

大规模 MIMO 系统中基于深度学习的新型信号检测网络^{*}

康晓非 井溢洋 郭含玉

(西安科技大学通信与信息工程学院 西安 710600)

摘要: 为了解决高阶调制场景下大规模 MIMO 检测算法的误码性能受限问题,提出一种新型信号检测网络。该网络基于投影梯度下降构建逼近最大似然解的迭代结构,并将迭代过程转化神经网络实现。在每个网络单元中首先通过神经网络学习参数,其次经过所设计的归一化多段激活函数进行非线性变换以增强网络在高阶调制下的映射能力,最后通过去噪器消除估计误差和信道噪声,此外,为了解决在网络深度增加时准确率下降问题,网络单元之间采用残差连接。仿真结果表明,当系统收发天线数为 64×32 ,信噪比为 16 dB 时,所提出检测网络的误码率接近 10^{-4} ,与其他检测算法相比,误码率降低一个数量级,表现出接近最优检测算法的误码性能,且具有较好的鲁棒性。

关键词: 大规模 MIMO;深度学习;信号检测;激活函数;去噪器

中图分类号: TN929.5 **文献标识码:** A **国家标准学科分类代码:** 510.4010

Deep learning-based novel detection networks for massive MIMO systems

Kang Xiaofei Jing Yiyang Guo Hanyu

(School of Communication and Information Engineering, Xi'an University of Science and Technology, Xi'an 710600, China)

Abstract: To address the limitation of error performance in Massive MIMO detection algorithms under high-order modulation scenarios, a novel detection network is proposed. This network is constructed based on a projected gradient descent framework that approximates the maximum likelihood solution through an iterative structure, which is then implemented using a neural network. In each network module, parameters are first learned through a neural network, followed by a nonlinear transformation using a designed normalized multi-segment activation function to enhance the network's mapping capability under high-order modulation. Finally, a denoiser is employed to eliminate estimation errors and channel noise. Furthermore, to tackle the issue of accuracy degradation with increased network depth, residual connections are introduced between network modules. Simulation results show that when the number of transmitting and receiving antennas of the system is 64×32 and the signal-to-noise is 16 dB, the bit error rate of the proposed detection network is close to 10^{-4} , and the bit error rate is reduced by an order of magnitude compared with other detection algorithms, showing the bit error performance close to the optimal detection algorithm, and has good robustness.

Keywords: massive MIMO; deep learning; MIMO detection; activation function; denoiser

0 引言

大规模 MIMO 技术具有数据传输速率高、信道容量大的显著优势,有效缓解频谱资源有限与不断增长的容量需求之间的矛盾^[1-2]。然而,大规模 MIMO 对高维信号处理时的复杂性也带来了新的挑战,其中最核心的挑战之一便是如何在庞大的天线规模和复杂信道条件下,快速且准确地从失真的接收信号中检测出各个数据流的信息。

为解决这一关键问题,研究者提出了多种信号检测方法,主要可分为传统检测方法和基于深度学习的方法。其中传统信号检测方法包括最大似然(maximum likelihood, ML)检测、线性检测和非线性检测。其中最大似然检测通过对所有可能的发射符号向量进行穷举搜索,计算其与接收信号在信道作用下的欧式距离,并选择距离最小者作为检测结果,具有理论上的最优误码性能,但其计算复杂度随系统规模呈指数级增长,已被证明为 NP-hard 问题^[3],因此

难以直接应用于大规模 MIMO 系统。为了在性能与复杂度之间取得平衡,随后提出了线性检测方法和非线性检测方法。常用的线性检测方法如迫零^[4] (zero forcing, ZF) 检测与最小均方误差^[5] (minimum mean square error, MMSE) 检测等。这类方法将期望信号以外的其他信号视为干扰,通过最小化干扰来进行信号检测,具有实现简单,复杂度较低等优点。非线性检测方法是线性检测的拓展,如近似消息传递 (approximate message passing, AMP)^[6] 是一种典型的非线性检测方法,其核心思想是将系统建模为图结构,通过引入统计平均和残差修正项,对传统贝叶斯消息传递过程进行近似,大幅降低了计算复杂度,但该方法对信道分布的依赖性较强,在非理想信道中稳定性不足。

基于以上分析,传统信号检测方法在误码性能和计算复杂度之间难以实现有效平衡,尤其在大规模 MIMO 系统中更显局限。为解决这一问题,近年来研究者开始借助深度学习的强大建模能力探索新的检测策略。随着其在图像识别^[7-8]、自然语言处理^[9-10] 和行为识别^[11-12] 等领域取得显著成果,深度学习已被广泛应用于通信系统的各类信号处理任务,其中 MIMO 信号检测成为研究的重点方向之一,并涌现出多种基于深度学习的信号检测网络。典型如 DetNet^[13], 基于投影梯度法构建可端到端训练的神经网络,在实现较高检测精度的同时具备可学习参数,但因采用全连接架构,导致参数规模庞大,复杂度随网络深度显著上升。为此, ScNet^[14] 则通过对 DetNet 分别从网络输入、连接结构和损失函数三方面的简化处理,明显降低了计算复杂,但 ScNet 对 DetNet 结构的过度简化导致其在复杂信道和高阶调制场景下,其检测精度和泛化能力受限。基于此,RPD-Net^[15] 引入结构随机性机制,在保持 DetNet 基本框架的同时,通过随机扰动加先验信息方式强化网络的表达能力,从而在提升检测性能的同时兼顾泛化能力与计算效率;为进一步增强检测网络对多用户干扰和噪声扰动的抑制能力,后续研究提出了 EScNe^[16]。该模型融合了卷积神经网络与传统干扰消除机制,通过串行或并行的干扰抵消结构强化特征提取过程,有效提升了网络在大规模 MIMO 场景中的干扰抑制能力和检

