

基于矩阵分块的大规模 MIMO 低复杂度检测算法

季雅惠

(东南大学, 江苏省南京市, 211100)

摘要: 本文提出了一种基于矩阵分块的大规模 MIMO 低复杂度检测算法。提出的多层迭代的块对角化算法(block diagonal with multiple-level iterations, BD-MLI)同时适用于单天线和多天线用户系统, 且在非理想传播条件中, 如相关信道和低配置比, 表现出鲁棒性。特别地, 本文中的算法针对单天线和多天线用户系统在非理想传播条件下, 分别在信噪比为 14dB 和 16dB 时取得 10^{-4} 的误码率表现。本文中的算法较传统的矩阵分块算法(matrix partition, MP), 有 47%的复杂度降低, 且在任何传播条件下都能收敛于最小均方误差检测。同时, 本文也设计了一种对应的基于超大规模集成电路的高效硬件架构。

关键词: 大规模 MIMO; 线性检测; 诺依曼级数; 矩阵分块; 鲁棒性

A Low-Complexity Massive MIMO Detection Algorithm Based on Matrix Partition

Yahui Ji

(Southeast University, Nanjing Jiangsu, 211100)

Abstract: In this paper, a low-complexity Massive MIMO uplink detection algorithm based on matrix partition is introduced. The proposed block diagonal with multiple-level iterations (BD-MLI) is suitable for both single and multiple-antenna user equipments (UEs) systems and shows robust bit-error-rate (BER) performance in poor propagation environments such as correlated channel and small antenna configuration ratio. Specifically, BD-MLI achieves $BER = 10^{-4}$ at signal-to-noise ratio (SNR) of 14 dB and 16 dB in non-ideal propagation environment in single and multiple-antenna UEs system, respectively. It outperforms conventional matrix partition (MP) method by 47% complexity reduction and consistent near MMSE detection performance. Also, an efficient VLSI architecture is designed for the proposed BD-MLI algorithm.

Key Word: Massive MIMO, linear detection, Neumann series, matrix partition, robust performance

本文涉及大规模 MIMO 系统上行链路信号检测环节, 属于无线通信技术领域。

1 研究背景

多输入多输出(Multiple-Input Multiple-Output, MIMO)技术是指在发射端和接收端分别使用多个发射天线和多个接收天线, 使信号通过发射端与接收端的多个天线进行传送和接收。长期研究表明, MIMO 无线传输技术可以在不增大额外频谱带宽的情况下, 大大提高传输链路的可靠性, 减少基站

的能量消耗, 并成倍地提高通信系统的容量。区别于第四代移动通信技术中的只能在基站配置 8 根天线的小规模 MIMO 技术, 大规模 MIMO 在基站配置数量较多的天线以同时服务于数量较少的移动终端, 被视为下一代移动通信的核心技术[1, 2]。

然而, 大规模 MIMO 的优势是以在基站端的急剧增加的计算复杂度为代价。考虑基站天线数为 N , 用户端天线数为 M 的大规模 MIMO 系统, 典型取值为 $N = 128, M = 16$, 定义配置比为 N/M 。以大规模 MIMO 系统的上行链路信号检测为例, 基于最小均方误差(minimum mean square error, MMSE)理论的线性检测方法中, 其主要复杂度来源于一个 $M \times M$ 阶矩阵的求逆。传统的矩阵精确

作者简介: 季雅惠, (1998-), 女, 本科生, E-mail: jyh@seu.edu.cn.

求逆方法，如 QR 分解法、Gauss 消元法、Cholesky 分解法等，其复杂度很高，在 $O(M^3)$ 数量级。为了克服高计算复杂度、高硬件代价的困难，矩阵近似求逆的思想被提出。矩阵近似求逆旨在以较低的复杂度来逼近矩阵精确求逆下的检测性能，其中较典型的为基于诺依曼级数的矩阵近似求逆算法 (Neumann series approximation, NSA)[3]。然而，NSA 在迭代次数较多时，与传统的矩阵精确求逆复杂度并无区别。且 NSA 仅在理想传播条件下，即高配置比和理想信道中，才具有较好的检测性能，在非理想传播条件下，即配置比较低和相关信道中，诺依曼级数收敛速度慢或几乎不收敛，导致检测性能变得非常差。为了解决这个问题，NSA 的改进算法被提出，如三对角阵近似算法 (tridiagonal matrix approximation, TMA)[4] 和矩阵分块算法 (matrix partition, MP)[5] 等。这些改进算法都取得了特定的优化，如 TMA 在一定程度上克服了相关信道带来的检测障碍，MP 能有效降低 MIMO 系统对高配置比的依赖且能降低 NSA 约 68% 的计算复杂度。但同时都有一定的局限性——TMA 在复杂度降低上没有取得突破，MP 在相关信道中检测性能急剧下降。因此，需要一种更高效的检测算法，在低复杂度的前提下，同时能克服相关信道、降低系统对高配置比的依赖。

