小信号短波较差怎么处理 27
1.2.12 PA 的输入功率去哪了 28
1.3 PA 的偏置设计 29
1.3.1 PA 工作电流跟哪些方面相关 29
1.3.2 PA 的静态工作点怎么确定 29
1.3.3 PA Ballast 电阻值如何选取 29
1.3.4 HBT PA 中基极流电阻和射极镇流电阻的优劣势 31
1.3.5 放大器馈电电感如何取值 31
1.3.6 放大器馈电电感第再次讨论 31
1.3.7 Class C PA 电源馈电是如何实现的 33
1.3.8 为什么功放的工作电流远大于偏置电流 33
1.3.9 为什么 HBT 管直和实测电流差了 50% 34
1.3.10 为什么 PA V_{c1}/V_{c2} 合路后电流会出现变化 34
1.3.11 PA 栅漏上电顺序有要求吗 35
1.3.12 npn 管不给 Base 偏置，为何自动开启了 35
1.4 PA 的大信号特性 36
1.4.1 多级 PA 如何判断每一级 AM/PM 指标是否合适 36
1.4.2 为什么 PA 增益曲线在压缩前会出现膨胀 36
1.4.3 为什么关 PA 增益在压缩之前为上升几个 dB 37
1.4.4 多级 PA AM-AM 拓扑哪种情况下会更好 37
1.4.5 有讲怎么测试 AM-AM/AM-PM 特性的资料嘛 38
1.4.6 多级 HBT PA 中 AM-PM 如何调试 39
1.4.7 GaAs HBT PA 设计时 AM-PM 不平坦如何调整 41
1.4.8 HBT PA 非线性电容对 AM-PM 的影响 41
1.4.9 为什么 PA 加入驱动级 AM-PM 恶化 43
1.4.10 为什么 PA 需要 AM-PM 比较平缓，才有好的 DPD 校正效果 43
1.4.11 PA 的 detrough table 有什么用 43
1.4.12 为什么 PA 过饱和时效率会下降 43
1.4.13 PA 在-13dB 就压缩了该怎么改善 44
1.4.14 为何 PA 大功率输出时增益会降为 0dB 45
1.4.15 PA 仿真为什么出现增益和功率回缩 45
1.4.16 PA 功率一直得到 P_{1dB}/P_{3dB} 工作会有什么影响 46
1.4.17 PA 双音仿真为什么 P_{1dB} 会减小 46
1.4.18 单音信号为何比调制信号饱和功率低 1dB 47
1.4.19 主信号 P_{1dB} 时，叠加一个小信号，会是什么状态 47
1.4.20 不同 Class 的 PA 的 P_{1dB} 和 OIP3 规范怎么比较 48
1.4.21 PA 什么情况下应用只需要考虑饱和功率，不需要考虑线性度 48
1.4.22 为何 PA 在某个频率下 PAE 下降 49
1.4.23 为什么 PA 饱和状态下的动态电流小于静态电流 52
1.4.24 有关 PA 电源推导的讨论 54
1.4.25 PA 邻低频 ACLR 变差的原因是什么 55
1.4.26 为什么 PA 小功率时 ACPR 有时候反而不好 55
1.4.27 为什么 PA 测试时 5M 8RB 线性度差，如何解决 56
1.4.28 为什么 Wi-Fi PA 不同信号需要下的功放饱功率相差 4-5dB 58
1.4.29 Wi-Fi PA 和 5G PA，设计上有什么不同吗 58
1.4.30 为什么双音测试中，同阶的信号幅度不等代表产生了记忆效应 59
1.4.31 为什么 PA 记忆效应在小信号下表现的更明显 59
1.4.32 PA 为什么 0 频内侧阻抗有不同 60
1.4.33 PA 单管为什么输入要控制二次谐波 61
1.4.34 为什么谐波能量随频率变化 61
1.5 PA 的 Load-line 和 Load-pull 63
1.5.1 有关共轭匹配和负载线匹配的之讨论 63
1.5.2 Load-pull 中，如何选择合适的负载匹配点 63
1.5.3 当 R_{Load} 点于最优 Load-line 时，为何要设置电压 64
1.5.4 负载大于最佳负载值的时候，电流是否继续增加吗 64
1.5.5 有关 PA 输出负载线及可靠性工作区域的讨论 65
1.5.6 大信号 S 参数是不是和 Load-pull 没有区别 69
1.5.7 为什么 PA 最大功率和最大效率很不重合 70
1.5.8 如何对晶体管进行大信号去嵌 71
1.5.9 PA Load-pull 为什么在 Smith 圆图的左上部分 72
1.5.10 Load-pull 仿真中谐波阻抗设置的影响 73
1.5.11 HBT PA 设计时二次谐波阻抗如何设置 74
1.5.12 有关 Load-line 电压/电流波形的讨论 75