测精度;文献[17]提出的基于迭代软阈值理论的 MMNet,通过并引入可学习参数以实现自适应调整,有效细化了信号估计过程,缓解了误差扩散问题,进一步为基于深度学习的检测算法在实际系统中的应用提供了新的思路。

综上所述,尽管现有基于深度学习的检测网络在误码性能与计算复杂度之间取得了一定平衡,但在高阶调制场景下仍存在误码性能受限问题,一方面,多轮迭代中缺乏对误差传播机制的有效控制,容易造成估计误差的累积放大;另一方面,复杂的星座结构对检测网络的特征提取与判决边界提出了更高要求,从而影响网络检测的误码性能。为克服上述问题,提出了一种基于投影梯度下降迭代展开的最大似然检测网络,命名为 MDNet (multi-segement denoise network)。其创新之处在于:1)设计了一种“激活函数-去噪器”顺序优化结构,使得在完成非线性激活后立即进行误差抑制,从而有效缓解误差放大与通道噪声的干扰;2)提出归一化多段激活函数,利用组合 sigmoid 结构增强其对复杂符号分布的适应能力,有效解决传统激活函数对高阶调制下星座映射能力不足的问题。

1 算法原理

MDNet 的整体迭代框架如图 1 所示,该框架主要由两个阶段构成:信号预处理阶段与信号估计阶段。在信号预处理阶段,接收端接收到的信号 \mathbf{y} 与已知信道矩阵 \mathbf{H}^H , 经线性变换后生成有用的中间信息 $\mathbf{H}^H \mathbf{y}$ 与 $\mathbf{H}^H \mathbf{H}$, 并结合上一轮迭代得到的估计信号 $\hat{\mathbf{x}}_k$ 将三者作为输入传递至网络单元 K 进行信号估计。在信号估计阶段中,来自预处理阶段的 $\mathbf{H}^H \mathbf{y}$ 、 $\mathbf{H}^H \mathbf{H} \hat{\mathbf{x}}_k$ 以及 $\hat{\mathbf{x}}_k$ 三个向量首先通过神经网络线性运算,随后通过提出的归一化多段激活函数进行非线性变换,从而获得中间估计信号。该中间估计信号与由信号估计误差累积项以及通道噪声项共同构成的总噪声方差一起作为输入送入去噪模块进行估计误差消除和噪声抑制处理。最终,去噪器输出当前迭代的信号估计结果 $\hat{\mathbf{x}}_{k+1}$, 并通过残差结构处理后以供下一轮迭代使用。

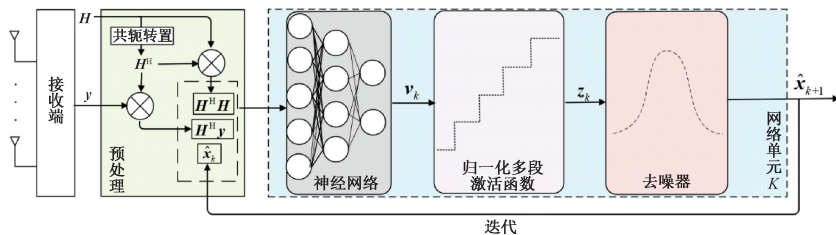


图 1 MDNet 的整体迭代框架

Fig. 1 The iterative framework of MDNet

2 系统模型

2.1 接收信号模型

在 MIMO 系统中发射天线数为 N_t , 接收天线数为

N_r , 其中 $N_t \leq N_r$, 其接收信号可以表示为:

$$\mathbf{y} = \mathbf{H}_c \mathbf{x}_c + \mathbf{n}_c \quad (1)$$

其中, 发射信号为 $\mathbf{x}_c \in \mathcal{X}^{N_t \times 1}$, \mathcal{X} 表示信号星座点的集合, 发射信号通过 MIMO 时变衰落信道 $\mathbf{H}_c \in \mathbf{C}^{N_r \times N_t}$ 传输

到达接收端, $\mathbf{n}_c \in \mathbf{C}^{N_r \times 1}$ 是均值为 0, 协方差矩阵为 $\sigma^2 \mathbf{I}$ 加性高斯白噪声。

为了应用于神经网络, 而需将复数域模型转换为实数域, 其等价的实数模型为:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \quad (2)$$

$$\mathbf{y} = \begin{bmatrix} \Re(\mathbf{y}_c) \\ \Im(\mathbf{y}_c) \end{bmatrix}, \quad \mathbf{H} = \begin{bmatrix} \Re(\mathbf{H}_c) & -\Im(\mathbf{H}_c) \\ \Im(\mathbf{H}_c) & \Re(\mathbf{H}_c) \end{bmatrix}, \quad (3)$$

$$\mathbf{x} = \begin{bmatrix} \Re(\mathbf{x}_c) \\ \Im(\mathbf{x}_c) \end{bmatrix}, \quad \mathbf{n} = \begin{bmatrix} \Re(\mathbf{n}_c) \\ \Im(\mathbf{n}_c) \end{bmatrix}$$

其中, $\Re(\cdot)$ 表示该值的实部, $\Im(\cdot)$ 表示该值的虚部,

$\mathbf{x} \in \mathbf{R}^{2N_t}$, $\mathbf{H} \in \mathbf{R}^{2N_r \times 2N_t}$, $\mathbf{y} \in \mathbf{R}^{2N_r}$, $\mathbf{n} \in \mathbf{R}^{2N_r}$ 。

2.2 提出的 MDNet 模型

MDNet 是基于投影梯度下降法求 ML 解问题的迭代展开, 而最大似然估计可表示为以下优化问题:

$$\hat{\mathbf{x}}_{ML} = \underset{\mathbf{x} \in \mathcal{X}^M}{\operatorname{argmin}} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|_2^2 \quad (4)$$