此外，当前绝大部分文献的研究对象为各用户端仅配置单根天线的 MIMO 系统(以下简称单天线用户系统)[6, 7]，然而面临用户端配置多根天线的 MIMO 系统(以下简称多天线用户系统)，同一个用户端的天线之间由于不充分分离而导致的强相关性，使得大多数对应于单天线用户系统提出的检测算法在应用于多天线用户系统时，性能急剧下降甚至几乎失效[8]。因此，需要提高 MIMO 检测算法在多天线用户系统中的性能。

2 系统模型

2.1 上行链路模型

考虑大规模 MIMO 系统的上行链路检

测模型，其中基站端配置 N 根接收天线同时服务于 m 个用户，且每个用户配置 m_{UE} 根发射天线，则用户端的天线总数为 $M = m \times m_{UE}$ 。显然，当 $m_{UE} = 1$ 时为单天线用户系统，当 $m_{UE} \neq 1$ 时为多天线用户系统。发射向量记为 $\mathbf{s} = [\mathbf{s}_1 \dots \mathbf{s}_i \dots \mathbf{s}_m]^T$ ，其中 $\mathbf{s}_i = [s_{i1} \dots s_{ij} \dots s_{im_{UE}}]^T$ 且发射能量归一化到

$$E\{|s_{ij}|^2\} = 1。上行链路信道矩阵定义为$$

$\mathbf{H} = [\mathbf{H}_1 \dots \mathbf{H}_i \dots \mathbf{H}_m]$ ，其中 $\mathbf{H}_i = [\mathbf{H}_{i1} \dots \mathbf{H}_{ij} \dots \mathbf{H}_{im_{UE}}]$ ， \mathbf{H}_{ij} 是一个 N 维列向量，表示从第 i 个用户的第 j 根天线到基站各天线之间的信道状况。基于以上定义，基站端的接收向量 \mathbf{y} 可表示为

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

其中， \mathbf{n} 是一个 N 维列向量，代表加性高斯白噪声，且每个元素服从均值为零，方差为 N_0 的高斯分布。

关于信道矩阵，由于信道之间的空间相关性会很大程度上影响算法检测性能，多数近期文献也不再仅基于理想信道来分析，而是引入经典的相关 MIMO 信道模型——Kronecker 模型[9]：

$$\mathbf{H} = \mathbf{R}^{1/2} \mathbf{H}_0 \mathbf{T}^{1/2} \quad (2)$$

其中 $\mathbf{H}_0 \in \mathcal{C}^{N \times M}$ 是理想信道，即瑞利衰落信道，每个元素都是独立同分布的高斯变量。 $\mathbf{R} \in \mathcal{C}^{N \times N}$ 和 $\mathbf{T} \in \mathcal{C}^{M \times M}$ 分别为表示基站端和用户端的空间相关性的矩阵， \mathbf{R} 和 \mathbf{T} 的元素定义为

$$R_{pq} = \begin{cases} (\zeta_r e^{j\theta})^{q-p} & p \leq q \\ R_{qp}^* & p > q \end{cases} \quad T_{pq} = \begin{cases} (\zeta_t e^{j\theta})^{q-p} & p \leq q \\ T_{qp}^* & p > q \end{cases} \quad (3)$$

其中相关系数 ζ_r 和 ζ_t 分别表示基站端和用户端的第 p 根天线与第 q 根天线之间的相关程度。有文献指出来自同一个用户端的发射天线之间因为不充分分离，而具有很强的相关性[8]。但近期很多以多天线用户系统为模型的检测算法并没有考虑这一点，这很大程度上偏离实际[6, 7]，因此在本文中，将所有用户的不同天线之间的相关系数统一记为 ζ_t ，同时为了合理简化模型，忽略不同用户端天

