人工智能赋能的混合MIMO波束成形

Nir Shlezinger，*IEEE*会员，Mengyuan Ma，*IEEE*学生会员，Ortal Lavi，*IEEE*学生会员，Nhan Thanh Nguyen，*IEEE*会员，Yonina C. Eldar，*IEEE*会士，以及 Markku Juntti，*IEEE*会士

arXiv:2303.01723v1 [cs.IT] 3 Mar 2023

摘要—混合多输入多输出（MIMO）是一项极具吸引力的技术，能够以可扩展且节能的方式实现未来无线通信所设想的超大规模MIMO系统。然而，混合MIMO系统将部分波束成形置于模拟域、部分置于数字域，这使其波束模式的优化相比传统全数字MIMO更具挑战性。因此，近年来人们对利用数据辅助人工智能（AI）工具进行混合波束成形设计的兴趣日益浓厚。本文综述了利用数据改进实时混合波束成形设计的候选策略。我们讨论架构约束，并概述混合波束成形优化所面临的核心挑战。随后，我们介绍如何通过传统优化应对这些挑战，并梳理不同的人工智能辅助设计方法。这些方法大致可分为纯数据驱动的深度学习模型，以及将人工智能与经典优化相结合的不同形式的深度展开技术。我们对现有方法进行系统的比较研究，其中包括数值评估和定性指标。最后，我们介绍将人工智能融入混合MIMO系统所带来的未来研究机遇。

I. 引言

大规模MIMO系统以及毫米波和亚太赫兹频段的高频通信，预计将在未来的第六代（6G）网络中发挥关键作用 [1]。这些技术天然互为补充：采用大型收发天线阵列的大规模多输入多输出技术能够生成高度聚焦的波束，这对于高频可靠通信至关重要；而短波长信号则使得在有限孔径内布置包含大量单元的MIMO配置成为可能。然而，实现此类大规模MIMO收发器会带来若干挑战。其中一项核心挑战在于高频工作的射频链成本高昂且功耗显著；在传统的全数字MIMO阵列中，每个天线单元都分别连接到数字信号处理单元。

混合波束成形 被认为是应对上述挑战的领先解决方案，可在射频链数量有限的情况下实现高频大规模MIMO通信 [2]。该方案将部分信号处理任务交由模拟域完成，从而将波束成形任务划分为数字和模拟两部分。模拟域中可实现的波束图取决于电路，典型实现方式包括基于移相器 [3],、矢量调制器 [4],以及动态超表面天线（动态超表面天线） [5]的方案。因此，与全数字收发机相比，混合收发机的波束成形能力受到固有约束。

尽管混合设计缓解了大规模MIMO系统的部分成本和功耗问题，但其受限的结构也带来了算法和信号处理方面的挑战。最值得注意的是，波束成形任务，即将信道状态信息（信道状态信息）转换为合适的波束图，通常需要求解一个非凸约束优化问题。人们提出了各种迭代优化算法，用于调整混合波束成形器 [6], ，这些算法所考虑的硬件约束和目标各不相同。这些迭代解决方案的一个主要局限在于其通常收敛缓慢，因为波束图设置必须实时完成，以应对信道变化。

深度学习作为人工智能的一项赋能技术，其兴起促使研究人员提出了人工智能赋能的混合波束成形设计。深度学习通常基于数据设定推理规则，但也可以训练深度神经网络（DNN）来解决具有挑战性的优化问题 [7]。训练完成后，深度神经网络可进行推理，其延迟固定，由层数决定，因此能够快速将信道状态信息映射为波束图 [8]。另一种利用数据进行混合波束成形的方法是采用基于模型的深度学习方法 [9]。在这种方法中，深度学习技术用于增强迭代式混合波束成形优化器，而非取代它们，同时利用数据实现快速收敛 [10]–[12]。混合MIMO波束成形方法日益多样，因而有必要对这些方法进行统一综述。

本文系统介绍了用于混合MIMO波束成形的人工智能辅助方法。成功实现混合MIMO收发器，必然需要将硬件开发与信号处理算法方面的考量相结合；本文聚焦于后者，不局限于特定实现方式。我们首先探讨混合MIMO系统，回顾具有代表性的架构，并说明其运行方式如何影响可实现的波束图。我们指出了混合波束成形带来的设计挑战，并确定了推动人工智能融入其中的关键因素。

接下来，我们将介绍混合波束成形设计方法，并将其划分为三大类：基于优化的方法，该类方法采用迭代优化器来设定波束图；基于深度神经网络的方案，通过预训练深度神经网络将信道状态信息映射为混合配置；以及深度展开设计，利用深度学习技术促进迭代优化。对于后者，我们将介绍不同类型的展开方法，并讨论每种方法如何形成不同的设计。基于这一划分，我们开展比较研究，包括数值研究和定性比较，并从多个关键性能指标方面分析这些方法之间的相互作用，

最后，我们将讨论有待未来探索的研究挑战，这些挑战有望为充分发挥人工智能在混合MIMO系统中的潜力铺平道路。

II. 混合波束成形

A. 混合*MIMO*收发器

大规模MIMO收发器配备了由大量单元组成的天线阵列，记为 M。在当前的5G基站（BS）中， M 可达数十个。随着6G从大规模MIMO向全息MIMO演进，天线数量预计可能增至数千个 [13]。在传统的全数字MIMO架构中，每根天线接收的信号都会由数字处理单元通过专用射频链分别进行数字处理。

