

毫米波 MIMO 系统中基于自适应梯度算法的混合预编码

张煜^{1,2}, 张治¹, 董晓岱²

(1. 北京邮电大学信息与通信工程学院, 北京 100876; 2. 维多利亚大学电子与计算机工程学院, 维多利亚 V8W 3P6)

摘 要: 为了降低毫米波 MIMO 系统中现有交替最小混合预编码算法和基于动量梯度下降的在线学习混合预编码方案的复杂度, 针对单用户通信场景, 重新考虑混合预编码器的设计问题, 提出一种等效的单隐藏层神经网络架构。在该架构下, 数字预编码矩阵和模拟预编码矩阵的每一个组成元素可等效为单隐藏层神经网络的连接权值, 其最优解可通过神经网络中的权值训练方法获得。在此基础上, 结合反向传播算法, 提出一种基于自适应梯度反向传播的混合预编码机制。进一步地, 将所提算法扩展到多用户通信场景。仿真结果表明, 在单用户场景和多用户场景下, 所提算法可实现的频谱效率均接近全数字预编码, 同时复杂度低于现有的基于交替最小的混合预编码算法和基于动量梯度下降的在线学习混合预编码方案。

关键词: 毫米波; 混合预编码; 神经网络; 自适应梯度

中图分类号: TN92

文献标识码: A

DOI: 10.11959/j.issn.1000-436x.2021207

Adaptive gradient algorithm for hybrid precoding in mmWave MIMO system

ZHANG Yu^{1,2}, ZHANG Zhi¹, DONG Xiaodai²

1. School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China

2. Department of Electrical and Computer Engineering, University of Victoria, Victoria V8W 3P6, Canada

Abstract: To reduce the complexity of existing hybrid precoding algorithms based on alternating minimization (AltMin) and online learning via gradient descent with momentum in mmWave MIMO systems, aiming at the single-user scenario, the problem of designing the hybrid precoder was reconsidered and an equivalent single hidden layer neural network was proposed. Under the new architecture, the elements of the digital and analog precoder were equivalent to the connecting weights of a single hidden layer neural network, and their optimal solution could be obtained via the weights training method. Inspired by the back propagation (BP) algorithm in feed forward neural networks, an adaptive gradient (AG)-based BP algorithm for hybrid precoding was proposed. Furthermore, the proposed algorithm was extended to the multi-user scenario. The numerical results show that the proposed algorithm achieves approximately the same spectral efficiency as the fully-digital precoding in both the single-user and multi-user scenarios, while has lower complexity than the existing AltMin-based hybrid precoding algorithms and online learning hybrid precoding based on gradient descent with momentum.

Keywords: mmWave, hybrid precoding, neural network, adaptive gradient

1 引言

作为 5G 及未来移动通信网络的主要技术之

—^[1-3], 毫米波通信能够提供较宽的带宽, 满足高速数据传输需求。然而, 高频率、短波长的电磁波具有严重的路径损耗和较差的绕射性能^[4-5], 导致通

收稿日期: 2021-06-20; 修回日期: 2021-09-16

通信作者: 张治, zhangzhi@bupt.edu.cn

基金项目: 国家重点研发计划基金资助项目 (No.2019YFC1511302); 国家自然科学基金资助项目 (No.61629101)

Foundation Items: The National Key Research and Development Program of China (No.2019YFC1511302), The National Natural Science Foundation of China (No.61629101)

信范围有限。幸运的是,多输入多输出(MIMO, multi-input multi-output)技术可以利用大规模天线阵列来补偿严重的衰减,延长信号的传输距离。

传统的全数字 MIMO 架构通常是基于数字预编码器实现的,其中所需射频(RF, radio frequency)链的数量等于天线的数量。射频链通常由混频器、功率放大器和数字-模拟转换器等器件构成。对于毫米波系统,由于过高的功耗和硬件成本^[6],部署全数字预编码具有挑战性。借助毫米波信道的稀疏特性,文献[7]提出了一种混合数字-模拟预编码架构,使用较少的射频链实现低维数字预编码器和高维模拟预编码器的连接,其中模拟预编码器采用模值恒定、相位可调的模拟移相器实现;同时把混合预编码等效为压缩感知中的稀疏信号恢复问题,并使用正交匹配追踪(OMP, orthogonal matching pursuit)算法求解,最终达到了降低功耗和硬件成本的目标。鉴于 OMP 算法的计算复杂度较高,有些文献提出了改进算法,包括并行索引选择-忽略矩阵求逆的同步正交匹配追踪(PIS-MIB-SOMP, parallel-index-selection matrix-inversion-bypass simultaneous OMP)算法^[8]、基于顺序迭代最小二乘的广义正交匹配追踪(ORLS-gOMP, order-recursive least squares-based generalized OMP)算法^[9]、基于正交性的匹配追踪(OBMP, orthogonality-based matching pursuit)算法^[10]等。另外,文献[11]根据天线阵列响应矩阵的相关性生成模拟预编码矩阵备选集,有效降低了 OMP 算法搜索空间的大小。针对码字搜索问题,文献[12]提出了基于傅里叶变换码本的低复杂度码本搜索方法,降低了搜索复杂度。总体而言,文献[7-11]所提算法都是基于码本的混合预编码方法,所实现的频谱效率还有很大的提升空间^[12]。

为了接近最优的频谱效率性能,学者们研究了非码本的混合预编码方案,将混合预编码视作恒模约束下的矩阵分解问题,基于交替最小(AltMin, alternating minimization)算法迭代设计数字预编码器和模拟预编码器。具体地,数字预编码器通过最小二乘准则计算,模拟预编码器的最优解则分别通过共轭梯度(CG, conjugate gradient)^[13]、内点法^[14]、Barzilai-Borwein 梯度^[15]、梯度投影(GP, gradient projection)^[16]等算法获得。尽管文献[13-16]提出的非码本混合预编码方案在频谱效率性能上接近全数字预编码,但实现复杂度均较高。虽然文献[13]也提出了一种通过求解正交 Procrustes 问题(OPP,

orthogonal Procrustes problem)获得混合预编码器的低复杂度算法,但需要假设数字预编码器满足比例酉性质。由于该假设仅在数据流的数目等于射频链的数目时有效^[17],因此当射频链的数目大于数据流的数目时,OPP 算法所实现的频谱效率在一定程度上会有所降低。

上述研究主要针对单用户场景。为了解决多用户场景下的混合预编码问题,文献[18]提出了两阶段混合预编码方案,即基站(BS, base station)首先和移动台(MS, mobile station)联合设计最优的模拟预编码器和组合器以使信道增益最大化,然后设计迫零(ZF, zero-forcing)数字预编码器以减少 MS 间的干扰。文献[19]提出了一种混合最小均方误差(MMSE, minimum mean-squared error)预编码方法,即每个 MS 首先独立配置模拟组合器,然后由 BS 为所有 MS 设计模拟预编码器和数字预编码器。文献[20]提出了一种混合块对角化(BD, block diagonalization)预编码方法,该方法首先设计模拟预编码器和组合器以最大化阵列增益,然后在等效基带信道上应用 BD 方法设计数字预编码器。然而,文献[18-20]所提方案获得的频谱效率与全数字预编码相比尚有很大的差距。另一方面,文献[13-16]中的 CG-AltMin、BB-AltMin、GP-AltMin 等非码本混合预编码方案同样也可应用于多用户场景,但复杂度较高的问题依然存在。

近年来,深度学习在混合预编码问题中的应用受到了广泛关注^[21-28]。从所采用网络模型的角度,基于深度学习的混合预编码方案分为基于实数神经网络的混合预编码^[21]、基于复数反向传播(BP, backpropagation)神经网络的混合预编码^[22]和基于复卷积神经网络的混合预编码^[23-28]三类,其中文献[21]将实数神经网络应用于混合预编码器的设计,文献[22]提出了基于复数 BP 神经网络的混合预编码方案,文献[23-28]分别提出了基于 ComcepNet、等效信道-卷积神经网络(CNN, convolutional neural network)、CovNet、CNN-MIMO、HPNet 和深度学习量化相位网络模型的混合预编码机制。

进一步地,还可从学习方式的角度,将上述基于实数神经网络的混合预编码和基于复卷积神经网络的混合预编码归纳为线下学习方案,将基于复数 BP 神经网络的混合预编码归纳为在线学习方案。线下学习方案分为线下训练和线上测试 2 个阶段,在线下训练阶段,随机生成大量信道训练样本,

以信道参数为网络输入，以每个信道条件下相应的最佳模拟预编码器或组合器为网络输出，执行训练过程获得网络配置参数；在线上测试阶段，将信道测试样本输入配置好的网络，获得该信道条件下相应的最佳模拟预编码器或组合器，然后，数字预编码器通过对等效信道进行奇异值分解或运用 ZF 准则获得。