目录

1.5.13 B 类的功放当考虑 C_{ds} 时，如何计算效率 75
1.5.14 Class-B PA 为什么偏置电流为 0 时，偏置电压不为零 76
1.5.15 有关 PA 功率计算的讨论 79
1.6 PA 的 Stability 及杂散 81
1.6.1 单级 PA 输入部在不稳定区，如何做输入匹配 81
1.6.2 多级 PA 不稳定的 Debug 82
1.6.3 为什么功放自激杂散是低频下 82
1.6.4 如何保证 PA 大信号下的稳定性 82
1.6.5 PA 大信号稳定性是如何仿真的 83
1.6.6 谐振和自激这两个概念有什么区别 86
1.6.7 为何在 PA V_{cc} 端口并联电容对部波抑制不明显 87
1.6.8 为何无源特性-40dB 的滤波器对部波只抑制了 10dB 88
1.6.9 PA 平坦度调如何解决 88
1.6.10 PA 输出信号为边有效包络数 90
1.6.11 PA 频谱泄露会跟的原因 91
1.6.12 PA 震荡问题的 debug 讨论 91
1.6.13 如何解决 PA Mismatch 下的啸波 Fail 92
1.6.14 GaN 功放高频自激如何消除 93
1.6.15 稳定性仿真架构设置 Frequency Step 94
1.6.16 放大器自激振荡是振态变化的现象讨论 95
1.7 PA 的 Ruggedness 及可靠性 99
1.7.1 GaAs HBT PA 测试容易烧毁，是什么原因引起的 99
1.7.2 如何设计 PA 能抗 10:1 的 VSWR 失配 99
1.7.3 CMOS 功管设计 PA 中的可靠性考虑 100
1.7.4 PA 出现金丝烧断的原因 101
1.7.5 为什么 GaN PA 过热动平由现电流消失，又可自动恢复 102
1.7.6 测试 PA ruggedness 时 CW 信号与调制信号哪个更严苛 103
1.7.7 GaN 器件栅极漏电导应现象 103
1.7.8 为什么 HBT 超过标称电流密度使用也不会损坏 104
1.7.9 功放烧毁原因讨论 104
1.7.10 负载调动半径仿真出来吗 105
1.7.11 有关 PA 可靠性的讨论 106
1.7.12 GaAs HBT PA 设计时的热效应应如何考虑 108
1.8 PA 的仿真技术 110
1.8.1 PA 设计中，代入实际器件后效率恶化的原因 110
1.8.2 ADS 仿真用 Permute 调数公式报错 110
1.8.3 是否可以用脉冲测试管都设计连续波 PA 111
1.8.4 多级 PA 如何仿真级间匹配 111
1.8.5 ADS 节点功率仿真结果显示不正常 111
1.8.6 ADS 大信号仿真不收敛如何解决 112
1.8.7 ADS 如何仿真管芯处的电压与电流波形 113
1.8.8 仿真中如何得到随输入功率的变化图 113
1.8.9 如何仿真 PA 工作时匹配网络的插损 116
1.8.10 ADS 仿真受未选择器件的影响 117
1.8.11 奇模 (Odd-mode) 稳定性如何仿真 117

5

目录

2.1.4 为什么需要在低噪放的末端加 testing buffer 155
2.1.5 如何减小 MOSFET 沟道电流热噪声 155
2.1.6 用 CMOS 管实现超宽带 LNA 用哪架构比较好 155
2.1.7 LNA 源极和栅极电感耦合太大有什么影响 156
2.1.8 LNA 的栅极加大电阻到地，为什么电阻越大，噪声越低啊 156
2.1.9 噪声消除电路为什么只有噪声被消除，而信号不被消除的讨论 156
2.1.10 在 Cascode 电路中，CG 管栅电阻如何选择 158
2.1.11 在 Cascode 电路中，CG 管栅电容如何选择 161
2.1.12 Stacked FET V_{dd} 电压如何选取 161
2.1.13 为什么 SOI 的 layout 都是下 gate 形式 164
2.1.14 有关放大器设计中，总宽度不变的情况下 2finger 和 3finger 的区别 164
2.1.15 ESD 设计中，栅保护管需要加或 ESD 画法吗 164
2.1.16 放大器 OIP3 是调定低吗 165
2.1.17 LNA 如何跑高 OIP3 165
2.1.18 LNA Layout 后性能恶化的讨论 166
2.1.19 差分放大器中共模电感的作用 166
2.1.20 毫米波 LNA 为何稳定回到地孔换成 EM 仿真模型后出现结果差异 167
2.1.21 LNA 设计中 RC 并联电路作用是什么 167
2.1.22 为什么 LNA 测试结果与仿真相差较大 168
2.2 LNA 的稳定性设计 169
2.2.1 LNA 设计必须全频带 $K_f > 1$ 吗 169
2.2.2 LNA 频段外 $K_f < 1$ 可以吗 169
2.2.3 多级 LNA 设计需要每级都看稳定吗 169
2.2.4 为了提高 LNA 的单级稳定性，能在源极加入电阻吗 170
2.2.5 这个混频器稳定是 LNA 引振吗 170
2.2.6 为什么没有电源滤波会导致低噪放自激 170
2.2.7 CS-CG 结构稳定性问题的探讨 172
2.2.8 差分放大器共模稳定性如何仿真 172
2.3 LNA 的仿真技术 176
2.3.1 ADS 和 Virtuso 如何仿真 k_f/f_{max} 176
2.3.2 怎么仿真耦合电容 178
2.3.3 在 ADS 中如何计算晶体管跨导 178
2.3.4 怎么在 ADS 看管子的二阶和三阶的导 179
2.3.5 Cadence 设计 LNA，到底才算是输入匹配好了 179
2.3.6 Cadence 或者 ADS 可以仿真电路某个部分的噪声或噪声 180
2.3.7 三阶 CMOS 提参和普通管子有什么不一样吗 180
2.3.8 ADS 仿真放大器 SimGamma 不能正常显示 180
2.3.9 有关 DRC 报错问题 183