在混合MIMO系统中，射频链的数量记为K，少于天线数量。这是通过在射频链与天线之间进行接口连接的模拟处理实现的，如图1所示。模拟处理器能够对信号进行不同的操作。与全数字架构相比，混合MIMO的一项显著优势在于其使用的射频链少于天线单元；在高频下使用大规模阵列时，这一点尤为关键。除了减少射频链数量外，混合设计还有助于抑制干扰，并减轻低分辨率模数转换器引起的失真 [4]。

B. 架构约束

混合MIMO系统结合了数字和模拟信号处理。在数字域中执行的处理具有很高的灵活性，可以有效地对不同频谱分量应用不同的映射。然而，模拟处理受到诸多限制，其能够实现的映射集合由硬件决定，文献中已提出多种不同的硬件架构。为说明不同设计所涉及的约束，我们简要回顾几种具有代表性的模拟架构，并重点介绍它们在发射时的工作方式：

1）移相器网络：最常见的模拟硬件采用具有可控相位的移相器 [2]。这些架构通常分为全连接架构和部分连接结构：在全连接架构中，每个射频链与每根天线之间都连接有一个专用移相器；而在部分连接结构中，每个射频链通过一个专用移相器连接至单根天线。在实际应用中，每个移相器施加的相位通常无法任意设置，而必须符合预定义的相位分辨率。此外，移相器通常设计为在所考虑的频带内保持（近似）相同的相移。因此，通常将其建模为对每个频谱分量应用相同的映射。

2）离散矢量调制器：移相器只影响信号的相位，而矢量调制器是能够实现不同相位移和信号衰减组合的模拟电路。与移相器相比，这类模拟电路具有更大的灵活性，因为它还能够以可控方式影响信号的幅度。不过，低功耗矢量调制器通常受到限制，只能实现预定义的有限数量的不同相位衰减组合 [4]。

3) 动态超表面天线：一种用于实现全息MIMO的新兴技术是将超表面用作天线。超表面是由可控超材料单元构成的平面配置。与上述依赖专用模拟电路的架构不同，动态超表面天线将可配置的模拟处理作为其天线结构的固有副产物来实现 [5]。在发射时，每个射频链输出的信号沿波导传播，并由连接到该波导的各单元辐射；每个单元都能实现一种频率选择性洛伦兹滤波器。因此，动态超表面天线能够固有地实现频率选择性模拟信号处理，但其形式受限于洛伦兹函数。

C. 混合波束成形设计挑战

混合波束成形设计 旨在联合设置模拟处理和数字处理，以针对当前信道实现优化预先定义的通信指标。典型指标包括可达速率，或多用户通信中的最小信干噪比（SINR）。对于采用常见线性波束成形的下行传输，任务归结为设计应用于待发送符号的预编码器（数字处理可对每个频谱分量分别进行预编码），并配置模拟处理。

混合波束成形设计面临多项核心挑战，包括：

C1 用于确定数字预编码器和模拟配置的优化问题通常是非凸的。即使设计目标采用二次形式，例如线性高斯信道的可达速率，将处理划分为数字和模拟两部分，以及模拟处理所受的硬件约束，通常也会导致非凸优化。

C2 由于混合波束成形是针对特定信道实现设计的，因此每当信道条件发生变化时，都需要重新执行设计，即每个相干时间都要执行一次（根据3GPP第17版，相干时间可能短至 125 μSec）。由于无线通信信道的相干时间通常会随着载波频率升高而缩短，因此必须快速完成设计过程，才能在每个相干时间内实现可靠通信。

C3 混合波束成形设计使用信道状态信息，而信道状态信息通常通过导频信令获取，因此可能受到噪声影响。因此，混合波束成形设计应能够应对可用信道状态信息中的一定程度误差。