在训练好网络模型之后，线下学习方案能够快速响应信道参数，获得相应的最佳模拟预编码器或组合器。然而，为了获得最佳的网络配置参数，线下训练阶段往往需要庞大的训练样本，以提高精确度。另外，训练好的网络模型往往针对具体的参数设置，在单用户场景下是具体的发射和接收天线数，在多用户场景下是基站的天线数和用户数。当天线数或用户数发生变化时，需要重新训练以保证网络模型的准确性，这会造成较大的训练开销和较长的训练时间，使算法效率降低^[29]。

与线下学习方案形成鲜明对比的是，在线学习方案所需训练样本较少，能够根据信道条件的变化实时地调整网络权值，最佳的数字预编码器和模拟预编码器从网络权值中提取。研究线下学习方案的文献相对较少，以文献[22]为主要代表，其针对多用户场景，提出了基于复数 BP 神经网络的混合预编码方案，以最优 ZF 数字预编码器和实际混合预编码器之间的最小均方误差为损失函数，采用动量梯度下降（GDM, gradient descent with momentum）算法对网络进行训练，获得的权值即混合预编码矩阵的组成元素。然而，文献[22]没有考虑模拟预编码器的恒模约束，通过训练得到的和模拟预编码器对应的权值没有归一化为统一的模值，这与实际的硬件设计不符。当考虑模拟预编码器的恒模约束时，文献[22]采用的算法收敛速度较慢。

受上述讨论的启发，为了降低现有的传统非码本混合预编码方案和在线学习方案的复杂度，本文针对毫米波大规模 MIMO 系统，分别提出适用于单用户场景和多用户场景的低复杂度在线学习混合预编码方案，所需训练样本较少，训练速度较快。本文所提混合预编码方案具体如下：在单用户场景下，提出一种新颖的等效神经网络架构，将数字预编码器和模拟预编码器的每个元素视为单隐藏层神经网络的参数，并将混合预编码器的求解视作该神经网络的参数训练过程；进一步地，设计适用于该等效神经网络的参数训练方法，提出一种基于自

适应梯度反向传播（AG-BP, adaptive gradient backpropagation）的混合预编码算法，以最小化损失函数为目标，通过迭代最终获得单用户场景下的最优混合预编码器；最后将所提算法扩展至多用户场景，获得多用户场景下的混合预编码方案。仿真结果表明，在单用户场景和多用户场景下，所提算法实现的频谱效率均接近全数字预编码，且复杂度低于现有的基于交替最小算法的混合预编码方案和基于 GDM 的在线学习混合预编码方案。

2 系统模型

本节首先给出单用户场景下的系统模型，多用户场景下的系统模型将在 3.5 节给出。考虑一个单用户毫米波 MIMO 系统，设该系统中发射机配置 N_t 根天线、 N_t^{RF} 个射频链，接收机配置 N_r 根天线、 N_r^{RF} 个射频链，它们之间通过 N_s 路数据流进行通信。

设发送信号 $\mathbf{s} = [s_1, s_2, \dots, s_{N_s}]^T$ 为 $N_s \times 1$ 维矢量，且在发射端依次通过 $N_t^{\text{RF}} \times N_s$ 维数字预编码器 \mathbf{F}_{BB} 和 $N_t \times N_t^{\text{RF}}$ 维模拟预编码器 \mathbf{F}_{RF} 进行处理。混合预编码器整体满足功率约束，即 $\|\mathbf{F}_{\text{RF}}\mathbf{F}_{\text{BB}}\|_{\text{F}}^2 = N_s$ ，其中， $\|\cdot\|_{\text{F}}$ 表示 Frobenius 范数。此外， \mathbf{F}_{RF} 的每个元素都满足恒模约束，即 $|\mathbf{F}_{\text{RF}}(m, n)| = 1/\sqrt{N_t}$ 。在接收端，经过 $N_r \times N_r^{\text{RF}}$ 维模拟组合器 \mathbf{W}_{RF} 和 $N_r^{\text{RF}} \times N_s$ 维数字组合器 \mathbf{W}_{BB} 处理后的接收信号可表示为

$$\mathbf{r} = \mathbf{W}_{\text{BB}}^H \mathbf{W}_{\text{RF}}^H \mathbf{H} \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}} \mathbf{s} + \mathbf{W}_{\text{BB}}^H \mathbf{W}_{\text{RF}}^H \mathbf{n} \quad (1)$$

其中， $(\cdot)^H$ 表示共轭转置， \mathbf{H} 为 $N_r \times N_t$ 维信道矩阵， $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \sigma_n^2 \mathbf{I}_{N_r})$ 表示 $N_r \times 1$ 维高斯噪声。 \mathbf{W}_{RF} 也满足恒模约束，即 $|\mathbf{W}_{\text{RF}}(m', n')| = 1/\sqrt{N_r}$ 。

毫米波传播过程中过高的路径损耗造成散射体有限，毫米波信道通常采用基于扩展 Saleh Valenzuela 模型的簇信道模型，信道矩阵 \mathbf{H} 被描述为 N_{cl} 个簇的组合，每个簇包含 N_{ray} 个射线^[7]。因此， \mathbf{H} 可表示为

$$\mathbf{H} = \sqrt{\frac{N_t N_r}{N_{\text{cl}} N_{\text{ray}}}} \sum_{i=1}^{N_{\text{cl}}} \sum_{\ell=1}^{N_{\text{ray}}} \alpha_{i\ell} \mathbf{a}_r(\phi_{i\ell}^r, \theta_{i\ell}^r) \mathbf{a}_t^H(\phi_{i\ell}^t, \theta_{i\ell}^t) \quad (2)$$

其中， $\alpha_{i\ell}$ 表示第 i 个簇中第 ℓ 个射线的复增益， $\phi_{i\ell}^t$ 和 $\theta_{i\ell}^t$ 分别表示发射机离开方向的水平角和俯仰角， $\phi_{i\ell}^r$ 和 $\theta_{i\ell}^r$ 分别表示接收机到达方向的水平角和俯仰

角, $\mathbf{a}_t^H(\phi_{ic}^t, \theta_{ic}^t)$ 和 $\mathbf{a}_r(\phi_{ic}^r, \theta_{ic}^r)$ 分别表示与发射机和接收机的水平角、俯仰角相关的阵列响应向量。考虑均匀平面阵 (UPA, uniform planar array), 阵列响应向量可表示为

$$\mathbf{a}_{\text{UPA}}(\phi, \theta) = \frac{1}{\sqrt{N_y N_z}} \begin{bmatrix} 1, \dots, e^{j\frac{2\pi}{\lambda}d(n_y \sin\phi \sin\theta + n_z \cos\theta)}, \dots, \\ e^{j\frac{2\pi}{\lambda}d((N_y-1)\sin\phi \sin\theta + (N_z-1)\cos\theta)} \end{bmatrix}^T \quad (3)$$

其中, λ 为载波波长, d 为相邻天线间距离, $n_y = 0, \dots, N_y - 1$ 和 $n_z = 0, \dots, N_z - 1$ 分别为天线单元在 y 轴和 z 轴的索引, $N = N_y N_z$ 为天线总数。

假设发送信号 \mathbf{s} 服从高斯分布并且满足 $\mathbb{E}\{\mathbf{s}\mathbf{s}^H\} = \rho/N_S \mathbf{I}_{N_S}$, 其中 ρ 为平均发送功率。则该毫米波通信系统的频谱效率为

$$R = \text{lb} \left| \mathbf{I}_{N_S} + \frac{\rho}{N_S} \mathbf{R}_n^{-1} \tilde{\mathbf{H}} \tilde{\mathbf{H}}^H \right| \quad (4)$$

其中, $\mathbf{R}_n = \sigma_n^2 \mathbf{W}_{\text{BB}}^H \mathbf{W}_{\text{RF}}^H \mathbf{W}_{\text{RF}} \mathbf{W}_{\text{BB}}$ 表示噪声协方差矩阵, $\tilde{\mathbf{H}} = \mathbf{W}_{\text{BB}}^H \mathbf{W}_{\text{RF}}^H \mathbf{H} \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}}$ 。

为使如上的系统频谱效率最大化, 可对 \mathbf{F}_{BB} 、 \mathbf{F}_{RF} 、 \mathbf{W}_{RF} 和 \mathbf{W}_{BB} 进行联合优化, 即

$$\begin{aligned} (\mathbf{F}_{\text{BB}}^{\text{opt}}, \mathbf{F}_{\text{RF}}^{\text{opt}}, \mathbf{W}_{\text{RF}}^{\text{opt}}, \mathbf{W}_{\text{BB}}^{\text{opt}}) &= \arg \max_{\mathbf{F}_{\text{BB}}, \mathbf{F}_{\text{RF}}, \mathbf{W}_{\text{RF}}, \mathbf{W}_{\text{BB}}} R \\ \text{s.t. } |\mathbf{F}_{\text{RF}}(m, n)| &= \frac{1}{\sqrt{N_t}} \\ |\mathbf{W}_{\text{RF}}(m', n')| &= \frac{1}{\sqrt{N_r}} \\ \|\mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}}\|_F^2 &= N_S \end{aligned} \quad (5)$$