第三章 开关设计技术 184
3.1 开关技术 185
3.1.1 开关设计中，串并联和四分组结构如何选择 185
3.1.2 除了 MOS，还有什么工艺能做射频开关 185
3.1.3 PIN 二极管做开关的原理 185
3.1.4 pHEMT 开关中，当源端电压高于开启电压的现象讨论 187

7

目录

3.1.5 开关仿真为啥漏为有电压 187

第四章 滤波器设计技术 190
4.1 滤波器技术 191
4.1.1 5G 频段里都会用到哪些滤波器技术 191
4.1.2 滤波器衰减点为何有大抽报 191
4.1.3 双工器 Tx 到 Rx 隔离度指标 191
4.1.4 电感空间辐射对双工器隔离度的影响 192
4.1.5 滤波器群延时的计算公式是什么 194
4.1.6 滤波器有工艺角偏差吗，为什么滤波器也有饱和功率 194
4.1.7 什么集总元件滤波器仿真与实测截止频率相差较大 195
4.1.8 ADS 优化滤波器时哪个算法比较好 196

第五章 无源设计技术 197
5.1 匹配网络设计 198
5.1.1 匹配网络损耗的计算 198
5.1.2 匹配网络箱计算 2 199
5.1.3 如何理解匹配网络 Q 值和带宽、损耗之间的关系 200
5.1.4 2pi 阻抗如何转到匹配到 50Ω 201
5.1.5 为什么源和负载阻抗相同，S11 却很差 202
5.1.6 如何往算相邻之接下就 204
5.1.7 给定架构，如何设计成宽带匹配 204
5.1.8 如何控制匹配网络随频率变化趋势 205
5.1.9 有关传输线阻抗匹配时阻抗发散问题 205
5.1.10 S11 在阻抗 Smith 圆图上形成一个振环，从物理意义上怎么解释 205
5.1.11 有关传输线阻抗匹配时阻抗突变问题 207
5.1.12 匹配设计的时候需要尽量走双层金属线吗 208
5.1.13 放大器级间 Transformer 匹配电路设计思路 208
5.1.14 片上匹配用变压器如何设计 209
5.1.15 哪里有介绍 Transformer 做匹配设计的资料 209
5.1.16 超宽带天线如何实现宽带匹配 210
5.2 无源集总元件 211
5.2.1 贴片电容的寄生电感如何估算 211
5.2.2 为什么大的 MIM 电容用两个小电容等效后，Q 值没有提高 211
5.2.3 为什么集片电感使用时要注意方向 212
5.2.4 大电压工作时电容应如何设计 213
5.2.5 宽频耦合 VCO 中 MOS 管可变电容如何推导 213
5.2.6 MIM 电容承受电流是看有效值还是瞬时峰值 215
5.2.7 Bonding 线寄生大约是多少 215
5.2.8 如何解决 Bonding 线寄生电感影响 PA 高频增益的问题 216
5.2.9 电感 Q 值有哪些优化手段 216
5.2.10 如何理解偏置电感会拉大电压摆幅的讨论 216
5.2.11 电感内部和周围 Dummy 金属填充应注意什么 219
5.2.12 为什么三个小电感串联时由现多个谐振点 219
5.2.13 为什么 ADS FEM 和 Momentum 仿真好片电感差别很大 220

8

目录

5.2.12 为什么三个小电感串联时由现多个谐振点 219
5.2.13 为什么 ADS FEM 和 Momentum 仿真好片电感差别很大 220

第二章 LNA 设计技术 150
2.1 LNA 设计技术 151
2.1.1 怎么理解 LNA 反射系数和噪声系数之间的 Trade-off 关系 151
2.1.2 如何从公式原理解释 LNA 增益和灵敏性的关系 152
2.1.3 如何确定 LNA 输入匹配的位置 154