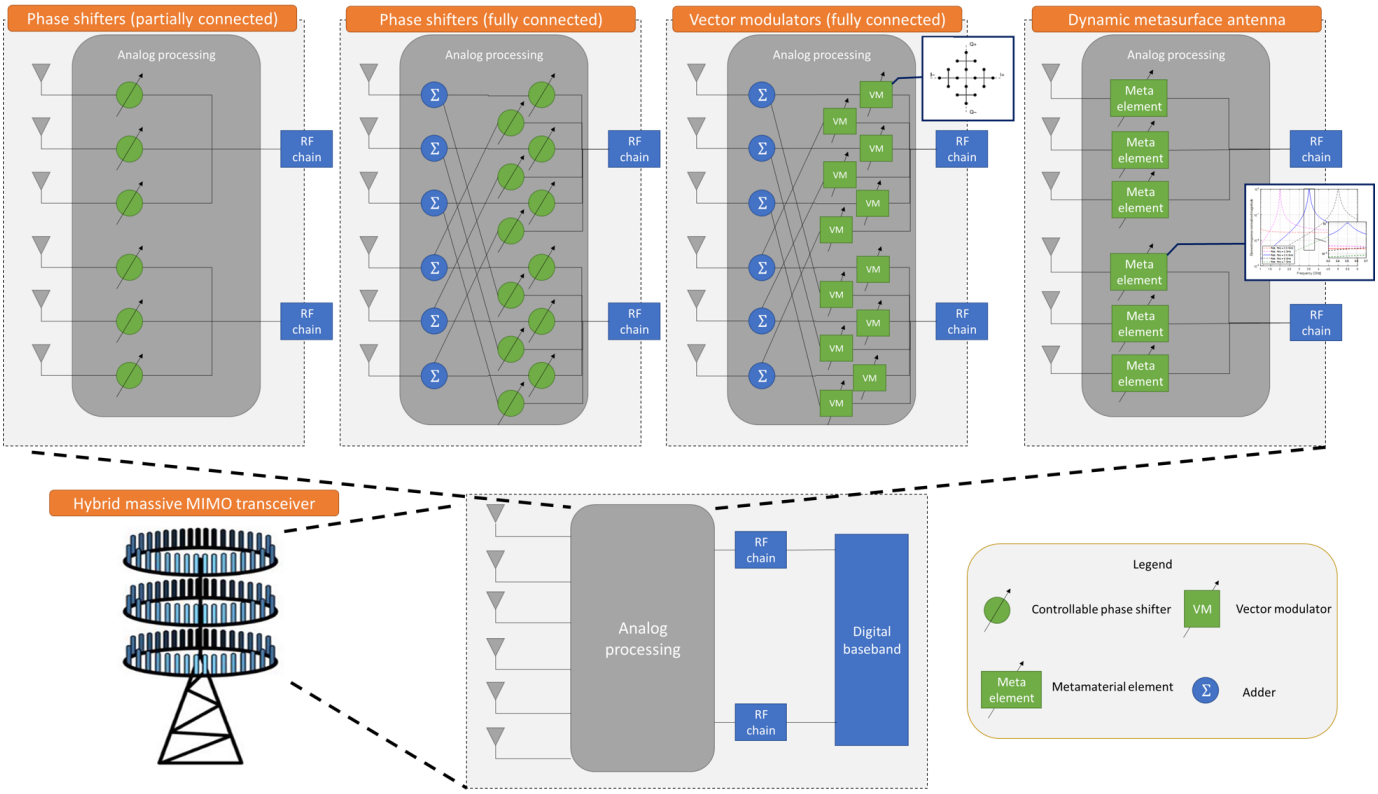


图 1：不同混合MIMO收发器架构及其相应模拟处理模型的示意图，包括部分连接和全连接移相器网络、矢量调制器和动态超表面天线。

上述挑战，尤其是 *C1* 和 *C2*，促使人们采用人工智能辅助设计，下一节将对此进行讨论。

III. 人工智能辅助混合波束成形设计

接下来，我们将详细介绍用于设计混合预编码器的主要框架。第一种框架采用针对具体问题的迭代优化器。第二种框架使用深度神经网络，即从数据中调优的抽象架构，用于将信道状态信息映射为混合波束成形器配置。最后一种框架采用深度展开，通过不同形式的基于模型的深度学习，将迭代优化与深度学习相结合 [9]。如图 2 所示，后一种方法在前两种技术之间取得平衡，兼具问题针对性和数据驱动的学习能力。

A. 基于优化的混合波束成形

如前所述，混合波束成形设计本质上是一个优化问题。因此，传统上通常借助优化工具来解决，常见做法是使用迭代求解器。概括而言，应对非凸性（*C1*）主要有两种方法：

- 一种主流方法采用凸松弛，即构造一个凸的替代问题。最常见的做法是，在混合配置中，用寻找最接近全数字速率最大化预编码器的混合配置，来替代非凸和速率目标函数 [3], [14]。与直接最大化和速率相比，松弛后的问题通常更容易求解，通常采用基于交替优化的迭代方法。然而，根据硬件约束的不同，所得问题仍可能是非凸的。可以证明，在某些情况下，所得解会接近速率最大化配置，尤其是在射频链数量 K 不小于接收天线数量 [3]时。
- 另一种方法是直接处理非凸目标，通常通过寻找合适的预编码器初始设置，并利用局部凸优化技术（例如投影梯度上升（PGA） [10]）对其进行改进。

虽然迭代优化器通常能够求得有效的混合波束成形器，但往往需要大量迭代才能收敛。由于迭代次数会转化为延迟和复杂度，这一特性限制了它们在时变场景中的适用性 *C2*。优化理论提供了通过回溯法等方法减少迭代次数的技术，但这类技术在推理过程中需要额外且耗时的计算。

B. 基于深度神经网络的混合波束成形

深度学习提供了用于调整以深度神经网络参数化的机器学习模型的工具，使其能够从数据中学习理想的复杂映射。深度神经网络还可以通过训练来解决具有挑战性的优化问题，例如混合波束成形设计中遇到的问题 [7]。卷积神经网络（CNN）等架构已被证明能够