然而, 鉴于 \mathbf{F}_{RF} 和 \mathbf{W}_{RF} 的恒模约束, 式(5)是不可解的。通过分开设计混合预编码器和组合器, 式(5)可以解耦为 2 个可解的子问题^[7]。最优的模拟预编码器 $\mathbf{F}_{\text{RF}}^{\text{opt}}$ 和数字预编码器 $\mathbf{F}_{\text{BB}}^{\text{opt}}$ 可以通过求解以下问题获得^[13]

$$\begin{aligned} (\mathbf{F}_{\text{RF}}^{\text{opt}}, \mathbf{F}_{\text{BB}}^{\text{opt}}) &= \arg \min_{\mathbf{F}_{\text{RF}}, \mathbf{F}_{\text{BB}}} \|\mathbf{F}_{\text{opt}} - \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}}\|_F^2 \\ \text{s.t. } |\mathbf{F}_{\text{RF}}(m, n)| &= \frac{1}{\sqrt{N_t}}, \|\mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}}\|_F^2 = N_S \end{aligned} \quad (6)$$

其中, \mathbf{F}_{opt} 表示 $N_t \times N_S$ 维无约束全数字预编码器, 由 \mathbf{H} 的 N_S 个最大奇异值对应的 N_S 个右奇异向量

组成。假设 \mathbf{H} 可通过相应的信道估计方法^[30]预先获取, 定义 \mathbf{H} 的奇异值分解为 $\mathbf{H} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$, 则 $\mathbf{F}_{\text{opt}} = \mathbf{V}(:, 1:N_S)$ 为 \mathbf{V} 的前 N_S 列元素。

3 算法设计

本节针对单用户场景, 首先提出一种描述混合预编码的等效神经网络架构, 分析其与传统神经网络架构的不同; 然后给出训练数据集的产生方法, 并推导出损失函数的表达式; 接着以最小化损失函数为目标, 提出基于 AG-BP 的混合预编码机制; 最后针对多用户场景, 将所提算法进行扩展, 获得多用户场景下的混合预编码方案。

3.1 等效神经网络架构

重新考虑混合预编码式(6), 假设 \mathbf{s} 经过混合预编码处理的信号为 $\mathbf{y} = [y_1, \dots, y_{N_t}]^T \in \mathbb{C}^{N_t \times 1}$, 那么混合预编码问题可以等效为确定 \mathbf{s} 与 \mathbf{y} 之间的映射关系, 即 $\mathcal{M}: \mathbf{s} \rightarrow \mathbf{y}$ 。

映射 \mathcal{M} 包含 2 个线性组合, 即 \mathbf{F}_{BB} 和 \mathbf{F}_{RF} 。因此, 混合预编码架构可视为一个单隐藏层神经网络, 如图 1 所示。

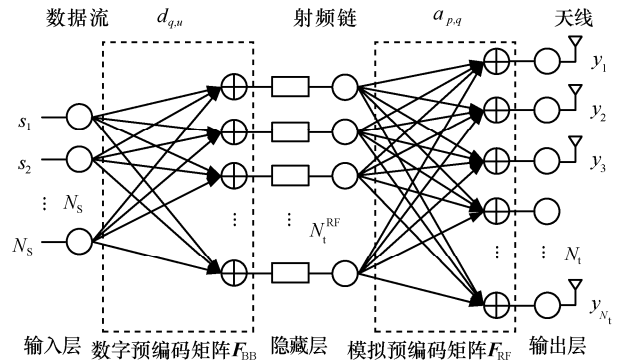


图 1 混合预编码的等效单隐藏层神经网络架构

N_S 路发送信号、 N_t^{RF} 个射频链、 N_t 根发射天线可分别类比成 N_S 个输入层神经元、 N_t^{RF} 个隐藏层神经元、 N_t 个输出层神经元, 则输出层信号可以重写为

$$\begin{bmatrix} y_1 \\ \vdots \\ y_{N_t} \end{bmatrix} = \begin{bmatrix} a_{1,1} & \cdots & a_{N_t^{\text{RF}},1} \\ \vdots & \ddots & \vdots \\ a_{1,N_t} & \cdots & a_{N_t^{\text{RF}},N_t} \end{bmatrix} \begin{bmatrix} d_{1,1} & \cdots & d_{N_S,1} \\ \vdots & \ddots & \vdots \\ d_{1,N_t^{\text{RF}}} & \cdots & d_{N_S,N_t^{\text{RF}}} \end{bmatrix} \begin{bmatrix} s_1 \\ \vdots \\ s_{N_S} \end{bmatrix} \quad (7)$$

图 1 所示网络的权值 $d_{q,u}$ 和 $a_{p,q}$ ($p=1, \dots, N_t$, $q=1, \dots, N_t^{\text{RF}}$, $u=1, \dots, N_s$) 与 F_{BB} 和 F_{RF} 的每个元素相对应, 可通过训练获得。因此, 可以利用神经网络的学习算法^[31]来设计混合预编码器。

神经网络的学习过程意味着调整相邻神经元之间的权值以及每个功能神经元的阈值以使网络最优。但本文所提出的混合预编码等效架构与传统的神经网络存在几点不同: 1) F_{BB} 和 F_{RF} 的元素均为复数; 2) 在混合预编码架构下不存在功能神经元; 3) 偏置全为 0 且激活函数为恒等函数; 4) 最重要的一点是隐藏层和输出层之间连接权存在恒模约束, 这使通用的神经网络训练算法不适用于混合预编码器的设计, 因此有必要设计混合预编码器专用的权值训练算法。

3.2 训练数据集和损失函数

对每一个发送信号实例 $\mathbf{s}^n = [s_1^n, \dots, s_{N_s}^n]^T$, 期待的经由无约束全数字预编码器处理的信号为 $\mathbf{y}^n = \mathbf{F}_{\text{opt}} \mathbf{s}^n$, 且 $\mathbf{y}^n = [y_1^n, \dots, y_{N_t}^n]^T \in \mathbb{C}^{N_t \times 1}$ 。一组 \mathbf{s}^n 和 \mathbf{y}^n 可视单隐藏层神经网络的一个输入输出样本, 目标是通过训练获得神经网络的最优权重。

由于 F_{RF} 存在恒模约束, F_{opt} 无法分解为 F_{RF} 和 F_{BB} 的乘积, 因此实际预编码后的信号表示为 $\hat{\mathbf{y}}^n = \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}} \mathbf{s}^n$, 且 $\hat{\mathbf{y}}^n = [\hat{y}_1^n, \dots, \hat{y}_{N_t}^n]^T \in \mathbb{C}^{N_t \times 1}$ 。优化式(6)可等效为设计合理的 F_{RF} 和 F_{BB} 以使理想输出和实际输出之间的误差 $E_n = \|\mathbf{y}^n - \hat{\mathbf{y}}^n\|_2^2$ 最小。即对一

组信号向量 $\{\mathbf{s}^1, \dots, \mathbf{s}^{N_d}\}$, 累积平均误差 $\bar{E} = \frac{1}{N_d} \sum_{n=1}^{N_d} E_n$ 应达到最小值。因此, 混合预编码问题可等效为

$$\begin{aligned} (\mathbf{F}_{\text{RF}}^{\text{opt}}, \mathbf{F}_{\text{BB}}^{\text{opt}}) &= \min_{\mathbf{F}_{\text{RF}}, \mathbf{F}_{\text{BB}}} \bar{E} \\ \text{s.t. } |a_{p,q}| &= \frac{1}{\sqrt{N_t}}, \|\mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}}\|_F^2 = N_s \end{aligned} \quad (8)$$

3.3 基于 AG-BP 的混合预编码

在复数神经网络的训练过程中, 通常将复数的实部和虚部分别进行更新, 然后合并^[31]。用字母 r 和 i 分别表示复数的实部和虚部, 则发送符号 s_u^n 、权值 $d_{q,u}$ 和 $a_{p,q}$ 可写成

$$\begin{cases} s_u^n = \text{sr}_u^n + \text{jsi}_u^n \\ d_{q,u} = \text{dr}_{q,u} + \text{jdi}_{q,u} \\ a_{p,q} = \text{ar}_{p,q} + \text{jai}_{p,q} \end{cases} \quad (9)$$