为求解上述问题, 引入梯度下降法, 并在每一步将结果投影回可行域 \mathcal{X}^M , 则对应的投影梯度下降更新公式为:

$$\hat{\mathbf{x}}_{k+1} = \Pi[\hat{\mathbf{x}}_k - \delta_k \nabla \|\mathbf{y} - \mathbf{H}\mathbf{x}\|_2^2] \quad (5)$$

其中, δ_k 为第 k 次迭代的步长, $\Pi[\cdot]$ 表示将向量投影到可行域 \mathcal{X}^M 。

将梯度展开, 式(5)可表示为:

$$\hat{\mathbf{x}}_{k+1} = \Pi[\hat{\mathbf{x}}_k - \delta_k \mathbf{H}^H \mathbf{H} \hat{\mathbf{x}}_k + \delta_k \mathbf{H}^H \mathbf{y}] \quad (6)$$

上式迭代可用神经网络实现, 其中, 神经网络有 K 层, 每一层为一个网络单元, K 层则对应 K 次迭代, 并用可学习参数 \mathbf{W}_k 和 \mathbf{b}_k 替代固定步长 δ_k , 用归一化激活函数替代投影算子。在网络单元中, 将 $\hat{\mathbf{x}}_k$, $\mathbf{H}^H \mathbf{y}$ 和 $\mathbf{H}^H \mathbf{H} \hat{\mathbf{x}}_k$ 通过先后通过向量拼接、神经网络、归一化多段激活函数和去噪器一系列过程处理。具体地, MDNet 网络结构如图 2 所示。

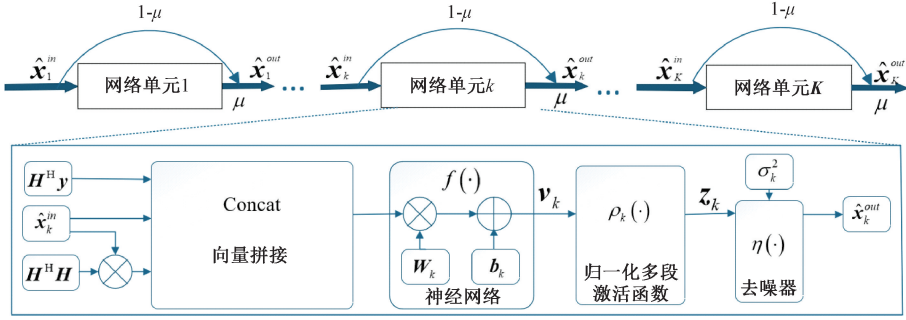


图 2 MDNet 网络结构

Fig. 2 Network structure of MDNet

首先, $\hat{\mathbf{x}}_k$, $\mathbf{H}^H \mathbf{y}$ 和 $\mathbf{H}^H \mathbf{H} \hat{\mathbf{x}}_k$ 进行拼接后输入神经网络经过线性变换得到如下形式:

$$\mathbf{v}_k = f \left(\begin{bmatrix} \mathbf{H}^H \mathbf{y} \\ \hat{\mathbf{x}}_k \\ \mathbf{H}^H \mathbf{H} \hat{\mathbf{x}}_k \end{bmatrix} \right) \quad (7)$$

其中, \mathbf{v}_k 是经过神经网络线性变换的输出信号, $f(\cdot)$ 代表神经网络中线性运算部分, 即:

$$\mathbf{v}_k = \mathbf{W}_k \begin{bmatrix} \mathbf{H}^H \mathbf{y} \\ \hat{\mathbf{x}}_k \\ \mathbf{H}^H \mathbf{H} \hat{\mathbf{x}}_k \end{bmatrix} + \mathbf{b}_k \quad (8)$$

式中, \mathbf{W}_k 和 \mathbf{b}_k 是分别是第 k 层的基础权重和偏置。进一步经过激活函数非线性变换得到:

$$\mathbf{z}_k = \rho_k(\mathbf{v}_k) \quad (9)$$

其中, \mathbf{z}_k 是经过激活函数非线性变换输出的中间估计信号, $\rho_k(\cdot)$ 是激活函数。然而传统激活函数(如 sigmoid、ReLU)的输出范围受限, 难以覆盖高阶调制星座点的动态范围, 导致在高阶调制场景下检测性能受限。为此, 以 sigS 函数^[18]为基础设计一种归一化多段激活函数 $\rho_k(\cdot)$, 其设计如下:

$$\rho_k(x) = \frac{L}{p_{norm}} + \frac{E}{p_{norm}} \sum_{t=1}^{2l-1} \{p_{norm} \mathbf{g}_t a \mathbf{v}_k + \mathbf{h}_t [5 + (-1)^t (10t - 5)]\} \quad (10)$$

为了使激活函数具有更好的泛化能力和灵活性, 进行了功率归一化。其中 L 为不同调制方式下星座点符号集合的最小值, E 为相邻符号间的最小欧氏距离, 星座点集的总数为 $2l$, \mathbf{g}_t 和 \mathbf{h}_t 为一组可训练参数并将初始值设为 1, 确保激活函数原有的结构不发生变化, a 代表斜率, 而 p_{norm} 用于归一化, 具体如下:

$$p_{norm} = 2 \times \frac{1}{M} \sum_{j=1}^M c_j^2 \quad (11)$$

其中, c_j 表示 M -QAM 星座点, M 是星座点总数。

为解决迭代过程中存在的估计误差累积问题和信道噪声干扰问题, 在激活函数非线性变换后引入去噪器模块, 通过将中间估计信号与总噪声方差作为输入传递给去噪器进行去噪处理:

$$\hat{\mathbf{x}}_{k+1} = \eta(\mathbf{z}_k; \sigma_k^2) \quad (12)$$

其中, $\eta(\cdot)$ 是去噪器, \mathbf{z}_k 是中间估计信号, σ_k^2 代表总噪声方差, 具体如下:

$$\hat{\mathbf{x}}_{k+1} = \mathbb{E}[\mathbf{x} | \mathbf{z}_k, \sigma_k^2] \quad (13)$$