线之间的相关性。而对于典型的的城市环境，基站一般建于比较高的位置，由于缺少局部散射体的包围，导致基站天线之间较显著的相关性，因此将基站各天线之间的相关系数统一记为 ζ_r 。需要指出，相关系数 ζ_r 和 ζ_t 的取值范围为 0 到 1，且越接近于 1 表示相关程度越强。

2.2 MMSE 检测与诺依曼级数定理

阐述完本文的系统模型后，进一步介绍信号检测理论。本文仍为基于 MMSE 理论的线性检测，对发射向量的估计为

$$\hat{\mathbf{s}} = \mathbf{W}^{-1} \mathbf{y}^{MF} \quad (4)$$

其中 $\mathbf{W} = \mathbf{H}^H \mathbf{H} + N_0 \mathbf{I}_M$ ， $\mathbf{y}^{MF} = \mathbf{H}^H \mathbf{y}$ 为基站端匹配滤波器的输出。可以看出，这里涉及 $M \times M$ 阶矩阵 \mathbf{W} 的求逆。

根据诺依曼级数定理[10]，如果矩阵 \mathbf{W} 存在一个正定可逆的相近矩阵 \mathbf{X} ，即

$$\lim_{n \rightarrow \infty} (\mathbf{I} - \mathbf{X}^{-1} \mathbf{W})^n = \mathbf{0} \quad (5)$$

那么 \mathbf{W}^{-1} 可以用诺依曼级数展开为

$$\mathbf{W}^{-1} = \sum_{n=0}^{\infty} (\mathbf{I} - \mathbf{X}^{-1} \mathbf{W})^n \mathbf{X}^{-1} \quad (6)$$

在实际应用中，取前 k 项作为 \mathbf{W}^{-1} 的近似，并令 $\mathbf{D} = \mathbf{X}$ 和 $\mathbf{E} = \mathbf{W} - \mathbf{X}$ ，可以得到 \mathbf{W}^{-1} 的近似表达式为

$$\mathbf{W}^{-1} \approx \sum_{n=0}^{k-1} (-\mathbf{D}^{-1} \mathbf{E})^n \mathbf{D}^{-1} \quad (7)$$

3 多层迭代的块对角化算法

3.1 矩阵分块

由于在 MIMO 系统中矩阵 \mathbf{W} 呈现对角占优的特性，因此对于相近矩阵 \mathbf{D} ，NSA 以矩阵 \mathbf{W} 的主对角线元素构造 \mathbf{D} ，TMA 以矩阵 \mathbf{W} 的三对角线元素构造矩阵 \mathbf{D} ，MP 以矩阵 \mathbf{W} 的块对角线元素构造 \mathbf{D} 。而在本文中，继续沿用 MP 的构造方式，即保留矩阵 \mathbf{W} 的块对角线元素。为了合理简化算法内容，令

$$\mathbf{D} = \begin{bmatrix} \mathbf{D}_1 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_2 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{D}_K \end{bmatrix} \quad (8)$$

其中 $\mathbf{D}_i (i = 1, 2, \dots, K)$ 均为 $O \times O$ 阶的分块子矩阵，显然有约束条件 $O \times K = M, K \in \mathbb{Z}^+, O \in \mathbb{Z}^+$ 。

可以看出，参数 K 可有不同的取值，且取值可能会影响算法性能。为了使本文具有说服力，以下从理论与仿真两点来说明最优参数 K 的选取。首先从理论上，考虑算法的误码率(bit-error-rate, BER)性能，有相关文献指出，当 \mathbf{D} 包含 \mathbf{W} 的尽可能多的元素时，算法的 BER 性能会有很大提升，显然当 $K = 2$ 时， \mathbf{D} 包含 \mathbf{W} 的元素最多；考虑算法的计算复杂度，可以证明当 $K = 2$ 时，算法的计算复杂度最低，这是因为以 $k = 3$ 为例，当采取 $K = 2$ 的分块方式后， $(-\mathbf{D}^{-1} \mathbf{E})^2 \mathbf{D}^{-1}$ 的迭代过程可以表示为