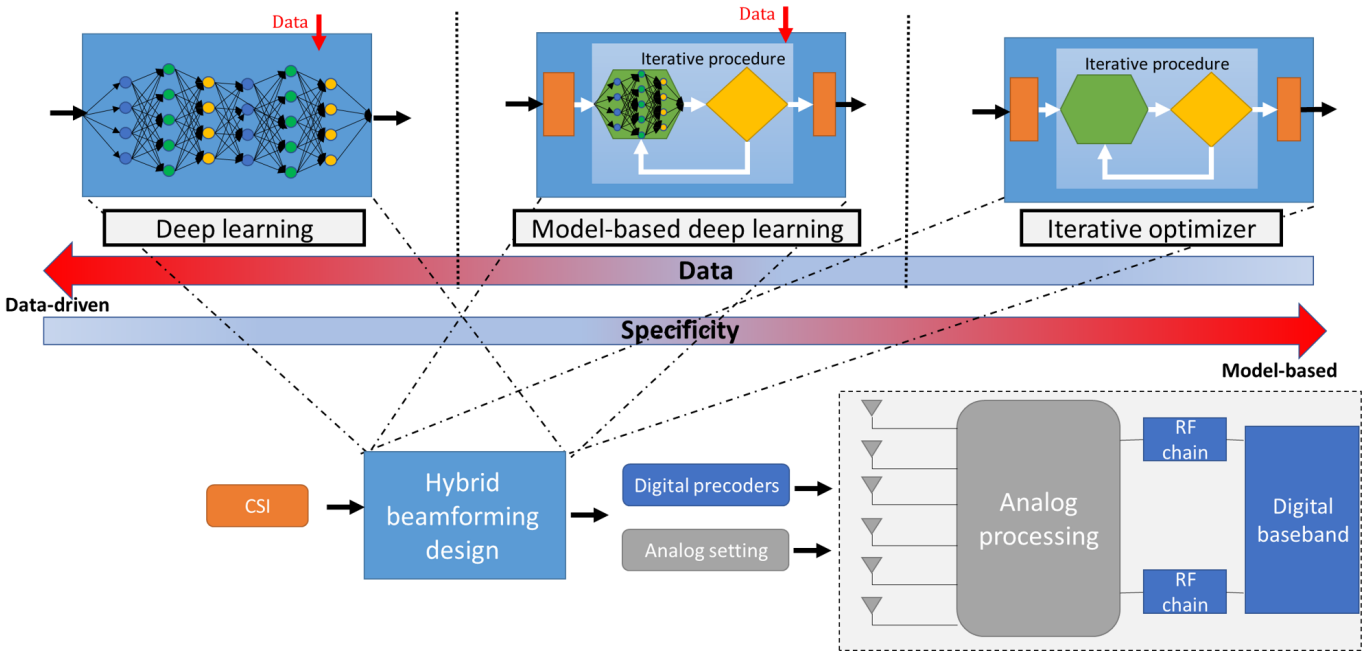


图 2：混合波束成形设计的不同方法示意图。

学习将 MIMO 信道状态信息映射为模拟和数字预编码器 [8]。

基于DNN的推理规则通常采用监督方式设计，即提供由输入及其期望输出组成的数据，让模型在训练过程中学习生成这些输出。然而，对于混合波束成形，通常可以采用无监督方式训练，也就是提供仅由信道实现组成的数据集，而不指定每个信道对应的期望波束成形器。这是可行的，因为可以针对每个选定的预编码器评估优化目标（例如和速率或信干噪比），且该目标关于预编码器可微。因此，可以使用（负的）优化目标作为无监督训练损失 [10]，应用传统的基于梯度的学习来训练基于DNN的混合波束成形器。

深度神经网络通常计算复杂度较高，因为其包含大量参数，且训练过程可能耗时较长。然而，其推理延迟由层数决定，是固定的，并且各种软硬件工具都能促进其并行化。因此，与迭代优化器相比，使用预训练的深度神经网络进行混合波束成形设计通常速度更快。然而，使用基于数据训练的通用高参数量模型来替代优化求解器，会带来若干缺点。首先，深度神经网络的训练通常耗时较长，需要大量数据（即信道实现）和繁琐的实验，才能学习到合适的映射。此外，虽然其推理延迟是固定的，但与迭代优化器相比，应用深度神经网络的复杂度通常较高，例如浮点运算量较大，这是由参数数量决定的。此外，深度神经网络的灵活性远不如优化方法；任务中的每项修改，例如在网络中增加一个用户，都需要耗时地重新训练。最后，深度神经网络的可解释性很差，通常只能为其输入和输出赋予操作层面的含义，因此往往被视为黑箱。

C. 深度展开混合波束成形

从数学原理出发的迭代优化器与数据驱动的深度学习系统，在混合波束成形场景中各有局限，因此需要结合两种方法各自优势的混合设计。基于模型的深度学习 [9] 涵盖了一系列策略，将基于严谨数学模型的推理与深度学习技术相结合；其中深度展开方法尤其适用于通常采用迭代方法解决的任务。

深度展开基于如下相似性：迭代优化器经过 L 次迭代的顺序操作，与深度神经网络经过 L 层的前向路径相似。其基本原理是将迭代视为参数化机器学习模型的归纳偏置，从而将经过 L 次迭代、因此具有固定延迟的优化器，有效转化为可训练的判别模型。这由此产生了三种不同形式的深度展开优化器 [9]：