其中, $E[\cdot]$ 是求条件期望, 对式 (13) 进一步展开得到:

$$\hat{\mathbf{x}}_{k+1} = \sum_{j=1}^M \frac{\exp(-(\mathbf{z}_{k,j} - \mathbf{c}_j)^2 / 2\sigma_k^2)}{\sum_{k=1}^M \exp(-(\mathbf{z}_{k,j} - \mathbf{c}_k)^2 / 2\sigma_k^2)} \mathbf{c}_j \quad (14)$$

其中, \mathbf{c}_j 是当前直接参与分子计算的星座点, \mathbf{c}_k 是分母中用于归一化的所有星座点的集合, 确保权重转化为概率分布。通过式 (14) 可以看出, 去噪器是对 \mathbf{z}_k 的每个元素 $\mathbf{z}_{k,j}$ 计算与星座点 \mathbf{c}_j 的差值, 平方后并归一化得到高斯似然项, 再用 softmax 式转化为概率分布, 最后将星座点与概率分布加权求和输出该次迭代的最终估计信号 $\hat{\mathbf{x}}_{k+1}$ 。

上式中总噪声方差 σ_k^2 的公式定义如式 (15) 所示。

$$\sigma_k^2 = \left\{ \frac{\theta_k}{N_t} \frac{\|\mathbf{I} - \mathbf{H}^H \mathbf{H}\|_F^2}{\|\mathbf{H}\|_F^2} [\|\mathbf{y} - \mathbf{H}\hat{\mathbf{x}}_k\|_2^2 - N_t \sigma^2] + \frac{\|\mathbf{H}^H\|_F^2}{\|\mathbf{H}\|_F^2} \sigma^2 \right\} \quad (15)$$

其中, σ^2 为信道噪声方差, θ_k 是可学习参数。由总噪声方差公式可以看出, 去噪器输入处的噪声由两部分组成: 估计 $\hat{\mathbf{x}}_k$ 与 \mathbf{x} 的真实值偏差引起的误差以及信道噪声带来的影响。第 1 个分量通过线性变换 $(\mathbf{I} - \mathbf{H}^H \mathbf{H})$ 放大, 第 2 个分量通过 \mathbf{H}^H 放大。

此外, 为避免出现梯度消失和爆炸现象, 以及在网络深度增加时准确率下降的问题, 在去噪器完成对误差和噪声的处理后引入残差神经网络 (residual neural network, ResNet)^[19] 中的残差结构, 以改善网络单元之间的信息传输。残差结构的表达形式如下:

$$\hat{\mathbf{x}}_{in,k+1} = \mu \mathbf{x}_k^{out} + (1 - \mu) \hat{\mathbf{x}}_k^{in} \quad (16)$$

式中: μ 为残差系数, $\hat{\mathbf{x}}_k^{in}$, $\hat{\mathbf{x}}_k^{out}$ 分别用于表示第 k 个单元的输入向量与输出估计向量。从上述公式可以看出, 残差结构的本质是在当前单元中, 通过将上一单元的输入与输出分别乘以残差系数并加权融合, 生成新的输入。

最后 $\hat{\mathbf{x}}_k$ 的初始值设定为 0。MDNet 的损失函数通过加权均方误差 (mean squared error, MSE) 衡量预测结果

$\hat{\mathbf{x}}_k$ 与真实值 \mathbf{x} 的偏差, 并结合 $\tilde{\mathbf{x}}$ 进行归一化:

$$Loss(\mathbf{x}; \hat{\mathbf{x}}_k) = \sum_{k=1}^K \log(k) \frac{\|\mathbf{x} - \hat{\mathbf{x}}_k\|_2^2}{\|\mathbf{x} - \tilde{\mathbf{x}}\|_2^2} \quad (17)$$

式中:

$$\tilde{\mathbf{x}} = (\mathbf{H}^H \mathbf{H})^{-1} \mathbf{H}^H \mathbf{y} \quad (18)$$

3 实验仿真与性能评估

3.1 实验环境及相关参数配置

实验基于 PyTorch 深度学习框架完成, 结合式 (1) 描述的模型通过代码生成仿真数据集。在该模型中有 3 个关键随机变量: 信号 \mathbf{x}_c , 信道矩阵 \mathbf{H}_c 和信道噪声 \mathbf{n}_c 。传输信号 \mathbf{x}_c 从归一化的星座点集合中均匀随机生成, 信道矩

阵 \mathbf{H}_c 选用时变瑞利衰落模型, 并满足 $N(0, 1)$ 的独立同分布 (i. i. d) 条件^[20], 信道噪声 \mathbf{n}_c 的每个元素从零均值、(i. i. d) 的复高斯分布中采样, 噪声功率 σ^2 根据目标信噪比 (signal-to-noise ratio, SNR) 动态调整。其中每一组 \mathbf{y} 、 \mathbf{H} 、 \mathbf{x} 被视为一组训练样本, 在训练阶段共生成约 1 000 万组样本, 测试阶段生成约 500 万组样本, 训练与测试数据通过不同批次在线生成方式构造, 逻辑上保持独立。在训练过程中, SNR 在给定区间内按整数均匀随机采样, 从而使训练集中各 SNR 对应的样本数量近似均匀分布; 而测试阶段则在每一个固定的 SNR 点上均生成相同数量的测试样本, 从而保证在不 SNR 水平下的性能评估具有可比性。此外, 在梯度下降过程中, 通过 Adam 优化器对参数进行更新, 实验中相关参数具体设定如表 1 所示。