$$\begin{aligned} & (-\mathbf{D}^{-1} \mathbf{E})^2 \mathbf{D}^{-1} \\ &= - \begin{bmatrix} \mathbf{0} & \mathbf{D}_{1,(1)}^{-1} \mathbf{E}_{1,(1)} \\ \mathbf{D}_{2,(1)}^{-1} \mathbf{E}_{2,(1)} & \mathbf{0} \end{bmatrix}^2 \begin{bmatrix} \mathbf{D}_{1,(1)}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_{2,(1)}^{-1} \end{bmatrix} \\ &= - \begin{bmatrix} \mathbf{0} & \mathbf{P}_1 \\ \mathbf{P}_2 & \mathbf{0} \end{bmatrix}^2 \begin{bmatrix} \mathbf{Q}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_2 \end{bmatrix} \\ &= - \begin{bmatrix} \mathbf{0} & \mathbf{P}_1 \mathbf{Q}_2 \\ \mathbf{P}_2 \mathbf{Q}_1 & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{0} & \mathbf{P}_1 \\ \mathbf{P}_2 & \mathbf{0} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{P}_1 \mathbf{P}_2 \mathbf{Q}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_2 \mathbf{P}_1 \mathbf{Q}_2 \end{bmatrix} \end{aligned} \quad (9)$$

可以看出每次的迭代结果，即

$$\begin{bmatrix} \mathbf{0} & \mathbf{P}_1 \\ \mathbf{P}_2 & \mathbf{0} \end{bmatrix}, \begin{bmatrix} \mathbf{0} & \mathbf{P}_1 \mathbf{Q}_2 \\ \mathbf{P}_2 \mathbf{Q}_1 & \mathbf{0} \end{bmatrix}, \begin{bmatrix} \mathbf{P}_1 \mathbf{P}_2 \mathbf{Q}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_2 \mathbf{P}_1 \mathbf{Q}_2 \end{bmatrix} \text{ 都 有 一 半 的 元 素 为 零， 也 就 大 大 降 低 了 计 算 复 杂 度。 而 对 于 其 他 分 块 方 式， 即 } K \neq 2, \text{ 当 迭 代 次 数 大 于 等 于 2 时， 迭 代 过 程 变 成 两 个 满 阵 相 乘， 计 算 复 杂 度 会 因 此 升 高。 综 合 算 法 的 BER 性 能 及 计 算 复 杂 度 的 理 论 分 析， 可 知 } K = 2 \text{ 为 最 优 的 分 块 方 式。 为 了 验 证 理 论 观 点， 图 1 给 出 了 仿 真 验 证， 可 以 清 晰 看 出 } K = 2 \text{ 的 分 块 方 式 较 其 他 分 块 方 式 有 明 显 优 势， 即 BER 性 能 最 好 且 迭 代 次 数 较 少。}$$

将 $\mathbf{D} = \begin{bmatrix} \mathbf{D}_{1,(1)} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_{2,(1)} \end{bmatrix}$ ， $\mathbf{E} = \begin{bmatrix} \mathbf{0} & \mathbf{E}_{1,(1)} \\ \mathbf{E}_{2,(1)} & \mathbf{0} \end{bmatrix}$ 代入式(7)中可得

$$\begin{aligned} & \mathbf{W}^{-1} \approx \sum_{n=0}^{k-1} \left(- \begin{bmatrix} \mathbf{0} & \mathbf{D}_{1,(1)}^{-1} \mathbf{E}_{1,(1)} \\ \mathbf{D}_{2,(1)}^{-1} \mathbf{E}_{2,(1)} & \mathbf{0} \end{bmatrix} \right)^n \begin{bmatrix} \mathbf{D}_{1,(1)}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_{2,(1)}^{-1} \end{bmatrix} \end{aligned} \quad (10)$$