6

1.8.12 Cadence 仿真 Load-pull 时，θ 和 gamma 这两个参数代表的意义是什么 118
1.8.13 ADS 链优化时报错 118
1.8.14 为什么 PA 的功率 Load-pull 仿真不收敛 119
1.8.15 Load-pull 仿真不收敛性的解决 119
1.8.16 为什么 ADS 仿真 PA 大信号，S21 是一条向上的直线 119
1.8.17 Fab 数据手册里写它们为 PA 制作，能否做 Switch、LNA 使用 121
1.8.18 ADS 瞬态仿真报错如何解决 121
1.8.19 PA 设计 AM-PM 仿真的 debug 124
1.8.20 有关 HB 仿真 debug 的讨论 126
1.8.21 有关 ADS 仿真的问题讨论 126
1.8.22 ADS 做 HB 优化仿真，有没有效果比较好的仿真速度优化方法 128
1.8.23 大功率增益曲线出现异常的讨论 128
1.8.24 HEMT 器件 g_m 与 C_{gs} 如何仿真得到 129
1.8.25 哪里可以找到 NXP 的器件模型 130
1.9 新架构、新器件 PA 设计 131
1.9.1 PA 的 Class 是怎么分的 131
1.9.2 电压合成型 PA 输入信号是都提，还是只提最下面晶体管 131
1.9.3 Cascode PA 中，共栅管漏端电容的作用 132
1.9.4 为什么 Cascode 放大器不做级间匹配 134
1.9.5 为什么 Balun 做差分电路功率后提升了 1.5dB 135
1.9.6 有关 Doherty PA 调试的讨论 138
1.9.7 Doherty PA 需要单独动放开启点对应的后级前端匹配吗 138
1.9.8 Doherty PA 的功率回退为什么设定成 6dB 139
1.9.9 1:1 Doherty 的匹配比 2，是怎么算出来的 139
1.9.10 Doherty PA 公式的推导 139
1.9.11 Doherty PA 偏置网络怎么设计 139
1.9.12 Doherty 中输功放电流应大有什么坏处 141
1.9.13 InGaP/GaAs 是一个东西吗 141
1.9.14 GaN 器件应用场景区分的依据是什么 142
1.9.15 GaN 功率管饱和时栅电流怎么理解 142
1.9.16 为什么 GaN 器件加偏压之后栅压被拉偏 144
1.9.17 GaN 器件关断状态下漏极电阻较小的讨论 144
1.9.18 PA 设计中，管子导通电阻 R_{on} 的大小选择 146
1.9.19 有关线性 PA 和 ET PA 的讨论 146
1.9.20 LDMOS 截面图中，N-well 是啥意思 147
1.9.21 PA 设计里为什么应用 NPN 做放大管子，而不用 PNP 147
1.9.22 如何提取功率管的寄生参数 147
1.9.23 功率管 RC 等效输的估算 148
1.9.24 有什么方法减少 PA 功率管流极到 GND 电感 149

目录

5.2.14 MLIN 是什么的缩写 220
5.2.15 为什么长度为 λ/8 的传输线可实现二次谐波短路点 221
5.2.16 怎么理解高阻线可等效成电感 221
5.2.17 微带线中拐角形状的选择 221
5.2.18 有关微带线的功率问题 223
5.2.19 微带线最大承受功率应该怎么计算 224
5.2.20 传输线传播常数如何设计 224
5.2.21 同轴线是如何实现相位差的 226
5.2.22 巴伦怎么做电容/电感等效 228
5.2.23 单转双巴伦仿真环境如何搭建 228
5.2.24 TFR 电阻误差来源是什么，在 wafer 上分布有规律吗 231
5.2.25 微带阵无边缘阻抗设计时，辅助微带线的长度如何选择 232
5.3 无源网络 233
5.3.1 一端开路的衰减器 VSWR 是多少 233
5.3.2 电路节点中等效电容如何计算 233
5.3.3 为什么 GaAs 工艺中 LC 串联和并联谐振频率不一样 234
5.3.4 为什么在电路中需要用 Bypass 电容创建低阻通路 234
5.3.5 差派电容上串个电阻是什么原理 235
5.3.6 无源电容会影响 PA 线性度吗 236
5.3.7 功分器作合路器时要求输入的两路信号相位要一致么 236
5.3.8 差压器如何设计输出比较平衡的差分信号 237
5.3.9 哪本书有讲 Smith 圆图详解比较详细吗 237
5.3.10 如何计算无源网络传输效率 237
5.4 无源电路实现 239
5.4.1 1mm 金环大约有多少电感量 239
5.4.2 如何计算共面波导中的间距 242
5.4.3 电感下方有 P_{PLUS} 或者 N_{PLUS} 的话，会有问题吗 243
5.4.4 为什么 PCB 50Ω 传输线实测不准 243
5.4.5 为什么实际生产中 Bypass 电容会使用同半径 0Ω 接地 244
5.4.6 金属屏蔽层接地应该由什么决定的 245
5.4.7 射频微波电路调试 Layout 用法 246
5.4.8 板级单独滤波器的位置是为了确定射频性能吗 246
5.4.9 Ku 波段电路板材厚度的选择 248
5.4.10 有关高频电流与电压分布的讨论 249
5.5 无源仿真技术 251
5.5.1 电容 Q 值是如何仿真的 251
5.5.2 电容的自谐振频率怎么仿真 251
5.5.3 ADS 可以做金线建模吗 252
5.5.4 如何仿真电感自谐振频率 252
5.5.5 为什么用 Y 参数和 Z 参数计算寄生电感值有差异 252
5.5.6 ADS 里计算两个电感互感公式是啥 254
5.5.7 管子漏源电容如何仿真 255
5.5.8 如何通过 SP 仿真得到传输线特征阻抗 255
5.5.9 为什么不同 Term 仿真时得到的 S 参数不同 257
5.5.10 ADS 中仿真出来的 S11 是匹配的 S11，还是输入端反射系数 258