表 1 实验相关参数设置

Table 1 Experimental parameters settings

参数	值
调制方式	QPSK, 16QAM
初始学习率 lr	0.001
衰减率 (decay_rate)	0.97
衰减步长 (decay_steps)	500
批处理大小 (batch_size)	2 000
训练轮数 (Epoch)	5 000

3.2 误码性能评估

误符号率 (symbol error rate, SER) 是评价 MIMO 信号检测算法检测性能的重要指标。图 3 为大规模 MIMO 系统中发送端天线数 32 接收端天线数 64 的配置下, QPSK 调制方式下多种检测算法的 SER 随信噪比变化的性能曲线。所对比的算法包括传统的零强迫 (ZF) 与最小均方误差方法 (MMSE), 以及 3 种基于深度学习的检测算法: DetNet, ScNet 与所提出的 MDNet。同时引入最大似然检测 (ML) 作为性能上界参考。

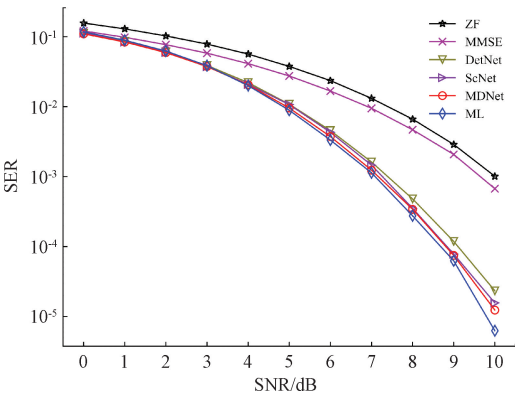


图 3 $N_t=32$, $N_r=64$ 时各算法误码性能比较

Fig. 3 Error performance comparison of algorithms when $N_t=32$, $N_r=64$

通过试验可以看出,随着 SNR 的升高,各检测方法的 SER 均呈现下降趋势。其中,经典检测方法 ZF 与 MMSE 在整个 SNR 区域内性能较弱,检测精度提升受限。相比之下,基于神经网络的检测方法表现出更优的鲁棒性与拟合能力,明显优于经典检测方法。而 MDNet 的 SER 曲线在全程优于 DetNet 和 ScNet,并在 10 dB 时 SER 接近 10^{-5} ,逼近最优的 ML 检测性能,实验结果验证了所提出算法的有效性。

为了验证提出的算法在不同调制方式下的误码性能,图 4 在发射天线数 32,接受天线数 64,不同调制方式下(QPSK 和 16QAM)进行了各算法误码性能比较实验结果表明,所有检测算法在 16QAM 调制下的误码性能均显著

低于 QPSK 调制。这一现象可归因于 16QAM 调制方式在相同带宽条件下传输更高阶的符号信息,使得信号星座点间的欧氏距离减小,从而对信道噪声和干扰更为敏感,显著增加了信号检测的复杂度。值得注意的是,在两种调制方式下,基于深度学习的检测方法均展现出优于传统算法的误码性能。这一优势主要源于神经网络强大的非线性建模能力。特别地,所提出的 MDNet 在低阶(QPSK)和高阶(16QAM)调制下,其性能均最接近最优 ML 检测算法,且在 16QAM 调制下表现出明显的性能优势。这一优异性能得益于所设计的归一化多段激活函数,该函数有效增强了网络的非线性映射能力,使其能够更好地适应高阶调制下更为复杂的信号检测需求。

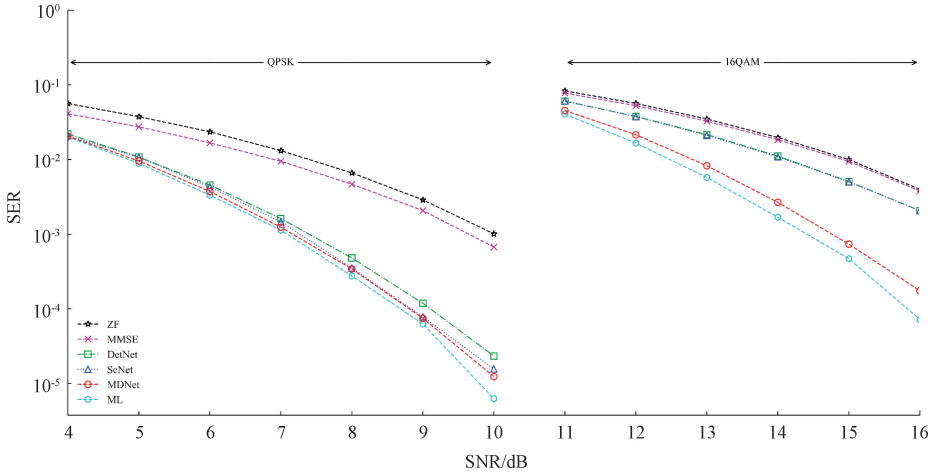


图 4 不同调制方式下各算法误码性能比较

Fig. 4 BER comparison of detection algorithms across modulation types

为验证网络结构中激活函数与去噪器的排列顺序对信号检测的影响,图 5 在 16QAM 调制方式、发射天线数为 32、接收天线数为 64 的设定下,进行了 3 组结构差异化的消融实验,并绘制了对应的 SER 性能曲线。

噪声器的结构(MDNet-no-denoise);先进行去噪处理后接激活函数的结构(MDNet-denoise-act);先进行激活操作后接去噪器的结构(MDNet-act-denoise)。从实验结果可观察到,先进行激活操作后接去噪器的网络结构在整个 SNR 区间内均展现出最低的 SER,特别是在 SNR 为 21 dB 时, SER 降至 10^{-8} ,相较其他两种结构提升了 1~2 个数量级。上述结果表明,将激活函数置于去噪器之前有助于增强网络的非线性表达能力,从而提升误码性能。因此,MDNet 最终选用该结构作为其核心组成。