U1 学习型超参数：迭代优化器具有超参数，即求解器的参数，例如步长。在允许优化器运行至收敛时，这些超参数的具体设置通常对优化结果影响不大，而且往往需要手动调优。然而，当优化器被限制在固定（且较少）的迭代次数内时，超参数设置可能会产生重大影响。因此，深度展开可以将迭代求解器的超参数转换为可训练参数，从而利用数据在预先设定的迭代次数内自动调优每次迭代所用的超参数。

U2 学习目标：迭代优化器基于目标函数（例如和速率）设计，通常会在每次迭代中修改优化变量，以进一步提升目标值。深度展开可以对每次迭代所用的目标函数进行参数化。这样便能从数据中学习，使每次迭代根据不同的目标调整其优化变量，从而使经过 L 次迭代后的输出在真实目标意义下达到最优。

U3 DNN 转换： 深度展开的第三种形式是设计一个深度神经网络来模仿迭代优化器的运作。这通常通过将每次迭代中的部分运算替换为可训练层来实现。这种展开支持不同程度的抽象。可以保留迭代优化器的运作，只将特定计算替换为可训练神经元；也可以设计一个参数量很大的深度神经网络，其架构以源自的优化器的运作为灵感。

对于混合波束成形，深度展开优化器具备深度神经网络以无监督方式进行训练的能力。此外，展开优化器的架构与迭代优化器相似，这也带来了有助于训练的其他因素。首先，迭代优化器可以为可训练架构提供有原则的初始化，确保训练从一个有效的运算开始；直观而言，这种运算只会随着训练推进而进一步改进。此外，由于每次可训练迭代的输出都可以与优化变量关联起来，训练计算损失时不仅可以依据经过 L 次迭代/层后的输出（如传统深度神经网络），还可以将中间特征纳入考量。这类训练损失不适用于黑盒架构，它能促使可训练模型在每次迭代/层都生成有效的设置，因此构成了一种已知有助于学习的正则化。

上述方法，尤其是*U1* （例如， [10],[11]）和*U3* （例如， [12]），支持针对混合波束成形设计的数据辅助固定延迟迭代优化。具体而言，采用学习型超参数的深度展开完整保留了迭代优化器的运作方式，从而维持其灵活性和可解释性。尽管如此，通过为每次迭代 [10], ，甚至为每个优化后的预编码器值 [11],学习不同的步长，仍可显著降低延迟。此外，学习型超参数还可纳入采用不同目标的优化器中，包括用于应对信道状态信息不确定性的鲁棒优化*C3*，如 [10]所示。

IV. 比较研究

本节将比较前文详述的混合波束成形设计方法。为此，我们将通过数值研究比较各类设计方法中的代表性方案，随后进行定性比较。

A. 数值评估

为了比较所考虑的混合波束成形设计方法，我们模拟采用全连接

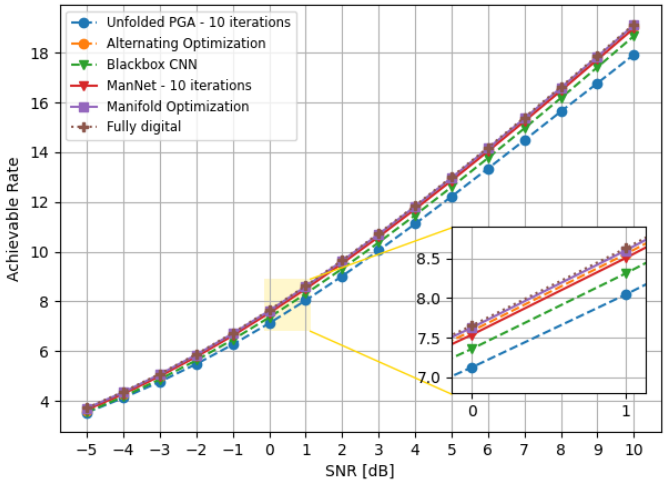


图 3：4 条射频链时的和速率与信噪比关系。

移相器网络进行模拟处理的混合MIMO系统。我们比较以下确定预编码器的方法：

- 对于基于优化的方法，我们评估了 [3]的黎曼流形优化器，以及 [14]，的交替优化器；二者均基于和速率目标函数的凸松弛。
- 对于基于深度神经网络的设计，我们采用了一个遵循 [8],架构的卷积神经网络，称为黑盒卷积神经网络。该架构由三个卷积层（使用 3×3 核）和三个全连接层组成。训练该卷积神经网络，使其既能同时生成模拟预编码器和数字预编码器，也能仅生成模拟预编码器，同时相应地调整数字预编码器，使其尽可能匹配全数字波束成形器。由于这两种实现得到的结果相似，此处仅报告后一种。
- 对于展开优化器，我们考虑 [11]，的 ManNet 模型，该模型对凸松弛优化进行展开；同时也考虑 [10], 的展开式 PGA，它将简单的 PGA 步骤应用于非凸和速率目标函数。虽然这两种展开方法仅使用 10 次迭代，但仍保留了其所源自的迭代优化器的*U1*运行方式。
- 为表示可达和速率的上限，我们评估了采用全数字波束成形所能达到的和速率。