9

目录

6.3.5 屏蔽腔体和 PA 布局一般有什么要求 289
6.3.6 地线周围边框的用途 289
6.3.7 放大器与邻近器之间高度的设计 289
6.3.8 Wi-Fi FEM PDET 引脚的作用 291
6.3.9 Wi-Fi FEM 数据手册中 PDET 口输出阻抗的含义 291
6.3.10 ASDiv 天线分集开关的功能是什么 292
6.3.11 LNA 和功率检波之间需要匹配吗 293
6.3.12 有关 LNA 与 Mixer 之间匹配设计的讨论 293
6.3.13 双平衡混频器如何匹配 294
6.3.14 从 VCO Pulling 角度，片内 PA 电感和 VCO 电感距离要求多少 294
6.3.15 RF Crossovers 这种器件都什么用的 295
6.3.16 Transceiver ADC 端口是否上下行反馈都能使用 295
6.3.17 什么是差分对的谐振“强制模拟” 295
6.3.18 有关电路 ESD 问题的讨论 297
6.4 系统问题排查 298
6.4.1 音内 De-sense 如何解决 298
6.4.2 LCD 干扰 TIS 的问题 debug 300
6.4.3 为什么 PA 加滤波器后 EVM 出现恶化 300
6.4.4 手机项目 SAR 指标不合格的潜在讨论 301
6.4.5 为什么系统会从端 RF 端引起大瓷砖 301
6.4.6 有人遇到过 NFC 影响 GPS 定位的问题吗 302
6.4.7 射频前端模组中移相器的作用 303
6.4.8 为什么本振与收发机频率不同时会出现杂散 303
6.4.9 PA 与 Transceiver 时序对频谱产生的影响 306
6.4.10 接收机系统问题的耗异常的 debug 讨论 306
6.4.11 Block 电源波形异常的原因 307
6.5 系统仿真技术 309
6.5.1 如何仿真 802.11ac 的 EVM 309
6.5.2 PA 仿真中，如何仿真 256QAM 调制并且生成 EVM 频谱原图 309
6.5.3 为什么 ADS 仿真混频器模板里，计算功率要 +30 313
6.5.4 EM 仿真得到的 S 参数，可以用于瞬态仿真吗 313
6.5.5 .wv 格式波形文件能在 ADS 读取吗 316
6.6 整机技术及系统知识运用 317
6.6.1 为什么 P_{1dB} 时，$IMD = P_{out} - IM3 = 19.26dBc$ 317
6.6.2 已知信号峰均比，如何算 PA 选型 317
6.6.3 带有 PAPR 信号时 TVS 箝定电压如何选择 318
6.6.4 谐波的频谱带宽和基频信号频谱带宽，有倍数关系吗 318
6.6.5 有关两个单音信号叠加的理解讨论 320
6.6.6 信号是否可以比噪低许低 324
6.6.7 为什么灵敏度可以比热噪声小 325
6.6.8 交调和互调有什么区别 325
6.6.9 无源混频器的增益就是它的变频损耗吗 326
6.6.10 DPD 为啥改善 AM-PM 和 AM-AM 效应 327
6.6.11 品质因数频率准确度指标的理解 328
6.6.12 相控阵可以提高天线辐射效率吗 328

11

目录

5.5.11 如何在 ADS 中提取 Transformer 参数 259
5.5.12 EM 仿真时 FEM 和 mom RF 仿真的差异原因 259
5.5.13 频率设置成 adaptive 之后不收敛 259
5.5.14 ADS RFpro 预留器件应选什么模型 260
5.5.15 EM 时如何加 Port 才更准确 261
5.5.16 GaAs 工艺无源电感用模型库还是 EM 仿真库 261
5.5.17 Layout 整体仿真和多个局部联合仿真的区别 262
5.5.18 有关整体 EM 仿真有问题的 debug 262
5.5.19 ADS EM 仿真中，整体仿真和分块仿真结果不一致 263
5.5.20 HFSS 仿真中使用波端口时是否需要勾选归一化选项 263
5.5.21 怎么用 HFSS 仿真验证电感的品质因数 264
5.5.22 HFSS 仿真 Wirebond 电感如何去除对地电容影响 264
5.5.23 怎么将 ADS 中的二维版图导入到 HFSS 里面 265
5.5.24 ADS 仿真中没连接不上还报点怎么解决 265
5.5.25 如何仿真大信号下的互调和干扰 265