给定训练集 $\mathcal{D} = \{(\mathbf{s}^n, \mathbf{y}^n) | n=1, \dots, N_d\}$, 对于第 n 个发送信号 \mathbf{s}^n , 第 q 个隐藏层神经元的输入为

$$b_q^n = \sum_{u=1}^{N_s} d_{q,u} s_u^n = \sum_{u=1}^{N_s} (\text{dr}_{q,u} \text{sr}_u^n - \text{di}_{q,u} \text{si}_u^n) + \text{j} \sum_{u=1}^{N_s} (\text{dr}_{q,u} \text{si}_u^n + \text{di}_{q,u} \text{sr}_u^n) \quad (10)$$

第 p 个输出层神经元的输出为

$$\hat{y}_p^n = \sum_{q=1}^{N_t^{\text{RF}}} a_{p,q} b_q^n = \sum_{q=1}^{N_t^{\text{RF}}} (\text{ar}_{p,q} \text{br}_q^n - \text{ai}_{p,q} \text{bi}_q^n) + \text{j} \sum_{q=1}^{N_t^{\text{RF}}} (\text{ar}_{p,q} \text{bi}_q^n + \text{ai}_{p,q} \text{br}_q^n) \quad (11)$$

累积平均误差可重写为

$$\begin{aligned} \bar{E} &= \frac{1}{N_d} \sum_{n=1}^{N_d} \sum_{p=1}^{N_t} |y_p^n - \hat{y}_p^n|^2 = \\ &= \frac{1}{N_d} \sum_{n=1}^{N_d} \sum_{p=1}^{N_t} (y_p^n - \hat{y}_p^n)(y_p^{n*} - \hat{y}_p^{n*}) \end{aligned} \quad (12)$$

对于 $a_{p,q}$, 其增量 $\Delta a_{p,q}$ 可以写为实部增量 $\Delta \text{ar}_{p,q}$ 和虚部增量 $\Delta \text{ai}_{p,q}$ 相加的形式, 即

$$\Delta a_{p,q} = \Delta \text{ar}_{p,q} + \text{j} \Delta \text{ai}_{p,q} \quad (13)$$

根据梯度下降法, $\Delta \text{ar}_{p,q}$ 和 $\Delta \text{ai}_{p,q}$ 沿负梯度方向更新, 即

$$\Delta \text{ar}_{p,q} = -\eta \frac{\partial \bar{E}}{\partial \text{ar}_{p,q}}, \Delta \text{ai}_{p,q} = -\eta \frac{\partial \bar{E}}{\partial \text{ai}_{p,q}} \quad (14)$$

其中, η 表示学习率。

根据链式准则, 有

$$\begin{cases} \frac{\partial \bar{E}}{\partial \text{ar}_{p,q}} = \frac{\partial \bar{E}}{\partial \hat{y}_p^n} \frac{\partial \hat{y}_p^n}{\partial \text{ar}_{p,q}} + \frac{\partial \bar{E}}{\partial \hat{y}_p^{n*}} \frac{\partial \hat{y}_p^{n*}}{\partial \text{ar}_{p,q}} = \\ \quad - \frac{2}{N_d} \sum_{n=1}^{N_d} (\text{zr}_p^n \text{br}_q^n + \text{zi}_p^n \text{bi}_q^n) \\ \frac{\partial \bar{E}}{\partial \text{ai}_{p,q}} = \frac{\partial \bar{E}}{\partial \hat{y}_p^n} \frac{\partial \hat{y}_p^n}{\partial \text{ai}_{p,q}} + \frac{\partial \bar{E}}{\partial \hat{y}_p^{n*}} \frac{\partial \hat{y}_p^{n*}}{\partial \text{ai}_{p,q}} = \\ \quad \frac{2}{N_d} \sum_{n=1}^{N_d} (\text{zr}_p^n \text{bi}_q^n - \text{zi}_p^n \text{br}_q^n) \end{cases} \quad (15)$$

其中, zr_p^n 和 zi_p^n 分别表示 y_p^n 与 \hat{y}_p^n 差值的实部和虚部, 即 $y_p^n - \hat{y}_p^n = \text{zr}_p^n + \text{jzi}_p^n$ 。

将式(15)代入式(13)和式(14)可得

$$\Delta a_{p,q} = -\frac{2\eta}{N_d} \sum_{n=1}^{N_d} (zr_p^n br_q^n + zi_p^n bi_q^n - jzr_p^n bi_q^n + jzi_p^n br_q^n) \quad (16)$$

更新 $a_{p,q}$ 为

$$a_{p,q} \leftarrow a_{p,q} + \Delta a_{p,q} \quad (17)$$

为了加速收敛，避免损失函数到达局部极小值，权值增量可基于自适应梯度（学习率）计算^[32]，沿负梯度方向的搜索步长与之前所有迭代的梯度值有关，即

$$\Delta a_{p,q}(t+1) = -\frac{\eta}{\sqrt{\sum_{\tau=0}^t \left[\frac{\partial \bar{E}}{\partial ar_{p,q}}(\tau) \right]^2} + \delta} \frac{\partial \bar{E}}{\partial ar_{p,q}}(t) - j \frac{\eta}{\sqrt{\sum_{\tau=0}^t \left[\frac{\partial \bar{E}}{\partial ai_{p,q}}(\tau) \right]^2} + \delta} \frac{\partial \bar{E}}{\partial ai_{p,q}}(t) \quad (18)$$

其中， $\delta > 0$ 为一个足够小的数值以保证数值稳定性。

考虑到 $a_{p,q}$ 的恒模约束，重构 $a_{p,q}$ 为

$$a_{p,q} \leftarrow \frac{1}{\sqrt{N_t}} e^{j\angle a_{p,q}} \quad (19)$$

其中， $\angle a_{p,q}$ 表示复数 $a_{p,q}$ 的相位。

\mathbf{F}_{RF} 的每个元素与 $a_{p,q}$ 一一对应，即

$$\mathbf{F}_{\text{RF}}(p,q) = a_{p,q} \quad (20)$$

$d_{q,u}$ 也可采用与 $a_{p,q}$ 相似的更新规则。但为了降低计算复杂度， $d_{q,u}$ 可直接根据最小二乘解获得

$$\mathbf{F}_{\text{BB}} = (\mathbf{F}_{\text{RF}}^H \mathbf{F}_{\text{RF}})^{-1} \mathbf{F}_{\text{RF}}^H \mathbf{F}_{\text{opt}} \quad (21)$$

$$d_{q,u} = \mathbf{F}_{\text{BB}}(q,u)$$

对 $a_{p,q}$ 和 $d_{q,u}$ 迭代更新，直到满足以下条件

$$\|\mathbf{F}_{\text{opt}} - \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}}\|_{\text{F}}^2 < \varepsilon \quad (22)$$

其中， $\varepsilon > 0$ 为一个足够小的正值；或者当迭代次数超过预定义的最大值 t_{max} 时，迭代更新也将停止。

最后，更新 \mathbf{F}_{BB} 以满足功率约束

$$\mathbf{F}_{\text{BB}} = \sqrt{N_s} \frac{\mathbf{F}_{\text{BB}}}{\|\mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}}\|_{\text{F}}} \quad (23)$$

基于 AG-BP 的混合预编码算法如算法 1 所示。

算法 1 基于 AG-BP 的混合预编码算法

给定无约束全数字预编码器 \mathbf{F}_{opt} 、训练数据集大小 N_d 、学习率 η 和最大迭代次数 t_{max}

1) 随机产生 N_d 个发送信号 $\{\mathbf{s}^n\}_{n=1}^{N_d}$ ；

2) 计算输出层期望的输出 $\mathbf{y}^n = \mathbf{F}_{\text{opt}} \mathbf{s}^n$ ；

3) 以随机相位初始化模拟预编码器 $\mathbf{F}_{\text{RF}}^{(0)}$ ，根据最小二乘计算初始数字预编码器 $\mathbf{F}_{\text{BB}}^{(0)}$ ，定义 $t = 0$ ；

4) while $\|\mathbf{F}_{\text{opt}} - \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}}\|_{\text{F}}^2 \geq \varepsilon$ 且 $t < t_{\text{max}}$ do

5) 根据式(10)计算隐藏层输出 $\{b_q^n\}_{n=1}^{N_d}$ ；

6) 根据式(11)计算输出层输出 $\{\hat{y}_p^n\}_{n=1}^{N_d}$ ；

7) 根据 $y_p^n - \hat{y}_p^n = zr_p^n + jzi_p^n$ 获得 zr_p^n 和 zi_p^n ；

8) 根据式(17)更新 $a_{p,q}$ ，然后根据式(19)重构

$a_{p,q}$ ；

9) 根据式(20)获得 \mathbf{F}_{RF} ；

10) 根据式(21)获得 \mathbf{F}_{BB} 并更新 $d_{q,u}$ ；

11) $t = t + 1$ ；

12) end while；

13) 根据式(23)更新 \mathbf{F}_{BB} 。

接下来，分析算法 1 的复杂度。算法 1 中每一步操作的复数乘法和除法个数如表 1 所示，其中

$$f_{\text{inv}}(m) = \frac{m^5 - 3m^4 + 6m^3 - 3m^2 + 2m - 3}{3} \quad (24)$$

表示利用初等行变换对 $m \times m$ 矩阵进行求逆所需的复数乘法和除法个数。算法 1 和对比算法的计算开销如表 2 所示，其中， N_{iter}^i 和 N_{iter}^o 分别表示内部迭代的总数和外部迭代的个数。对于 CG-AltMin、BB-AltMin 和 GP-AltMin， N_{iter}^i 为所有外部迭代所需内部迭代个数的累加值。