所考虑的 MIMO发射机具有 $M = 12$ 根天线，并通过 16 个频率子带进行信号传输，为 4 个单天线用户提供服务。我们使用 QuaDRiGa 模型生成了 1000 个中心频率为 30 GHz 的毫米波信道。

我们首先将射频链的数量设置为 $K = 4$，即与用户数量相同。图 3所示的和速率与信噪比（SNR）的关系表明，所有基于凸松弛的优化器，即 [3], [14]的迭代优化器和人工智能辅助的ManNet [11],，其和速率均接近全数字波束成形。黑盒卷积神经网络和展开式PGA与全数字波束成形所达到的速率之间也仅存在较小差距。尽管如此，展开式设计相较于纯粹的

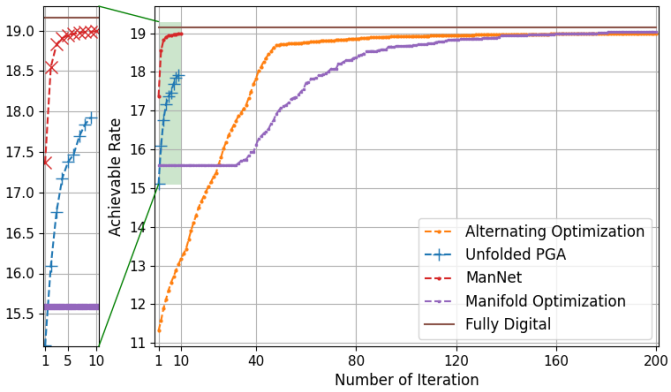


图 4：每次迭代的和速率。

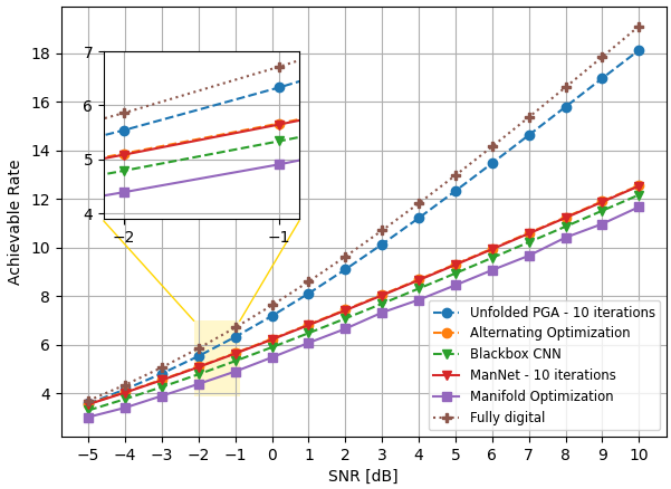


图 5：和速率与信噪比的关系，2 条射频链。

基于优化的方法的优势，则体现在达到这一性能所需的迭代次数上。图 4报告了信噪比为 10 dB 时各迭代方法的和速率与迭代次数的关系。由此可以观察到，与传统迭代优化器相比，展开方法利用数据以更少的迭代次数获得合适的设置，这表明人工智能辅助设计能够显著降低延迟和计算复杂度。

人工智能辅助设计相较于基于模型的优化器的另一项性能优势，是能够从数据中学习以应对非凸性。为此，我们重复图3 中的研究，同时将射频链数量设为 $K = 2$，即少于用户数量，这表明混合波束成形面临着具有挑战性的工作条件。图5所示结果表明，在这种情况下，直接处理非凸和速率目标函数并利用数据学习如何优化的展开式 PGA，与全数字上界之间的差距仍然很小。这里，基于凸松弛的优化器表现明显较差，因为它们旨在逼近全数字波束成形器，而在这种设置下这是无法实现的。

B. 定性比较

前文详述的混合波束成形器优化方法各有特点，分别适用于不同类型的场景。上述数值研究使我们能够从可达速率的角度比较这些方法，即解决*C1*问题。为了进一步探讨其他有意义的比较方面，下面我们将讨论五项关键指标——设计时延、计算复杂度、数据需求、灵活性和可解释性。下文的比较总结见表I。

1）延迟：混合波束成形的一项核心挑战，是需要在每个相干时*C2*间更新波束图。传统迭代优化器通常耗时较长，因需多次迭代而产生显著延迟。深度展开可以缓解这一问题，尤其是通过超参数学习*U1*，如图4所示。采用深度神经网络进行混合波束成形设计通常延迟较低，因为计算具有若干层的神经网络的前向传播所需时间是固定的；通过并行化和硬件加速器还可进一步缩短这一时间，不过其延迟未必能达到无线信道的相干时间量级。

2）复杂度：虽然深度神经网络通常能够实现快速且延迟固定的混合波束成形设计，但其计算复杂度较高，因为包含大量参数；而其延迟较低通常得益于并行化和硬件加速。迭代优化器的复杂度要低得多，因为每次迭代通常只涉及少量运算，但由于其按顺序运行，这种低复杂度并未转化为低延迟。深度展开设计，尤其是采用学习型超参数的设计*U1*，兼具迭代优化器的低复杂度，并且由于迭代次数固定，能够实现快速推理。