第六章 系统设计技术 267
6.1 协议需求和系统指标 268
6.1.1 5G-A 和 6G 分别是什么 268
6.1.2 LTE QPSK 的峰均比是多少 268
6.1.3 NR 信号不同调制下的峰均比大约是多少 269
6.1.4 系统动平回退的标准是什么 271
6.1.5 23dBm Full RB 和 24dBm Partial RB 哪种场景 PA 非线性严重 272
6.1.6 LTE 中的 Cat 指的是什么 272
6.1.7 Wi-Fi FEM TDD 切换时间上下限是多少 273
6.1.8 Wi-Fi FEM Vbat 上电启动时间要求是多少 273
6.1.9 为什么 Wi-Fi 图重于 DEVM，LTE 侧重于 ACPR 274
6.1.10 华硕网都需要什么芯片 275
6.1.11 卫星通信用的是什么射频方案 275
6.1.12 MIPI 协议里定义 LNA 的 trigger mode 是什么意思 277
6.1.13 MIPI RFFE 中寄存器如何和 Trigger 模式关联 278
6.1.14 衡量接收机带内 block 信号的影响用的什么指标 278
6.1.15 DPD 方案对 ACLR 有没有要求 279
6.1.16 相位噪声与 EVM 的关系 279
6.1.17 EVM 单位噪比有关系吗 279
6.2 射频前端方案 282
6.2.1 什么是 Phase 系列 282
6.2.2 射频前端未来是分立还是 PAMiD 282
6.2.3 Phase8 里有什么新特性吗 283
6.2.4 哪里可以学习手机类系统知识 284
6.3 系统互联技术 286
6.3.1 PA/FEM 里多个 V_{cc} 脚的作用是什么 286
6.3.2 放大器偏置网络中，电阻与电感的位置可以交换吗 286
6.3.3 Tx 大功率下，输出开关管截的影响 288
6.3.4 用 MCU 控制 PA，PA 的线性信号质量有啥要求 288

10

目录

6.6.13 设计射频 MMIC 有推荐资料 329
6.6.14 北京为什么闻名香港同半导体 329

第七章 系统测试技术 330
7.1 测试组策 331
7.1.1 不同 SMA 头类型的名称 331
7.1.2 EVB SMA 接头处走线处理 332
7.1.3 有关 EVB 板 SMA 头长波线问题 332
7.1.4 SMA 座头需要打螺丝吗 333
7.1.5 双面 SMA 头使用的讨论 334
7.1.6 0201 器件组装有什么快速方法 334
7.1.7 PCB 上芯片整合之后，还能进回流焊吗 334
7.1.8 晶体管器件底部不平如何组装 335
7.1.9 手工贴装是否需要用钢网 335
7.1.10 焊空洞有什么影响，如何分析 335
7.2 测试方案设计 337
7.2.1 芯片的可靠性评估方法 337
7.2.2 开关 C_{off} 电容的测试方法 337
7.2.3 何测试 PA 负载失衡的性能 337
7.2.4 衰减器的输出 1dB 压缩点功率是测试积分功率还是峰值功率 338
7.2.5 如何测试 Tx 泄漏到另外信道 Rx 的信号强度 338
7.2.6 差分 LNA 如何测试 338
7.2.7 如何测试大信号 blocker 下接收机的噪声性能 339
7.2.8 如何测试耦合器的 IP3 339
7.3 测试知识与经验 340
7.3.1 网分测试如何去除 PCB 的影响 340
7.3.2 频评估板上的校准通道的作用 340
7.3.3 噪声测试中，为什么去嵌只去除一半线损 341
7.3.4 有关毫米波 LNA 探针是 bonding 线等生如何考虑 341
7.3.5 为什么相矢网和信号源测试的大功率芯片增益不一致 342
7.3.6 PA Ruggedness 反射门选择 VSWR 检测界面 342
7.3.7 P_{1dB} 是相对小信号增率点还是大功率增益扩张起来的点 343
7.3.8 功放呆用古空比测试时功率会减小吗 343
7.3.9 为什么 GaN 毫米波功放实测比仿真低 2~3dB 343
7.3.10 频带带内倾斜的原因是什么 344
7.3.11 EVM 测试为什么要解调 344
7.3.12 为什么单音和调制信号用检波管检测出来的 AD 值不一样 345
7.3.13 发机双音测试时，信号交调现象的讨论 345
7.3.14 峰值功率计的视频带宽如何理解 346
7.3.15 相对功率测试指标如何理解 347
7.3.16 Signal Studio 里 NR 不同 Config 的区别是什么 347
7.4 测试问题排查 348
7.4.1 为何测试中无法看到 EVM 信息点 348
7.4.2 关 GPS LNA 测试 S21 异常的 debug 348
7.4.3 两路 PA 测试时，连接方法不同，测试电流不同的原因 351

12

6.4 《无线重构世界》正式全面开售

作者：小慧，发布时间：2024 年 8 月

经过一周多的紧张筹备，众多读者已经陆续收到了慧智微的新书《无线重构世界》。在出版社及各平台的共同努力下，此书现已正式全面开售。



在此，我们衷心感谢每一位在预售期间给予支持与帮助的读者。得益于大家的热情，该书在预售阶段便取得了令人瞩目的成绩：

- **在电子通信类新书中荣登榜首**
- **在整个科技大类中位列第二**

这无疑是对本书的巨大鼓励。

全面开售后，购书平台将确保图书的全面发货，力求让读者在最短的时间内收到本书。另外，优惠购书活动仍在继续，图书原价 108 元，当前 59 元包邮，部分城市可当天到货。再次感谢大家的鼎力支持与厚爱，真诚地希望《无线重构世界》能够得到您的喜爱，并期待您的宝贵反馈！