鉴于激活函数在网络中的关键作用,图 6 在 16QAM 调制方式,64×32 的天线数配置下,比较了不同激活函数对 MDNet 的误码性能影响。对比的 3 种激活函数分别为: $\phi_k(\cdot)$ 激活函数^[14],sigS 激活函数,以及采用所提出的归一化多段激活函数 $\rho_k(\cdot)$ 。

从实验结果可以看出,3 种激活函数在 SNR 为 9 dB 表现相近, SER 皆维持在 10^{-1} 左右,差异尚不明显。但随着 SNR 提升,误码性能差异迅速拉开。在 SNR 为 15 dB, MDNet- $\rho_k(\cdot)$ 的 SER 已下降至 10^{-3} ,相比其他两种激活函数已经可以观察到明显优势。而在 SNR 为 21 dB, MDNet- $\rho_k(\cdot)$ 的 SER 更是突破至 10^{-8} 数量级,而其他

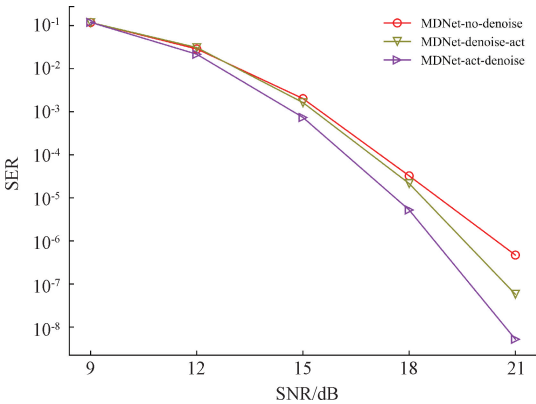
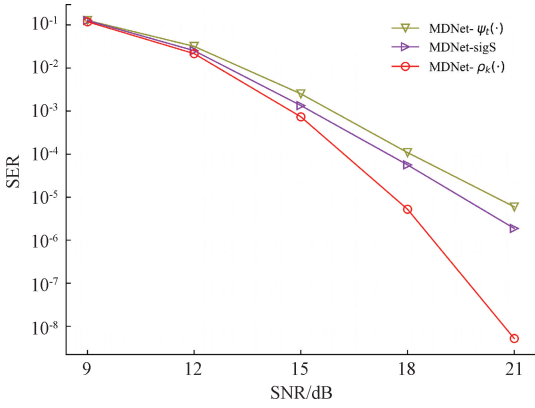


图 5 $N_t=32$, $N_r=64$ 时不同网络结构对 SER 的影响

Fig. 5 The impact of different network architectures on SER when $N_t=32$, $N_r=64$

先后测试了 3 种网络结构:仅包含激活函数、不含去

图 6 $N_t=32$, $N_r=64$ 时不同激活函数对 SER 的影响Fig. 6 The impact of different activation functions on SER when $N_t=32$, $N_r=64$

两种激活函数的检测误差仍停留在 10^{-6} 左右,说明所提出的归一化多段激活函数在高阶调制下具有更好的误差收敛性与映射精度。

MIMO 的核心优势在于多天线架构能够实现空间复用与分集,显著改善系统性能^[21],因此对 MIMO 空间复用的应用场景进行了性能评估。

图 7 为 16QAM 调制方式下,发射天线数 16,接收天线数分别为 16、24、32 以及 64 的 4 种不同检测曲线的对比。结果表明,当接收天线数量上升时误码性能逐渐提升并在 N_r 为 64 时 SER 达到 10^{-9} 。值得注意的是,更多的接收天线在相同的误码率下需要的信噪比更低,这是由于接收端获得的高分集增益,对于提高系统性能贡献大于高信噪比。因此,对于提高系统的可靠性,适当增加接收天线数是一种有效措施。

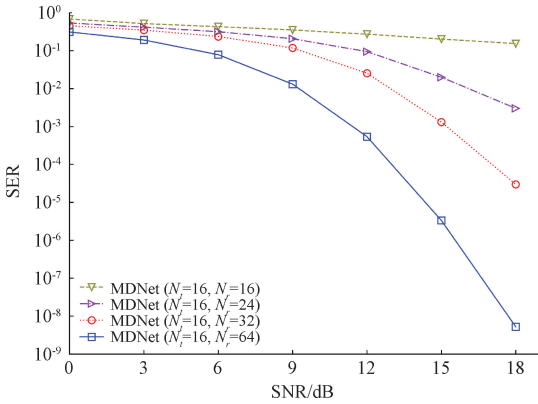


图 7 不同接收天线数量对 SER 的影响

Fig. 7 Impact of receive antenna number on SER

基于网络收敛速度与稳定性分析,图 8 是在 QPSK 调制方式以及 64×32 的天线数目下,所提出的 MDNet 模型与 ScNet 在训练过程中的损失函数 (Loss) 随训练轮数 (Epoch) 变化的演化轨迹。横轴为训练轮数,纵轴为损失值,采用对数尺度以增强对模型收敛特性的解析能力。从

结果可见,MDNet 在训练初期即展现出显著的收敛速率,仅经过 1 000 Epoch 即将损失值压降至 10^{-1} 以下,随后维持平稳下降并在约 2 000 Epoch 后趋于收敛,整体过程波动幅度较小,体现出优良的训练稳定性与优化效率。相比之下,ScNet 虽同样具备初始快速下降的趋势,但在整体收敛过程中表现出更高的损失值,且后期曲线波动较为明显。综上所述,MDNet 在训练过程中展现出更低的损失值与良好的稳定性。