表 1 算法 1 每步操作所需复数乘法和除法个数

操作	复数乘法和除法个数
计算 $\{y\}_{n=1}^{N_d}$	$N_d N_t N_s$
计算 $\{b_q^n\}_{n=1}^{N_d}$	$N_d N_t^{\text{RF}} N_s$
计算 $\{\hat{y}_p^n\}_{n=1}^{N_d}$	$N_d N_t N_t^{\text{RF}}$
更新 $a_{p,q}$	$(4N_d + 2) N_t N_t^{\text{RF}}$
重构 $a_{p,q}$	$2N_t N_t^{\text{RF}}$
更新 $d_{q,u}$	$2N_t (N_t^{\text{RF}})^2 + N_t N_t^{\text{RF}} N_s + f_{\text{inv}}(N_t^{\text{RF}})$

3.4 基于 AG-BP 的混合组合器设计

最优的模拟组合器 $\mathbf{W}_{\text{RF}}^{\text{opt}}$ 和数字组合器 $\mathbf{W}_{\text{BB}}^{\text{opt}}$ 可以通过求解以下问题获得^[13]

表 2 算法 1 与对比算法的计算开销

算法	复数乘法和除法个数
算法 1	$N_{\text{iter}} \left((5N_d + 4)N_t N_t^{\text{RF}} + N_d N_t N_s + N_d N_t^{\text{RF}} N_s + 2N_t (N_t^{\text{RF}})^2 + N_t N_t^{\text{RF}} N_s + f_{\text{inv}}(N_t^{\text{RF}}) \right)$
CG-AltMin	$N_{\text{iter}}^i \left(2(N_t)^2 N_t^{\text{RF}} N_s + 11N_t N_t^{\text{RF}} + 1 \right) + N_{\text{iter}}^o \left(2N_t (N_t^{\text{RF}})^2 + N_t N_t^{\text{RF}} N_s + f_{\text{inv}}(N_t^{\text{RF}}) \right)$
BB-AltMin	$N_{\text{iter}}^i \left(2(N_t)^2 N_t^{\text{RF}} N_s + 8N_t N_t^{\text{RF}} + 1 \right) + N_{\text{iter}}^o \left(2N_t (N_t^{\text{RF}})^2 + N_t N_t^{\text{RF}} N_s + f_{\text{inv}}(N_t^{\text{RF}}) \right)$
GP-AltMin	$N_{\text{iter}}^i \left(2(N_t)^2 N_t^{\text{RF}} N_s + 2N_t N_t^{\text{RF}} + 1 \right) + N_{\text{iter}}^o \left(2N_t (N_t^{\text{RF}})^2 + N_t N_t^{\text{RF}} N_s + f_{\text{inv}}(N_t^{\text{RF}}) \right)$
GDM	$N_{\text{iter}}^{\text{GDM}} \left((5N_d + 4)N_t N_t^{\text{RF}} + N_d N_t N_s + N_d N_t^{\text{RF}} N_s + 2N_t (N_t^{\text{RF}})^2 + N_t N_t^{\text{RF}} N_s + f_{\text{inv}}(N_t^{\text{RF}}) \right)$

$$\begin{aligned} (\mathbf{W}_{\text{RF}}^{\text{opt}}, \mathbf{W}_{\text{BB}}^{\text{opt}}) &= \arg \min_{\mathbf{W}_{\text{RF}}, \mathbf{W}_{\text{BB}}} \left\| \mathbf{W}_{\text{opt}} - \mathbf{W}_{\text{RF}} \mathbf{W}_{\text{BB}} \right\|_{\text{F}}^2 \\ \text{s.t. } \left| \mathbf{W}_{\text{RF}}(m', n') \right| &= \frac{1}{\sqrt{N_r}} \end{aligned} \quad (25)$$

其中, \mathbf{W}_{opt} 表示 $N_r \times N_s$ 维无约束全数字组合器, 由 \mathbf{H} 的 N_s 个最大奇异值对应的 N_s 个左奇异向量组成, 即 $\mathbf{W}_{\text{opt}} = \mathbf{U}(:, 1:N_s)$ 。

显而易见, 除了功率约束, 式(25)和式(6)有相同的形式。因此, 可通过把 \mathbf{F}_{opt} 、 \mathbf{F}_{RF} 和 \mathbf{F}_{BB} 替换成 \mathbf{W}_{opt} 、 \mathbf{W}_{RF} 和 \mathbf{W}_{BB} 并运用算法 1 对式(25)进行求解。

3.5 多用户场景

本节在上述工作的基础上, 将所提算法扩展至多用户场景。设配置 N_{BS} 根天线的 BS 同时和 U 个 MS 进行通信, 每个 MS 均具有 N_{MS} 根天线。假设每个 MS 仅采用模拟预编码架构, 即每个 MS 仅有一路数据流, 则 BS 总共有 $N_s = U$ 路数据流。不失一般性, 可假设 BS 使用 N_{RF} 个可用数据链中的 U 个数据链对这 U 个 MS 进行服务, 且 $N_{\text{RF}} \geq U$ 。在 BS 侧, $U \times 1$ 维发送符号 $\mathbf{s} = [s_1, \dots, s_U]^T$ 依次经过 $U \times U$ 维数字预编码器 $\mathbf{F}_{\text{BB}} = [\mathbf{f}_{\text{BB},1}, \dots, \mathbf{f}_{\text{BB},U}]$ 和 $N_{\text{BS}} \times U$ 维模拟预编码器 \mathbf{F}_{RF} 的处理, 其中, s_u 表示发送给第 u 个 MS 的信号, $\mathbf{f}_{\text{BB},u}$ 表示与第 u 个 MS 相对应的数字预编码向量。第 u 个 MS 经过 $N_{\text{MS}} \times 1$ 维模拟组合器 \mathbf{w}_u 处理后的接收信号可表示为

$$y_u = \sqrt{\rho} \mathbf{w}_u^H \mathbf{H}_u \sum_{j=1}^U \mathbf{F}_{\text{RF}} \mathbf{f}_{\text{BB},j} s_j + \mathbf{w}_u^H \mathbf{n}_u \quad (26)$$

其中, \mathbf{H}_u 为 $N_{\text{MS}} \times N_{\text{BS}}$ 维信道矩阵, 且与式(2)有相同的表达式; $\mathbf{n}_u \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I}_{N_{\text{MS}}})$ 为 $N_{\text{MS}} \times 1$ 维高斯噪声向量。则第 u 个 MS 可实现的速率为

$$R_u = \text{lb} \left(1 + \frac{\frac{\rho}{U} \left| \mathbf{w}_u^H \mathbf{H}_u \mathbf{F}_{\text{RF}} \mathbf{f}_{\text{BB},u} \right|^2}{\frac{\rho}{U} \sum_{j \neq u} \left| \mathbf{w}_u^H \mathbf{H}_u \mathbf{F}_{\text{RF}} \mathbf{f}_{\text{BB},j} \right|^2 + \sigma^2} \right) \quad (27)$$

其中, ρ 表示平均发送功率。

与单用户场景下类似, 多用户混合预编码设计的优化目标是最大化所有 MS 的和速率, 即

$$\begin{aligned} \left\{ \mathbf{F}_{\text{RF}}^{\text{opt}}, \left\{ \mathbf{f}_{\text{BB},u}^{\text{opt}} \right\}_{u=1}^U, \left\{ \mathbf{w}_u^{\text{opt}} \right\}_{u=1}^U \right\} &= \arg \max \sum_{u=1}^U R_u \\ \text{s.t. } \mathbf{F}_{\text{RF}}(\cdot, u) &\in \mathcal{F}, \mathbf{w}_u \in \mathcal{W}, \left\| \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}} \right\|_{\text{F}}^2 = U \end{aligned} \quad (28)$$

其中, \mathcal{F} 和 \mathcal{W} 分别表示模拟预编码向量和模拟组合向量的可能取值所组成的码本。式(28)是一个混整数规划问题, 它的解需要搜索整个 $\mathcal{F}^U \times \mathcal{W}^U$ 空间的所有 \mathbf{F}_{RF} 和 $\{\mathbf{w}_u\}_{u=1}^U$ 的可能组合。因此, 和速率最大化问题的直接解决方案既不实际也不可行。为此, 本文提出一种两阶段法, 将混合预编码矩阵和模拟组合向量分开进行设计。