《无线重构世界》实拍图：


专家推荐
RECOMMENDATION
慧智微电子


目录
CONTENTS
第一篇 射频简史
人类沟通的渴望
射频的诞生
高速发出，射频的有限发展
高速加速，射频的蓬勃发展
第二篇 现代射频
人多好办事，射频产生碰撞协同
无线标示实方案：独立的标准化
现代射频系统：认识“三大件”
射频实现方案：示意图决策的搭建
第三篇 射频芯片
沙子是怎么变成芯片的
射频芯片中的半导体
收发机芯片：射频中的“集野猪”
射频 PA：强得好好聊一聊
射频 LNA：低调又强大
003
004
006
010
017
026
027
033
041
052
071
076
116
公众号 · 慧智微电子


无线重构世界
射频技术的过去、现在与未来
邬贺铨
传序力荐
公众号 · 慧智微电子





在线阅读

本文摘自慧智微微信公众号，扫描以下二维码，在线阅读

6.5 射频人专属鼠标垫更新

作者：小慧，发布时间：2024 年 10 月

自 2022 年慧智微首次推出射频人专属鼠标垫以来，它便凭借其实用性受到了广大射频工程师的欢迎，已经成为不少射频人桌上的必备工具。慧智微原版射频人鼠标垫包含：

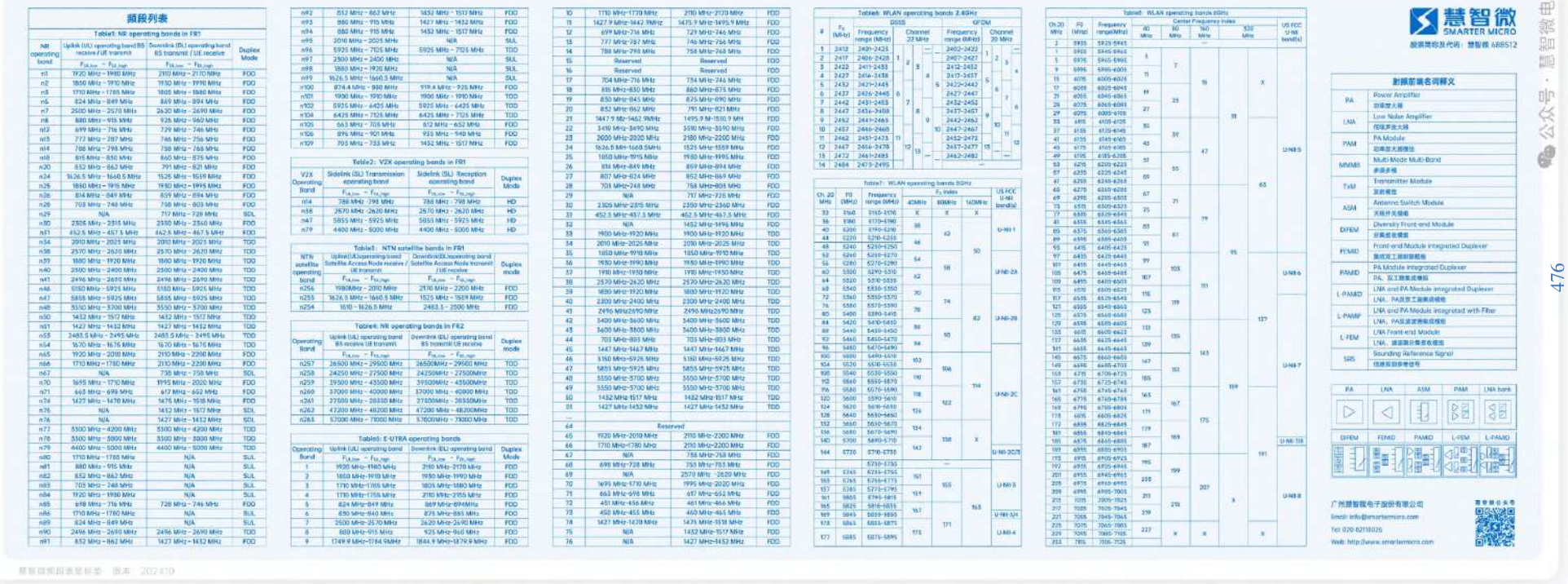
- 4G/5G 射频常用频段列表
- 4G 射频前端基础名词解释
- 4G 慧智微微信公众号二维码

为了更好地满足射频工程师的需求，我们根据大家的反馈对鼠标垫进行了升级。最新版本新增及更新了以下内容：

- **Wi-Fi 频段列表（含最新 Wi-Fi7 频段）**
- **3GPP 最新定义的 5G 频段列表**
- **5G NTN/V2X 最新频段信息**



如果您对这款升级版射频人专属鼠标垫感兴趣，欢迎联系慧智微当地销售团队获取。慧智微始终致力于成为您身边专业的射频合作伙伴。


频段列表
慧智微
SMARTER MICRO
射频前端名词释义
广州慧智微电子股份有限公司
Email: info@smartermicro.com
Tel: 020-62118026
Web: http://www.smartermicro.com