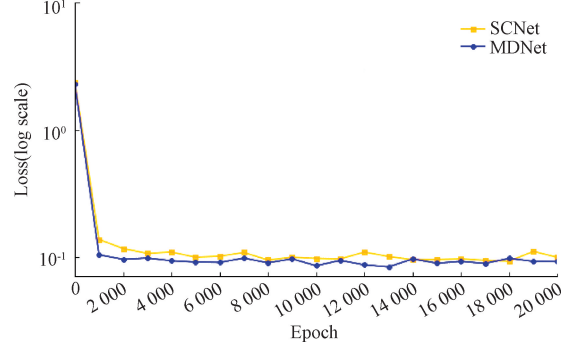


图 8 网络收敛速度与稳定性对比

Fig. 8 Comparison of network convergence speed and stability

3.3 复杂度分析

在基于深度学习的 MIMO 信号检测算法中,通常以实乘次数作为衡量其计算复杂度的标准,实数乘法数量与系统中的天线和符号数量相关。考虑到实数加法的线性特性,其对计算复杂度影响不大。在基于 MMSE 的检测策略中,计算量主要来自滤波矩阵的搭建与匹配滤波过程。而在深度学习检测方法中,计算量集中在各层所执行的矩阵乘法上。PyTorch 以张量 (Tensor) 为基础进行运算,其乘法次数由相关矩阵的维度决定。为更直观地评估模型,可进一步引入 FLOPs (floating point operations) 指标,用于衡量网络在推理过程中的实际浮点运算次数,从而更准确地反映网络的计算资源需求。

首先在 MDNet 中,复杂度分为 4 部分计算。第 1 部分为输入数据预处理时所做的乘法,其中 $\mathbf{H}^H \mathbf{y}$ 、 $\mathbf{H}^H \mathbf{H}$ 和 $\mathbf{H}^H \mathbf{H} \hat{\mathbf{x}}_k$ 分别作了 $4N_t N_r$ 、 $8N_r N_r^2$ 以及 $4N_r^2$ 次乘法运算;第 2 部分是中间估计信号 \mathbf{z}_k 的计算复杂度,预处理的输出数据总拼接维度为 $6N_t$,则该部分计算复杂度为 $6DN_t$, D 表示线性变换权重矩阵 \mathbf{W}_k 的输出维度;第 3 部分是去噪器中的乘法运算,去噪器的本质是将输入 \mathbf{z}_k 作为先验估计值,并使用高斯似然加 Softmax 归一化,对星座点集合做加权平均。一次迭代中高斯似然项计算涉及一次平方加一次除法加一次指数运算,算作常数阶,下一步分母求和并 Softmax 不涉及乘法操作,最后加权求和本质是对星座点集的向量点积操作,因此复杂度为 $O(M)$, M 为星座点数;第四部分执行的是残差特征的乘法操作,在每次迭代中涉及的乘法次数为: $4N_t$ 。最后,MDNet 网络结构中各模块所涉及的乘法运算次数可进行汇总,得出其所需的总

乘法次数为: $(8N_rN_t^2 + 4N_rN_t) + (4N_t^2 + QN_t + M)K$, $Q = 6D + 4, K$ 为 MDNet 迭代次数。其算法复杂度为 $O(N_rN_t^2)$ 。不同检测算法复杂度具体如表 2 所示。分析可得,传统检测算法 MMSE、ZF 矩阵求逆操作带来了较大的计算负担,其复杂度上升至 $O(N_t^3)$,MDNet 等基于深度学习的检测算法不涉及矩阵求逆,相较于 MMSE 降低了一个数量级的复杂度并且 MDNet 的计算复杂度随迭代次数呈线性增长,表明该迭代结构在计算资源消耗方面具有良好的可控性。进一步地,通过统计实际模型训练过程中的 FLOPs 指标,可以更直观地评估不同网络的运算资源需求。在 16QAM 调制、64 接收天线与 32 发射天线的环境下,对 DetNet、ScNet 与 MDNet 三种基于深度学习检测网络的 FLOPs 进行对比分析,分别为 50.58 GMac、2.38 GMac 及 202.56 GMac。尽管 MDNet 在计算量上明显高于其余两种方法,但得益于其可控的迭代结构设计与更强的非线性特征提取能力,在高阶调制下误码性能的提升足以弥补 FLOPs 方面开销过大的不足,从而在误码性能和复杂度之间取得了平衡。

表 2 检测算法复杂度对比

检测算法	复杂度
ZF	$8N_t^3 + 4N_t^2(1 + 2N_r) + 4N_tN_r$
MMSE	$8N_t^3 + 4N_t^2(1 + 2N_r) + 4N_tN_r$
DetNet	$8N_rN_t^2 + 4N_rN_t + 32N_t^2K$
ScNet	$8N_rN_t^2 + 4N_rN_t + 16N_t^2K$
MDNet	$(8N_rN_t^2 + 4N_rN_t) + (4N_t^2 + QN_t + M)K$

4 结 论

本文针对高阶调制场景下现有大规模 MIMO 检测算法误码性能受限问题,基于投影梯度下降法近似求解 ML 提出了 MDNet 检测网络。该网络设计了先激活函数后去噪器的优化结构,通过去噪器的引入,有效抑制了估计误差积累与信道噪声的影响,提升了检测的误码性能。并在此基础上设计了适用于高阶调制的归一化多段激活函数,解决了现有检测方法在高阶调制时检测性能受限的问题,从而进一步提高了误码性能。实验仿真表明,与现有信号检测方案进行对比时,所提出的 MDNet 检测网络具有更好的误码性能、更快的收敛速度和稳定性。所构建的检测网络能够更有效地支持大规模 MIMO 系统中信号检测的高效实现。未来将继续研究如何通过网络稀疏化和自注意力机制等技术的应用降低算法复杂度,进一步提升算法的实用性。

参考文献

[1] LIN Q, SHEN H, ZHAO C. Learning linear MMSE precoder for uplink massive MIMO systems with one-

bit ADCs [J]. IEEE Wireless Communications Letters, 2022, 11(10): 2235-2239.