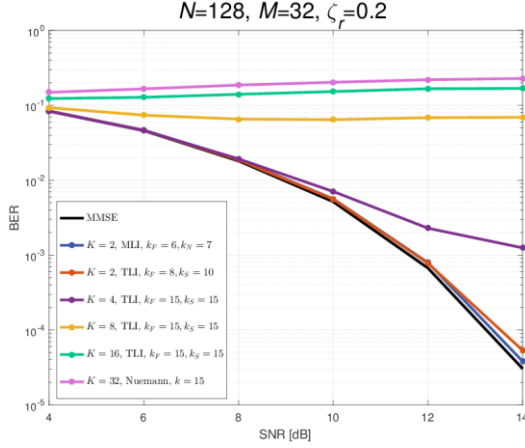


图 1 $K=2$ 的分块方式较其他的分块方式收敛性能对比图

3.2 多层迭代

在式(10)中, 子矩阵 $\mathbf{D}_{1,(1)}$ 和 $\mathbf{D}_{2,(1)}$ 的逆都是未知的, 为了解决这个问题, MP 通过 NSA 来求 $\mathbf{D}_{1,(1)}^{-1}$ 和 $\mathbf{D}_{2,(1)}^{-1}$ 的近似, 因此该算法对 \mathbf{W}^{-1} 的近似由两层迭代 (two-level iterations, TLI) 组成。区别于 TLI, 本文提出一种多层迭代 (multiple-level iterations, MLI) 的方法。考虑到 $\mathbf{D}_{1,(1)}$ 和 $\mathbf{D}_{2,(1)}$ 为 $M/2$ 阶矩阵, 以和第一层迭代相同的方式进行第二层迭代得到 $\mathbf{D}_{1,(1)}^{-1}$ 和 $\mathbf{D}_{2,(1)}^{-1}$, 以 $\mathbf{D}_{1,(1)}^{-1}$ 为例, 有

$$\mathbf{D}_{1,(1)}^{-1} \approx \sum_{n=0}^{k-1} \left(- \begin{bmatrix} \mathbf{0} & \mathbf{D}_{1,(2)}^{-1} \mathbf{E}_{1,(2)} \\ \mathbf{D}_{2,(2)}^{-1} \mathbf{E}_{2,(2)} & \mathbf{0} \end{bmatrix} \right)^n \begin{bmatrix} \mathbf{D}_{1,(2)}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_{2,(2)}^{-1} \end{bmatrix} \quad (11)$$

考虑到 $\mathbf{D}_{1,(2)}$ 和 $\mathbf{D}_{2,(2)}$ 为 $M/2^2$ 阶矩阵, 以和第一层迭代相同的方式进行第三层迭代得到 $\mathbf{D}_{1,(2)}^{-1}$ 和 $\mathbf{D}_{2,(2)}^{-1}$ 。以此类推, 直到最后一层。

在最后一层迭代中, 需要有不同的额外操作。在单天线用户系统中, 需要对 \mathbf{W} 主对角线的元素求倒数。而多天线系统中, 由于同一用户天线之间的强相干性使得 \mathbf{W} 呈块对角占优而不是主对角占优, 且理论上块对角占优的这些子矩阵的维度为 $m_{\text{UE}} \times m_{\text{UE}}$, 由于这些子矩阵更加接近于满阵, 因此不能继续采用矩阵近似求逆法, 而只能采用矩阵精确求逆法。虽然矩阵精确求逆的复杂度很高, 但由于该处子矩阵的规模 $m_{\text{UE}} \times m_{\text{UE}}$ 较小, 因为一般同一个用户端不会配置过多的天线, 子矩阵的个数也较少, 为 $m = M/m_{\text{UE}}$,

因此该部分的计算复杂度相对于整体的计算复杂度是可忽略的。

3.3 复杂度分析

与大多数 MIMO 检测算法的复杂度分析一样, 仅考虑算法中的复数乘法次数, 本文中的算法 BD-MLI 与另两种 NSA 优化算法 TMA 和 MP 的复杂度列于表 1 中。

表 1 TMA, MP, BD-MLI 的复杂度对比

算法	复数乘法次数
TMA	$(k-2)M^3$
MP	$k_F/4M^3 + (k_S-2)/4M^3$
BD-MLI	$k_F/4M^3 + k_N/12M^3$

对于 TMA, k 为算法的迭代次数;

对于 MP, k_F 为算法的第一层迭代次数, k_S 为算法的第二层迭代次数;

对于 BD-MLI, k_F 为算法的第一层迭代次数, k_N 为算法余下层数的迭代次数。

结合图 1 和表 1, 需要指出图 1 中的“ $K=2$, MLI”正是本文中的 BD-MLI 算法, 而“ $K=2$, TLI”正是上面提到的 MP 算法, 可以看出图 1 中 BD-MLI 与 MP 均取得接近 MMSE 的 BER 性能, 但通过表 1 中的复杂度公式可得, BD-MLI 较 MP 降低了约 47% 的复数乘法次数。