3）数据：人工智能辅助混合波束成形设计利用数据学习如何将信道状态信息映射为混合预编码器。虽然可以采用无监督方式进行此类学习，但为此类任务训练深度神经网络仍需要大量数据，即来自与部署时预期分布相同的信道实现。深度展开通过对学习模型施加归纳偏置，在数据依赖程度上实现了平衡，以特定性换取参数化 [9]，，其中，抽象参数化（*U3*）相比参数较少的模型（*U1*）需要更多数据。

4) 灵活性：混合波束成形设计需要具备一定的灵活性，因为信道配置（例如用户数量）可能会随时间变化。迭代优化器具有极高的灵活性，同一个优化器可应用于不同的设置。同样，完全保留迭代优化器（*U1*）运算的展开方法也具备这种灵活性。然而，深度神经网络是针对固定配置训练的，因此灵活性很低，因为配置发生变化时必须重新训练。

5) 可解释性：混合波束成形设计的一个重要属性，是能够理解其如何将信道状态信息映射为混合预编码器，并跟踪其处理链。迭代优化器具有完全的可解释性，不改变其运行方式的展开优化器（*U1*）也是如此。偏离优化器的更抽象展开形式（*U3*）的可解释性较低，但仍然可以跟踪其过程，因为每次迭代仍与一个操作

方法	延迟	复杂度	Data	灵活性	可解释性
迭代式 优化器 优化器	高——迭代次数多 迭代	低——运算次 数少但迭代次数多 迭代	无 - 无需数据 需要	完全灵活- 适用于不同的 不同 配置 项	完全 可解释
DNNs	中等——由前向 传播固定 次 DNN	High - 复杂 high 参数化 模型	高 - 大规模 所需数据集 用于训练	无需重新训练—— 需要进行 切换配置—— tion	无法解释
深度展开 优化器*U1*	最低- 少量预先 定义的迭代 低复杂度	最低- 少量操作 少量迭代 迭代	低—少量参数 使用小型数据集训练 小型数据集	灵活 - 适用 with 不同的 配置 不过 性能 可能会受到影响	完全 可解释 - 保持操作 as 迭代 优化器
深度展开 优化器 *U3*	Low - few 预定义 迭代 with 中等 复杂度	中 - 复杂 参数化 少量 迭代	中等 - 相对- 较大的数- 参数数量 用于训练	无需重新训练—— ing 通常 is 需要切换 配置	部分 可解释 因为可以追踪 中间 特征

TA表 I：所考虑的混合波束成形方法之间的定性比较 .

含义相关。对于黑盒深度神经网络，只有输入和输出具有可解释的含义。

V. 总结与未来研究方向

人工智能辅助设计和基于模型的深度学习有望显著促进实时高吞吐量混合波束成形，从而为实现可持续、可扩展的大规模MIMO部署铺平道路。然而，要充分发挥人工智能辅助波束成形的潜力，仍需探索多个研究方向。下面我们回顾一些候选主题。

A. 集成感知混合多输入多输出

6G网络预计将使用多输入多输出收发器，不仅用于通信，也用于感知。这种运行方式会带来波束成形设计方面的诸多考量，涉及感知与通信频谱共享设备之间的共存，以及双功能信令等。这些考量显著增加了混合波束成形的复杂性，因为优化过程必须兼顾与感知功能相关的其他方面。这进一步推动了对集成感知混合多输入多输出人工智能辅助技术的探索。

B. 面向功耗与硬件的设计

尽管大多数关于混合MIMO的研究都考虑基于移相器的模拟电路，但实际上混合架构有多种形式，每种形式都会带来不同的约束，进而影响波束成形设计。此外，现有的混合波束成形方法往往忽视了这样一个事实：模拟电路的不同配置会消耗不同的功率。例如，在混合设计中关闭部分矢量调制器的能力已被证明可以显著降低功耗 [4]。这促使人们探索将功耗和硬件因素纳入优化过程的混合波束成形算法，而相关的高复杂度也使得采用所倡导的人工智能辅助策略成为必要。此外，典型的混合波束成形设计假设天线阵列是理想线性的，而实际的功率放大器则具有不同程度的非线性。线性化技术广为人知，但整体阵列响应设计需要相当复杂的解决方案。

C. 分布式混合*MIMO*

未来无线通信预计将偏离传统蜂窝架构，采用多重连接和无蜂窝拓扑 [1]。这一运作方式通过部署多个协作式多输入多输出发射机，将传统的集中式波束成形扩展为分布式波束成形。混合架构成本较低，适合大规模部署；同时，协作运行能够克服其受限波束图带来的局限。在此类场景中使用人工智能，可以显著促进实时协作式混合波束成形，并有望利用联邦学习和多智能体强化学习等分布式机器学习范式。

D. 从远场到近场

影响未来无线通信中波束成形的另一项考虑因素是，预计将从远场通信过渡到近场。具体而言，预计 MIMO 收发器孔径的增大以及

高频的使用意味着，通信很可能将在辐射近场中进行，而非传统无线收发器设计中所假设的常规远场。

在辐射近场中运行催生了波束成形的新形式，尤其是生成聚焦波束的能力，这些波束能够显著减轻干扰。初步研究表明，采用不同形式的混合波束成形，也可以通过耗时的优化实现聚焦波束 [15]。未来的研究将进一步探索同时支持远场和近场用户的能力，以及人工智能辅助混合波束成形在实现实时、精确地为近场通信生成聚焦波束图方面的能力。此外，原则上，近场的球面波前能够提高定位及其他感知应用的精度，而深度展开优化也是实现这一目标的一种潜在工具。