在线阅读

本文摘自慧智微微信公众号，扫描以下二维码，在线阅读



6.6 慧智微射频工具箱小程序上线

作者：小慧

随着科技的不断发展，射频技术在无线通信、雷达、卫星通信等领域得到了广泛应用，也成为了现代通讯、电子、航空等领域中不可或缺的一部分。作为射频人，我们日常工作中会需要进行各种公式计算，但很多人是在 Excel 里面计算，这样的方法需要自己手动收集整理公式，在使用时不够便利。

于是我们开发了一款专为射频人使用的计算工具，包含了常见的射频计算公式和实用工具，现慧智微射频工具箱小程序正式上线。

期待这款工具箱能成为大家的的得力助手，给大家的工作提供便利。

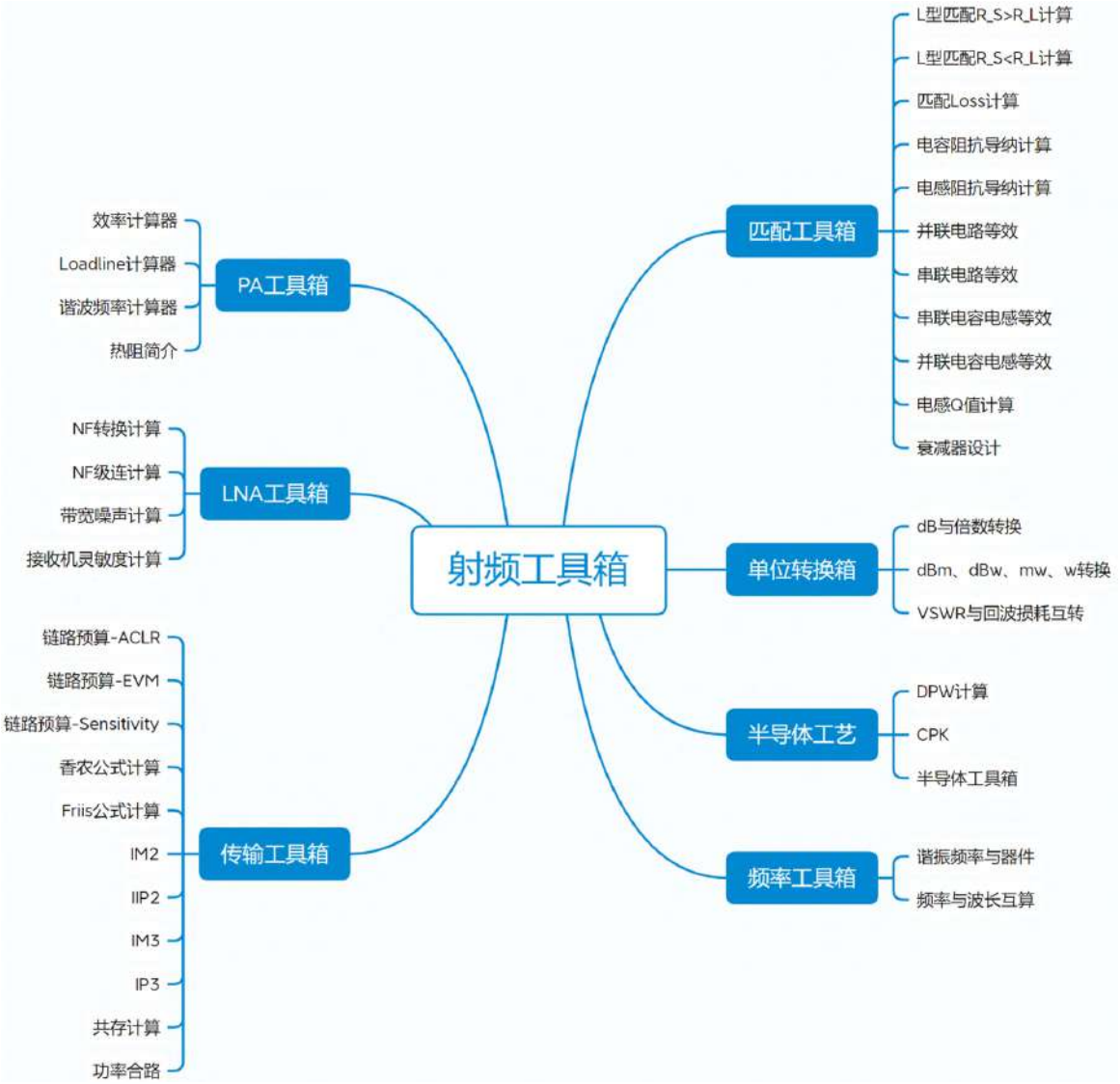


图 6.58: 射频工具箱内容



图 6.59: 慧智微射频工具箱计算页面



图 6.60: 慧智微射频工具箱收藏页面

图 6.61: 扫码添加小程序