以上完整叙述了本文, 即多层迭代的块对角化算法 (block diagonal with multiple-level iterations, BD-MLI) 的内容。

4 仿真结果

图 2-7 给出了不同类型系统、不同传播条件下的仿真结果。图 2-4 分别为单天线用户系统在理想传播条件 (理想信道、高配置比), 非理想传播条件 (相关性较弱的信道、低配置比), 恶劣传播条件 (相关性较强的信道、低配置比) 下的仿真结果, 图 5-7 分别为多天线用户系统在理想传播条件, 非理想传播条件, 恶劣传播条件下的仿真结果, 均以误码率随信噪比 (signal-to-noise ratio, SNR) 的变化来表示算法性能。

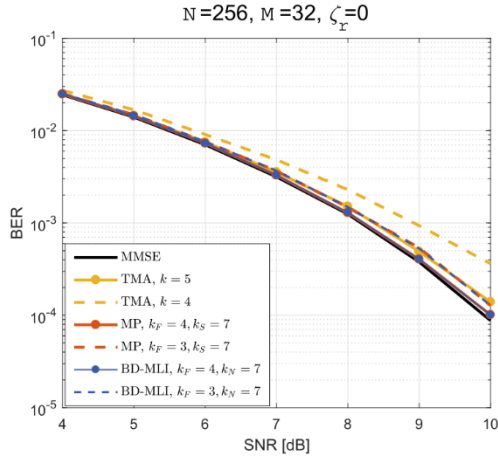


图 2 $N = 256, M = 32, \zeta_r = 0$ 的单天线用户系统中, BD-MLI 与 TMA、MP 误码率性能对比图。

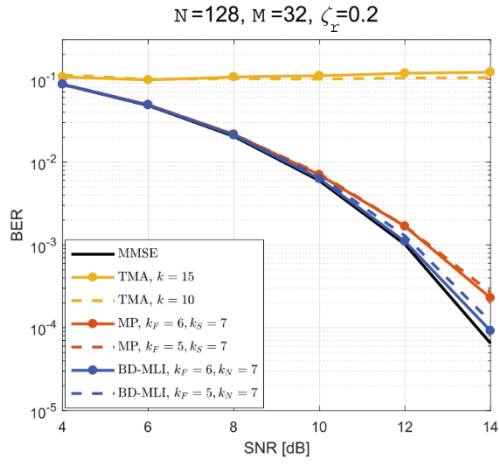


图 3 $N = 128, M = 32, \zeta_r = 0.2$ 的单天线用户系统中, BD-MLI 与 TMA、MP 误码率性能对比图。

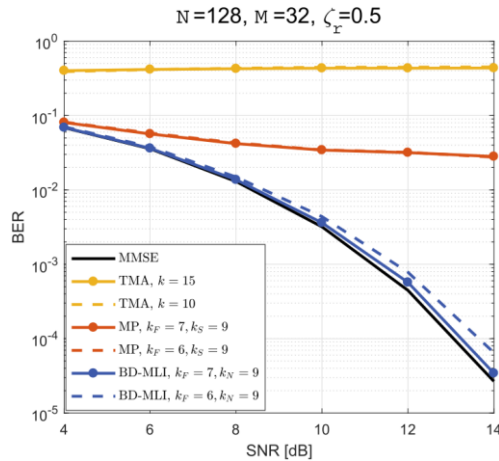


图 4 $N = 128, M = 32, \zeta_r = 0.5$ 的单天线用户系统中, BD-MLI 与 TMA、MP 误码率性能对比图。

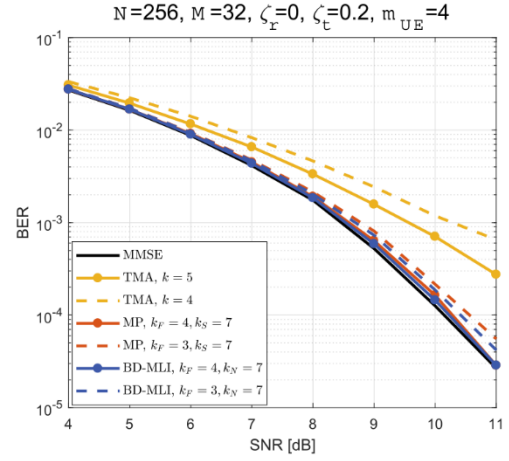


图 5 $N = 256, M = 32, m_{UE} = 4, \zeta_r = 0, \zeta_t = 0.2$ 的多天线用户系统中, BD-MLI 与 TMA、MP 误码率性能对比图。