具体地, 在第一阶段, 每个 MS 设计各自的模拟组合向量, 以最大化信道增益, 即

$$\begin{aligned} \mathbf{w}_u^{\text{opt}} &= \arg \max \left\| \mathbf{w}_u^H \mathbf{H}_u \right\|_2^2 \\ \text{s.t. } \mathbf{w}_u &\in \mathcal{W} \end{aligned} \quad (29)$$

在第二阶段, 用户计算各自的等效信道 $\mathbf{h}_u = \mathbf{w}_u^H \mathbf{H}_u$ 并反馈给 BS。据此 BS 将设计理想情况下的 ZF 全数字预编码矩阵 $\mathbf{F}_{\text{ZF}} = \bar{\mathbf{H}}^H (\bar{\mathbf{H}} \bar{\mathbf{H}}^H)^{-1}$, 其中 $\bar{\mathbf{H}} = [\mathbf{h}_1, \dots, \mathbf{h}_U]^H$ 。然后 BS 以最小化 $\left\| \mathbf{F}_{\text{ZF}} - \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}} \right\|_{\text{F}}^2$ 为优化目标, 采用算法 1 设计 \mathbf{F}_{RF} 和 \mathbf{F}_{BB} 。最后, BS 通过归一化获得对应每个 BS 的数字预编码向量, 即 $\mathbf{f}_{\text{BB},u} = \frac{\mathbf{f}_{\text{BB},u}}{\left\| \mathbf{F}_{\text{RF}} \mathbf{f}_{\text{BB},u} \right\|_{\text{F}}}$ 。

4 仿真分析

本节使用 MATLAB 软件对所提算法的性能进行详细的仿真与评估,所用处理器的型号为 Intel(R) Core(TM) i5-4210M,主频 2.60 GHz,机带 RAM 8 GHz。首先考虑单用户场景,设仿真中的发射机和接收机分别具有 64 根和 16 根天线;发射机和接收机射频链个数相等,统一用 N_{RF} 表示;数据流个数 $N_S = 2$ 。毫米波信道的实例依据式(2)和式(3)进行生成,其中簇个数 $N_{cl} = 5$,每个簇的中心到达角和离开角均匀分布在 $[0, 2\pi)$ 内;每个簇有 $N_{ray} = 10$ 个射线,到达角和离开角服从角度扩展为 10° 的拉普拉斯分布;第 i 个簇中第 l 个射线的复增益为 $\alpha_{il} \sim \mathcal{CN}(0, \sigma_{\alpha,i}^2)$,其中 $\sigma_{\alpha,i}^2 = 1$ 表示第 i 个簇的平均功率。设天线间距 d 为半波长,噪声功率 σ_n^2 为 1。信噪比 (SNR, signal-to-noise-ratio) 定义为 ρ/σ_n^2 。 ε 、 δ 和 t_{max} 分别设置为 10^{-3} 、 10^{-6} 和 500。以下分析将在随机产生 1 000 个信道实例的情况下,将本文算法与传统的 OMP^[7]、OPP^[13]、CG-AltMin^[13]、BB-AltMin^[15]、GP-AltMin^[16]等算法和基于 GDM^[20] 的在线学习方案进行对比。在 GDM 中,最大迭代次数和动量因子 α 分别设置为 2 000 和 0.9,为区别于本文算法,学习率用 μ 表示,另外,获得的模拟预编码器通过式(19)进行模值归一化处理。

图 2 为 $N_{RF} = 4$ 、SNR=0 时,不同学习率情况下本文算法和 GDM 的可实现频谱效率与训练数据集大小 N_d 的关系。对于 BP 神经网络,网络输出与输入的映射关系为非线性函数。对于非线性函数,训练数据集越大,函数的拟合效果越好,训练得到的网络模型才最接近真实函数。

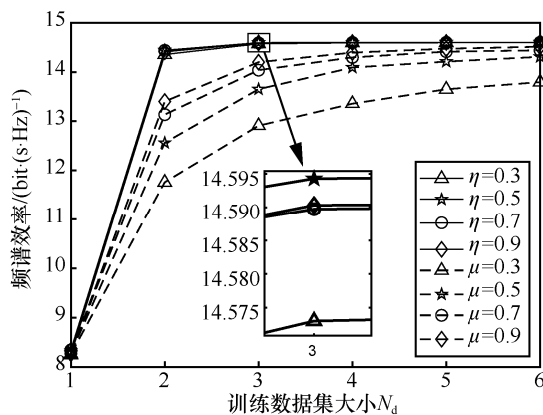


图 2 不同学习率情况下频谱效率与训练数据集大小的关系

从图 2 可以看到,频谱效率随着 N_d 的增加而增加,对于本文算法, $N_d = 3$ 时频谱效率趋于最大值;对于 GDM, $N_d = 4$ 时频谱效率趋于最大值。由此可见,本文算法仅需较少的训练数据集即可达到较满意的结果。在本文算法中,每个权重的搜索步长与之前迭代的一阶导数的平方和有关,仅利用了一阶导数信息即可起到与二阶方法和模拟退火相同的性能,将会使频谱效率收敛到一个稳定值。因此,给定 N_d 和学习率时,本文算法的频谱效率性能优于 GDM。从图 2 中还可以看出,在给出的几个学习率中, $\eta = 0.5$ 时本文算法的频谱效率最大, $\mu = 0.9$ 时 GDM 的频谱效率最大。

图 3 为 $N_d = 3$ 、SNR=0 时不同学习率的收敛情况对比。本文算法的收敛结束条件是式(6)和式(25)中的优化目标趋于最小,等价于使优化式(5)中的频谱效率趋于最大,因此迭代次数会收敛于频谱效率。对于本文算法, $\eta = 0.5$ 时的收敛速度较快(平均迭代次数为 192,平均收敛时间约为 0.131 4 s),且实现的频谱效率较高,平均约为 14.58 bit/(s·Hz);对于 GDM, $\mu = 0.9$ 时的收敛速度较快,且实现的频谱效率较高。给定学习率时,本文算法的收敛速度远远快于 GDM。从图 2 和图 3 可以看到,当 $N_d = 3$ 、 $\eta = 0.5$ 时,本文算法性能最好;当 $N_d = 4$ 、 $\mu = 0.9$ 时, GDM 性能最好。因此在后续仿真中,对于本文算法,均令 $N_d = 3$ 、 $\eta = 0.5$;对于 GDM,均令 $N_d = 4$ 、 $\mu = 0.9$ 。

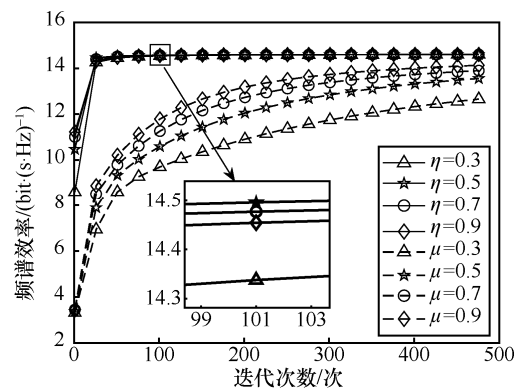


图 3 不同学习率的收敛情况对比

图 4 为 $N_{RF} = 4$ 时不同算法的性能对比。根据式(4),随着 SNR 的增加,所有算法可实现的频谱效率均增加。在给定 SNR 的前提下,本文算法可实现的频谱效率均高于参与对比的算法,最接近无约束全数字预编码的性能。在本文算法中,经过收

敛, 模拟预编码矩阵与数字预编码矩阵的乘积 $F_{RF}F_{BB}$ 接近无约束全数字预编码矩阵 F_{opt} , 模拟组合矩阵和数字组合矩阵的乘积 $W_{RF}W_{BB}$ 接近无约束全数字组合矩阵 W_{opt} , 因此所实现频谱效率会接近无约束全数字预编码。

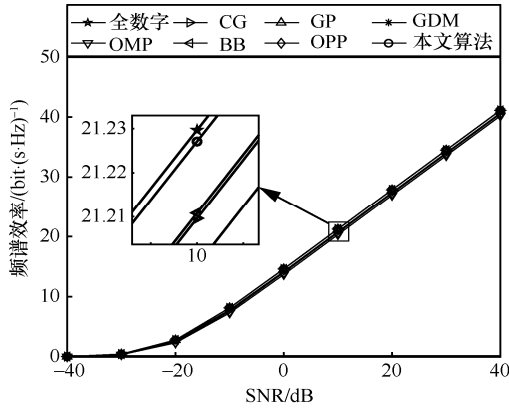


图 4 $N_{RF} = 4$ 时不同算法的性能对比

图 5 为不同算法可实现的频谱效率随 N_{RF} 的变化情况。随着 N_{RF} 的增加, 本文算法、OMP、CG-AltMin、BB-AltMin、GP-AltMin 和 GDM 可实现的频谱效率均增加。在给定 N_{RF} 的条件下, 本文算法可实现的频谱效率高与其他算法, 且在 $N_{RF} = 4$ 时接近无约束全数字预编码的性能。文献[33]证明, 在混合预编码架构下, $N_{RF} \geq 2N_S$ 是实现无约束全数字预编码性能的充分条件, 图 5 的仿真结果也证实了这一结论。尽管 OPP 相比上述算法具有较低的复杂度, 但可实现的频谱效率随着 N_{RF} 的增加几乎不发生变化。这是由于 OPP 假设 F_{BB} 满足比例酉性质, 但这一假设仅在 $N_{RF} = N_S$ 时成立。因此, 当 $N_{RF} > N_S$ 时, OPP 所实现的频谱效率在一定程度上会有所降低。