[2] 代涛,李正权,王舟明,等. 基于改进 Richardson 的大规模 MIMO 信号检测算法[J]. 电子测量与仪器学报, 2023, 37(2): 171-178.

DAI T, LI ZH Q, WANG ZH M, et al. Signal detection algorithm for large-scale MIMO based on improved Richardson method[J]. Journal of Electronic Measurement and Instrumentation, 2023, 37 (2): 171-178.

[3] LIU Y, QIAN Y, WANG L, et al. Rapidly trainable and shallow-compiled quantum approximate optimization algorithm for maximum likelihood detection[J]. Physics Letters A, 2025, 548: 130541.

[4] MOHAMMAD A, ALI M A. A modified ZF algorithm for signal detection in an underwater MIMO STBC-OFDM acoustic communication system[J]. Annals of Telecommunications, 2023, 78(7-8): 491-507.

[5] KUUMAR A, GAUR N, NANTHAAMORNPHONG A. Signal detection of M-MIMO-orthogonal time frequency space modulation using hybrid algorithms: ZFE + MMSE and ZFE + MF[J]. Results in Engineering, 2024, 24: 103311.

[6] ZHANG Z, LI Y, YAN X, et al. A low-complexity AMP detection algorithm with deep neural network for massive MIMO systems[J]. Digital Communications and Networks, 2024, 10(5): 1375-1386.

[7] 李苗苗,华才健,谢涛,等. 融合多尺度特征及注意力机制的食品图像识别[J]. 电子测量技术, 2024, 47(18): 164-171.

LI M M, HUA C J, XIE T, et al. Food image recognition integrating multi-scale features and attention mechanism [J]. Electronic Measurement Technology, 2024, 47(18): 164-171.

[8] 李冰锋,刘帅,杨艺. 基于改进的 Transformer 细粒度图像识别算法研究[J]. 电子测量技术, 2024, 47(2): 114-120.

LI B F, LIU SH, YANG Y. Research on fine-grained image recognition algorithm based on improved Transformer [J]. Electronic Measurement Technology, 2024, 47(2): 114-120.

[9] ALSHAMMARI A, ALZAIDII A S M, ALRUSAINI O, et al. Robust speech perception and classification-driven deep convolutional neural network with natural language processing [J]. Alexandria Engineering Journal, 2025, 123: 358-368.

[10] AHLAWAT H, AGGARWAL N, GUPTA D. Automatic speech recognition: A survey of deep

- learning techniques and approaches[J]. International Journal of Cognitive Computing in Engineering, 2025, 6: 201-237.
- [11] 周璇,易剑平. 基于优化 CBAM 改进 ResNet50 的异常行为识别方法[J]. 国外电子测量技术, 2024, 43(5): 36-41.
- ZHOU X, YI J P. Abnormal behavior recognition method based on optimized CBAM and improved ResNet50 [J]. Foreign Electronic Measurement Technology, 2024, 43(5): 36-41.
- [12] 杨彦侃,马鑫宇,郁林. 一种基于 WiFi 信道状态信息的课堂行为识别方法[J]. 电子测量技术, 2025, 48(3): 118-127.
- YANG Y K, MA X Y, YU L. A classroom behavior recognition method based on WiFi channel state information[J]. Electronic Measurement Technology, 2025, 48(3): 118-127.
- [13] SAMUEL N, DISKIN T, WIESEL A. Deep MIMO detection[C]. 2017 IEEE 18th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC). IEEE, 2017: 1-5.
- [14] GAO G, DONG C, NIU K. Sparsely connected neural network for massive MIMO detection[C]. 2018 IEEE 4th International Conference on Computer and Communications (ICCC). IEEE, 2018: 397-402.
- [15] 李素月,贾鹏,王安红. 改进的 DetNet 大规模 MIMO 检测器[J]. 电讯技术, 2023, 63(2): 220-225.
- LI S Y, JIA P, WANG AN H. Improved DetNet detector for large-scale MIMO[J]. Telecommunication Engineering, 2023, 63(2): 220-225.
- [16] JIN L, TU Y, YANG J, et al. Interference cancellation assisted enhanced sparsely connected neural network for signal detection in massive MIMO systems[J]. Physical Communication, 2024, 66: 102438.
- [17] KHANI M, ALIZADEH M, HOVYDIS J, et al. Adaptive neural signal detection for massive MIMO[J]. IEEE Transactions on Wireless Communications, 2020, 19(8): 5635-5648.
- [18] 王建国. 基于深度学习的 Massive MIMO 信号检测研究[D]. 哈尔滨: 哈尔滨工程大学, 2022.
- WANG J M. Research on Massive MIMO signal detection based on deep learning[D]. Harbin: Harbin Engineering University, 2022.
- [19] CHAKRAVARTY N, DUA M. Erlang spectrogram and residual network-based features for fake audio detection[J]. IETE Journal of Research, 2025, 71(4): 1134-1140.
- [20] JIIN F, LIU Q, LIU H, et al. A low complexity signal detection scheme based on improved Newton iteration for massive MIMO systems [J]. IEEE Communications Letters, 2019, 23(4): 748-751.
- [21] LEE H, CHOI H, KIM H, et al. Downlink channel reconstruction for spatial multiplexing in massive MIMO systems[J]. IEEE Transactions on Wireless Communications, 2021, 20(9): 6154-6166.

作者简介

康晓非(通信作者), 博士, 副教授, 主要研究方向为宽带无线通信中的信道估计和信号检测技术、室内无线定位技术、3D 目标检测、深度学习及其在上述研究方向中的应用。

E-mail: kangxiaofei@xust.edu.cn

井溢洋, 硕士研究生, 主要研究方向为无线通信。

郭含玉, 硕士研究生, 主要研究方向为室内无线定位。