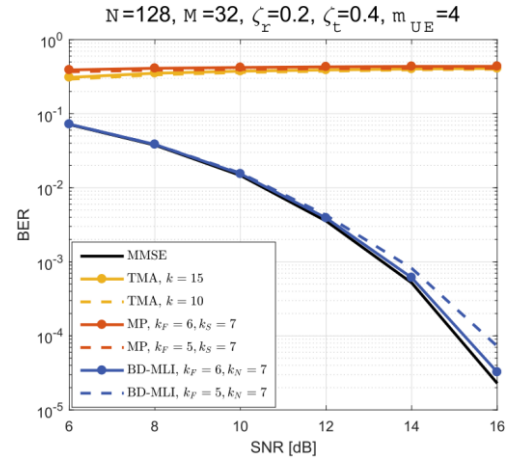


图 6 $N = 128, M = 32, m_{UE} = 4, \zeta_r = 0.2, \zeta_t = 0.4$ 的多天线用户系统中, BD-MLI 与 TMA、MP 误码率性能对比图。

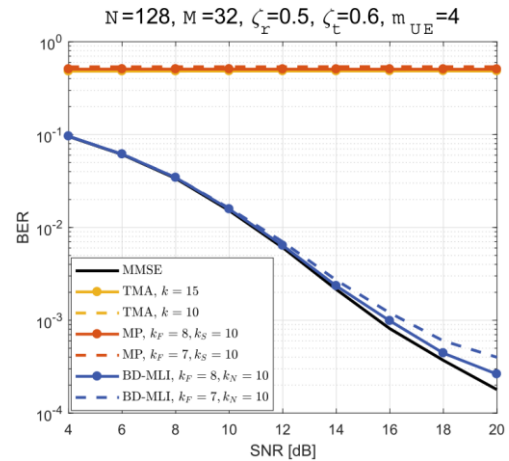


图 7 $N = 128, M = 32, m_{UE} = 4, \zeta_r = 0.5, \zeta_t = 0.6$ 的多天线用户系统中, BD-MLI 与 TMA、MP 误码率性能对比图。

可以看出, TMA 仅在理想传播条件下才能有较好的检测性能(如图 2、5 所示), 在非理想和恶劣传播环境中, 均已失效(如图 3、4、6、7 所示)。MP 在理想传播条件及单天线用户系统中的非理想传播条件能有较好的检测性能(如图 2、5、3 所示), 在恶劣传播条件的多天线用户系统中性能非常差(如图 5 所示), 在非理想及恶劣传播环境下的多天线用户系统中也已失效(如图 6、7 所示)。但是对于本文中的算法 BD-MLI, 无论是单天线还是多天线用户, 无论是理想、非理想、恶劣传播环境, 均能以较低的复杂度取得接近 MMSE 的检测性能。

由此可见, 本文方法提出了一种更高效的 MIMO 检测算法, 在低复杂度的前提下, 同时能克服相关信道、降低系统对高配置比的依赖, 并且提高算法在多天线用户系统中的性能。

5 硬件架构

最后为本文中的算法设计一套相应的硬件架构。为了方便叙述, 这里以 $m = 8$ 为例, 展开相应的设计过程。

根据式(4), 仅需要利用 \mathbf{W}^{-1} 和匹配滤波器的输出 \mathbf{y}^{MF} 就可得到发射向量的估计。由于关于计算 \mathbf{W} 和 \mathbf{y}^{MF} 的硬件架构是比较简单直接的, 因此我们仅关注计算 \mathbf{W}^{-1} 的硬件架构。对于本文中的算法, 总迭代层数可记为 $T = \log_2 m$, 可见这是一个与 N 和 m_{UE} 均无关的量。为了方便起见, 将第 $(T + 1 - t)$ 层迭代($t \in [1: T]$)中的子矩阵记为 $\mathbf{D}_{:,t}$ 和 $\mathbf{E}_{:,t}$ 。