为了比较不同算法的复杂度, 表 3 列出了本文算法、CG-AltMin、BB-AltMin、GP-AltMin 和 GDM

在不同 N_{RF} 时的平均迭代次数。将表 3 的数据代入表 2 的公式中, 得到不同算法的计算开销随 N_{RF} 的变化曲线, 如图 6 所示。从图 6 中可以看出, 本文算法的复杂度明显低于 CG-AltMin、BB-AltMin、GP-AltMin 和 GDM, 且在 $N_{RF} > 3$ 时随着 N_{RF} 的增加而降低。因为表 3 的数据显示, 当 $N_{RF} > 3$ 时, 随着 N_{RF} 的增加, 本文算法的迭代次数 N_{iter} 不断降低, 将 N_{iter} 代入表 2 的公式中, N_{iter} 降低的幅度大于复杂度仅在 N_{RF} 作用下增加的幅度, N_{iter} 和 N_{RF} 的联合作用造成复杂度的降低。而参与对比的交替最小算法和 GDM 算法的迭代次数尽管也随着 N_{RF} 的增大而减小, 但减小的程度远不如本文算法, 因此没有出现复杂度随 N_{RF} 增加而降低的情况。

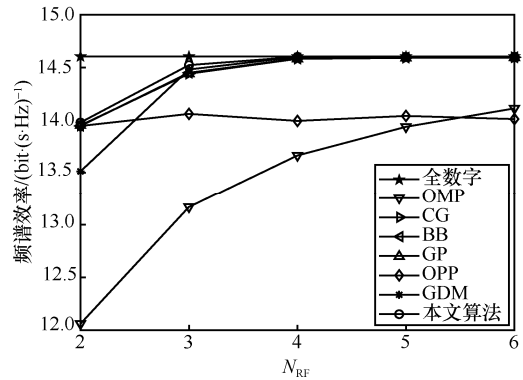


图 5 不同算法可实现的频谱效率随 N_{RF} 的变化情况

最后, 评估多用户场景下不同算法的性能。设配置 64 根天线的基站同时和 U 个用户进行通信, 每个用户均具有 16 根天线。每个用户与基站间毫米波信道实例的产生方式与相应参数的设置与单用户场景一致。以下在随机产生 1 000 个信道实例的条件下, 将本文算法与 ZF^[18]、MMSE^[19]、CG-AltMin^[13]和 GDM^[22]进行对比分析。

图 7 为 $U = 4$ 时不同算法的和速率与 SNR 的关系。根据式(27), 随着 SNR 的增加, 所有算法可实现

表 3 不同算法的平均迭代次数

N_{RF}	本文算法	CG		BB		GP		GDM
		N_{iter}^o	N_{iter}^i	N_{iter}^o	N_{iter}^i	N_{iter}^o	N_{iter}^i	N_{iter}^{GDM}
2	314.503	3.644	80.213	3.727	102.717	3.846	63.102	2 000
3	453.824	4.016	111.852	3.981	187.170	4.715	126.886	2 000
4	192.281	3.720	104.735	3.755	249.381	3.992	113.690	1 955.35
5	81.701	3.565	94.382	3.604	231.386	3.679	124.113	1 547.26
6	56.548	3.391	86.065	3.427	220.753	3.528	190.989	1 236.19

的和速率均增加。从图 7 中可看出，本文算法实现的和速率高于其他算法，这是因为本文算法设计的混合预编码矩阵最接近理想情况下的 ZF 全数字预编码矩阵。

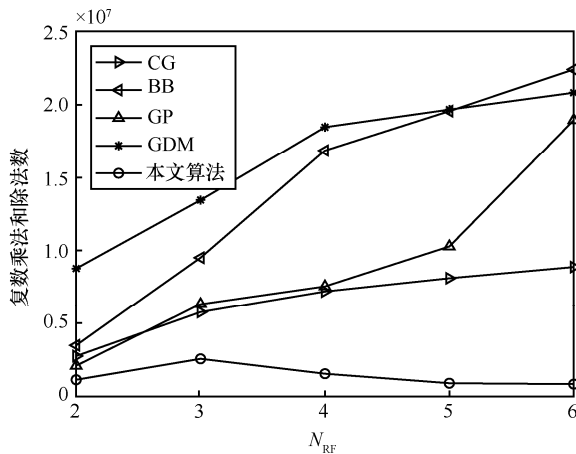


图 6 不同算法的计算开销随 N_{RF} 的变化曲线

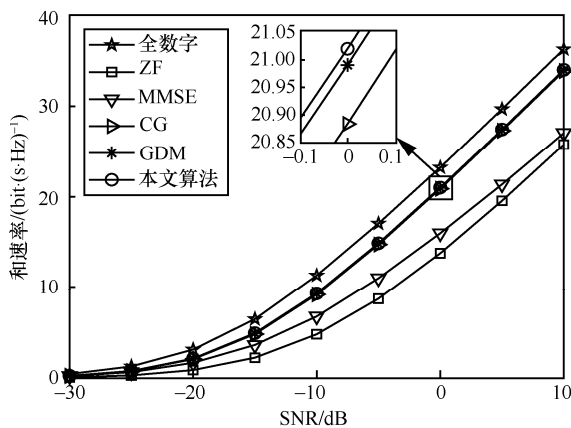


图 7 $U = 4$ 时不同算法的和速率与 SNR 的关系

图 8 为 SNR=0 时不同算法的和速率与 U 的关系。随着 U 的增加，所有算法实现的和速率均增加，但 U 的增加会造成每个用户所受干扰的增大，因此和速率增长的趋势会变缓。在给定 U 的前提下，本文算法实现的和速率与 GDM 基本相同，且均高于 ZF 和 MMSE。与 CG-AltMin 相比，当 $U \leq 4$ 时，本文算法实现的和速率与之相同；当 $U > 4$ 时，本文算法实现的和速率更高；且随着 U 的增加，本文算法和 CG-AltMin 所实现和速率的差值不断增大。

综上所述，在单用户场景和多用户场景下，本文算法可实现的频谱效率均高于现有的算法，接近全数字预编码的性能。而在算法复杂度方面，与基于交替最小的次优算法和基于 GDM 的在线学习方案相比，本文算法的复杂度明显降低。

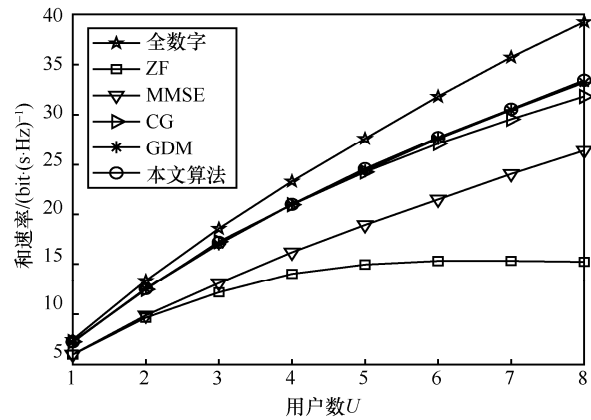


图 8 SNR=0 时不同算法的和速率与 U 的关系

5 结束语

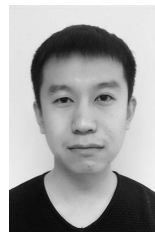
本文针对毫米波大规模 MIMO 系统，分别提出了适用于单用户场景和多用户场景的低复杂度混合预编码方法。首先针对单用户场景，提出了一种等效的单隐藏层神经网络架构，将混合预编码的设计问题等效为该网络架构下的权值训练过程。然后以最小化损失函数为目标，提出了基于 AG-BP 算法的混合预编码机制。最后将所提算法扩展至多用户场景。仿真结果表明，所提算法可实现的频谱效率高于现有的算法，接近全数字预编码的性能，且复杂度低于基于交替最小的次优算法和基于动量梯度下降的在线学习方案。

参考文献:

- [1] RAPPAPORT T S, SUN S, MAYZUS R, et al. Millimeter wave mobile communications for 5G cellular: it will work![J]. IEEE Access, 2013, 1: 335-349.
- [2] ROH W, SEOL J Y, PARK J, et al. Millimeter-wave beamforming as an enabling technology for 5G cellular communications: theoretical feasibility and prototype results[J]. IEEE Communications Magazine, 2014, 52(2): 106-113.
- [3] SHAFI M, MOLISCH A F, SMITH P J, et al. 5G: a tutorial overview of standards, trials, challenges, deployment, and practice[J]. IEEE Journal on Selected Areas in Communications, 2017, 35(6): 1201-1221.
- [4] RANGAN S, RAPPAPORT T S, ERKIP E. Millimeter-wave cellular wireless networks: potentials and challenges[J]. Proceedings of the IEEE, 2014, 102(3): 366-385.
- [5] SALOUS S, ESPOSTI V D, FUSCHINI F, et al. Millimeter-wave propagation: characterization and modeling toward fifth-generation systems[J]. IEEE Antennas and Propagation Magazine, 2016, 58(6): 115-127.
- [6] HEATH R W, GONZÁLEZ-PRELCIC N, RANGAN S, et al. An overview of signal processing techniques for millimeter wave MIMO systems[J]. IEEE Journal of Selected Topics in Signal Processing, 2016, 10(3): 436-453.
- [7] AYACH O E, RAJAGOPAL S, ABU-SURRA S, et al. Spatially sparse

- precoding in millimeter wave MIMO systems[J]. IEEE Transactions on Wireless Communications, 2014, 13(3): 1499-1513.
- [8] LEE Y Y, WANG C H, HUANG Y H. A hybrid RF/baseband precoding processor based on parallel-index-selection matrix-inversion-by-pass simultaneous orthogonal matching pursuit for millimeter wave MIMO systems[J]. IEEE Transactions on Signal Processing, 2015, 63(2): 305-317.
- [9] ZHANG Y, HUANG Y Z, QIN X Q, et al. Low complexity hybrid precoding based on ORLS for mmWave massive MIMO systems[C]//Proceedings of 2018 IEEE Wireless Communications and Networking Conference (WCNC). Piscataway: IEEE Press, 2018: 1-6.
- [10] HUNG W L, CHEN C H, LIAO C C, et al. Low-complexity hybrid precoding algorithm based on orthogonal beamforming codebook[C]//Proceedings of 2015 IEEE Workshop on Signal Processing Systems (SiPS). Piscataway: IEEE Press, 2015: 1-5.
- [11] 廖勇, 杨馨怡, 杜洁汝. 基于两阶段的毫米波大规模 MIMO 低复杂度混合预编码算法[J]. 电子学报, 2021, 49(7): 1298-1304.
LIAO Y, YANG X Y, DU J R. A two-stage based low complexity hybrid precoding algorithm for millimeter-wave massive MIMO[J]. Acta Electronica Sinica, 2021, 49(7): 1298-1304.
- [12] 刘剑飞, 何利平, 陶颖, 等. 大规模 MIMO 系统中低复杂度的码本搜索方法[J]. 通信学报, 2019, 40(1): 79-86.
LIU J F, HE L P, TAO Y, et al. Low complexity codebook search method in massive MIMO system[J]. Journal on Communications, 2019, 40(1): 79-86.
- [13] YU X H, SHEN J C, ZHANG J, et al. Alternating minimization algorithms for hybrid precoding in millimeter wave MIMO systems[J]. IEEE Journal of Selected Topics in Signal Processing, 2016, 10(3): 485-500.
- [14] NI W H, DONG X D, LU W S. Near-optimal hybrid processing for massive MIMO systems via matrix decomposition[J]. IEEE Transactions on Signal Processing, 2017, 65(15): 3922-3933.
- [15] MULLA M, ULUSOY A H, RIZANER A, et al. Barzilai-borwein gradient algorithm based alternating minimization for single user millimeter wave systems[J]. IEEE Wireless Communications Letters, 2020, 9(4): 508-512.
- [16] CHEN J C. Gradient projection-based alternating minimization algorithm for designing hybrid beamforming in millimeter-wave MIMO systems[J]. IEEE Communications Letters, 2019, 23(1): 112-115.
- [17] RUSU C, MÈNDEZ-RIAL R, GONZÁLEZ-PRELCIC N, et al. Low complexity hybrid precoding strategies for millimeter wave communication systems[J]. IEEE Transactions on Wireless Communications, 2016, 15(12): 8380-8393.
- [18] ALKHATEEB A, LEUS G, HEATH R W. Limited feedback hybrid precoding for multi-user millimeter wave systems[J]. IEEE Transactions on Wireless Communications, 2015, 14(11): 6481-6494.
- [19] NGUYEN D H N, LE L B, LE-NGOC T. Hybrid MMSE precoding for mmWave multiuser MIMO systems[C]//Proceedings of 2016 IEEE International Conference on Communications (ICC). Piscataway: IEEE Press, 2016: 1-6.
- [20] NI W H, DONG X D. Hybrid block diagonalization for massive multiuser MIMO systems[J]. IEEE Transactions on Communications, 2016, 64(1): 201-211.
- [21] HUANG H J, SONG Y W, YANG J, et al. Deep-learning-based millimeter-wave massive MIMO for hybrid precoding[J]. IEEE Transactions on Vehicular Technology, 2019, 68(3): 3027-3032.
- [22] CHEN K, YANG J, GE X H, et al. Complex-BP-neural-network-based hybrid precoding for millimeter wave multiuser massive MIMO systems[C]//Proceedings of 2019 Computing, Communications and IoT Applications (ComComAp). Piscataway: IEEE Press, 2019: 100-105.
- [23] SIDHARTH C, HIREMATH S M, PATRA S K. Deep learning based hybrid precoding for mmwave massive MIMO system using concept[C]//Proceedings of 2020 International Conference on Communication and Signal Processing (ICCCSP). Piscataway: IEEE Press, 2020: 1317-1321.
- [24] BAO X L, FENG W J, ZHENG J L, et al. Deep CNN and equivalent channel based hybrid precoding for mmWave massive MIMO systems[J]. IEEE Access, 2020, 8: 19327-19335.
- [25] ELBIR A M. A deep learning framework for hybrid beamforming without instantaneous CSI feedback[J]. IEEE Transactions on Vehicular Technology, 2020, 69(10): 11743-11755.
- [26] ELBIR A M, PAPAFAEIROPOULOS A K. Hybrid precoding for multiuser millimeter wave massive MIMO systems: a deep learning approach[J]. IEEE Transactions on Vehicular Technology, 2020, 69(1): 552-563.
- [27] CHAI M Y, TANG S H, ZHAO M, et al. HPNet: a compressed neural network for robust hybrid precoding in multi-user massive MIMO systems[C]//Proceedings of GLOBECOM 2020 - 2020 IEEE Global Communications Conference. Piscataway: IEEE Press, 2020: 1-7.
- [28] MA W Y, QI C H, ZHANG Z C, et al. Sparse channel estimation and hybrid precoding using deep learning for millimeter wave massive MIMO[J]. IEEE Transactions on Communications, 2020, 68(5): 2838-2849.
- [29] 章坚武, 王路鑫, 孙玲芬, 等. 人工智能在 5G 系统中的应用综述[J]. 电信科学, 2021, 37(5): 14-31.
ZHANG J W, WANG L X, SUN L F, et al. An survey on application of artificial intelligence in 5G system[J]. Telecommunications Science, 2021, 37(5): 14-31.
- [30] ALKHATEEB A, EL AYACH O, LEUS G, et al. Channel estimation and hybrid precoding for millimeter wave cellular systems[J]. IEEE Journal of Selected Topics in Signal Processing, 2014, 8(5): 831-846.
- [31] LEUNG H, HAYKIN S. The complex backpropagation algorithm[J]. IEEE Transactions on Signal Processing, 1991, 39(9): 2101-2104.
- [32] DUCHI J, HAZAN E, SINGER Y. Adaptive subgradient methods for online learning and stochastic optimization[J]. Journal of Machine Learning Research, 2011, 12: 2121-2159.
- [33] SOHRABI F, YU W. Hybrid digital and analog beamforming design for large-scale antenna arrays[J]. IEEE Journal of Selected Topics in Signal Processing, 2016, 10(3): 501-513.

[作者简介]



张煜 (1992-), 男, 山东滨州人, 北京邮电大学博士生, 主要研究方向为大规模 MIMO、混合预编码。

张治 (1977-), 男, 河北安平人, 博士, 北京邮电大学教授、博士生导师, 主要研究方向为无线信号智能检测与高效通信的相关理论、关键技术与系统设计等。

董晓岱 (1970-), 女, 河南洛阳人, 博士, 维多利亚大学教授、博士生导师, 主要研究方向为 5G、毫米波通信、太赫兹通信等。