参考文献

[1] M. Giordani、M. Polese、M. Mezzavilla、S. Rangan 和 M. Zorzi，“迈向6G网络：应用场景与技术”，IEEE Commun. Mag.，第58卷，第3期，第55–61页，2020年。[2] A. F. Molisch、V. V. Ratnam、S. Han、Z. Li、S. L. H. Nguyen、L. Li 和 K. Haneda，“大规模MIMO的混合波束成形：综述”，IEEE Commun. Mag.，第55卷，第9期，第134–141页，2017年。[3] X. Yu、J.-C. Shen、J. Zhang 和 K. B. Letaief，“毫米波MIMO系统混合预编码的交替最小化算法”，IEEE J. Sel. Topics Signal Process.，第10卷，第3期，第485–500页，2016年。[4] T. Zirtiloglu、N. Shlezinger、Y. C. Eldar 和 R. T. Yazicigil，“采用低分辨率模数转换器和任务专用波束成形的高能效混合MIMO接收机”，载于IEEE ICASSP会议论文集，2022年，第5338–5342页。[5] N. Shlezinger、G. C. Alexandropoulos、M. F. Imani、Y. C. Eldar 和 D. R. Smith，“用于6G极端大规模MIMO通信的动态超表面天线”，IEEE Wireless Commun.，第28卷，第2期，第106–113页，2021年。[6] X. Qiao、Y. Zhang、M. Zhou 和 L. Yang，“基于交替优化的毫米波MIMO系统混合预编码策略”，IEEE Access，第8卷，第113 078–113 089页，2020年。[7] A. Zappone、M. Di Renzo 和 M. Debbah，“深度学习时代的无线网络设计：基于模型、基于人工智能，还是两者兼有？”，IEEE Trans. Commun.，第67卷，第10期，第7331–7376页，2019年。[8] A. M. Elbir 和 A. K. Papazafeiropoulos，“多用户毫米波大规模多输入多输出系统的混合预编码：一种深度学习方法”，IEEE Trans. Veh. Technol.，第69卷，第1期，第552–563页，2019年。[9] N. Shlezinger、Y. C. Eldar 和 S. P. Boyd，“基于模型的深度学习：深度学习与优化的交汇”，IEEE Access，第10卷，第115 384–115 398页，2022年。[10] O. Agiv 和 N. Shlezinger，“学习快速优化混合预编码”，载于IEEE SPAWC会议论文集，2022年。[11] N. T. Nguyen、M. Ma、N. Shlezinger、Y. C. Eldar、A. L. Swindlehurst 和 M. Juntti，“面向太赫兹大规模多输入多输出系统的深度展开混合波束成形设计”，arXiv预印本 arXiv:2302.12041，2023年。

[12] [ ] 2302 12041 在线资源。网址：https://arxiv.org/abs/。E. Balevi 和 J. G. Andrews，“采用生成对抗网络压缩超低反馈开销的展开式混合波束成形”，IEEE Trans. Wireless Commun.，第20卷，第12期，第8381–8392页，2021年。[13] C. Huang、S. Hu、G. C. Alexandropoulos、A. Zappone、C. Yuen、R. Zhang、M. Di Renzo 和 M. Debbah，“面向6G无线网络的全息多输入多输出表面：机遇、挑战与趋势”，IEEE Wireless Commun.，第27卷，第5期，第118–125页，2020年。[14] F. Sohrabi 和 W. Yu，“面向大规模天线阵列的数字与模拟波束成形设计”，IEEE J. Sel. Topics Signal Process.，第10卷，第3期，2016年。[15] H. Zhang、N. Shlezinger、F. Guidi、D. Dardari、M. F. Imani 和 Y. C. Eldar，“近场多用户多输入多输出通信的波束聚焦”，IEEE Trans. Wireless Commun.，第21卷，第9期，第7476–7490页，2022年。

Nir Shlezinger 是以色列本-古里安大学电气与计算机工程学院的助理教授。

Mengyuan Ma (mengyuan.ma@oulu.fi) 是芬兰奥卢大学无线通信中心的博士生。

Ortal Lavi (agivo@post.bgu.ac.il) 是以色列本-古里安大学电气与计算机工程学院的研究生。

Nhan Thanh Nguyen （nhan.nguyen@oulu.fi）获首尔科技大学博士学位。目前就职于芬兰奥卢大学。

Yonina C. Eldar （yonina.eldar@weizmann.ac.il）是以色列魏茨曼科学研究所数学与计算机科学系教授，负责领导该校生物医学工程与信号处理中心。她是以色列科学院与人文学院成员、IEEE会士和EURASIP会士。

Markku Juntti （markku.juntti@oulu.fi）获芬兰奥卢大学理学博士学位，自2000年以来一直在该校任教授，并担任CWC无线电技术研究组主任。他还是莱斯大学电气与计算机工程系的兼职教授。