图 8 给出了本文算法对用于 $m = 8$ 的硬件架构, 注意这里仅画出了对应于矩阵 $\mathbf{D}_{:,t}$ 的处理单元(processing elements, PE)。这套流水线架构进行了不同阶段的分化, 因而是结构良好的。图 8 中第 t 个阶段对应于算法中的第 $(T + 1 - t)$ 层迭代。在阶段 1 前, $\mathbf{D}_{i,(1)}^{-1} (i \in [1:m])$ 由 INV 模块求出, 这里需要说明, 在单天线用户系统中, INV 的作用为求倒数, 在多天线用户系统中, INV 的作用为进行一个 $m_{UE} \times m_{UE}$ 阶的矩阵精确求逆, 由于对于小规模矩阵的精确求逆有很多现存的高效硬件架构, 因此我们不再给出这

部分的具体细节。显然, 该硬件架构中的总阶段数与算法中的总迭代层数对于并且相等的。对于阶段 t , PE 的数目应为 2^{T-t} , 因此总的 PE 数目应为 $(2^T - 1)$ 。经过阶段 T 后, 可以获得矩阵 \mathbf{W} 的近似求逆。

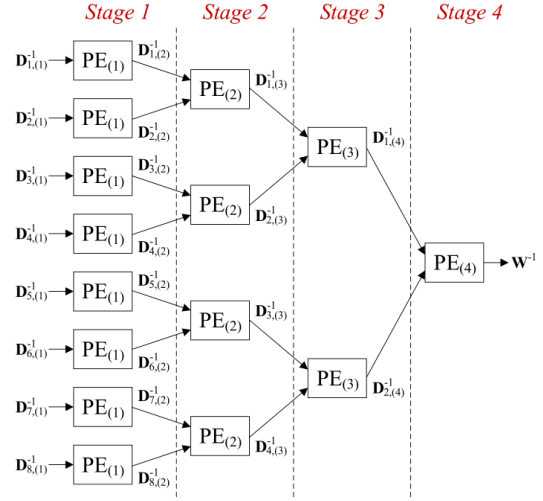


图 8 本文中的算法对于 $m = 8$ 的流水线架构

PE 的作用是在每层迭代中计算式(7)。对于不同的阶段, $\text{PE}_{(t)}$ 本质上是相似的, 不同之处仅在于输入输出矩阵的阶数。图 9 画出了 $\text{PE}_{(t)}$ 进行计算 $\mathbf{D}_{1,(t+1)}^{-1}$ 的架构细节, $\mathbf{D}_{1,(t+1)}^{-1}$ 对应于图 8 最上端的 $\text{PE}_{(t)}$ 部分。首先, $\text{PE}_{(t)}$ 将 $\mathbf{D}_{1,(t)}^{-1}$, $\mathbf{D}_{2,(t)}^{-1}$ 和 $\mathbf{E}_{1,(t)}$, $\mathbf{E}_{2,(t)}$ 作为输入并为下一步运行构建过渡矩阵(如式(10, 11)所示)。然后阶段 t 的诺依曼级数可以由矩阵加法器和矩阵乘法器求出。这里, 我们利用寄存器来节省硬件消耗。

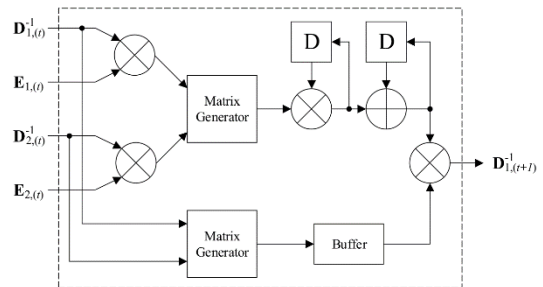


图 9 $\text{PE}_{(t)}$ 进行计算 $\mathbf{D}_{1,(t+1)}^{-1}$ 的架构细节

表 2 给出了改架构的硬件复杂度, 这里我们仅以复数加法器和复数乘法器的数量

表示硬件消耗。

表 2 硬件复杂度分析

	$PE_{(t)}$	BD-MLI
延时(时钟)	$4 \times 2^t m_{UE}$	$8 \times (2^T - 1) m_{UE}$
加法器	$3 \times 4^{t-1} m_{UE}^2$	$3 \times 2^{T-1} \times (2^T - 1) m_{UE}^2$
乘法器	$2 \times 4^{t-1} m_{UE}^2$	$2^T \times (2^T - 1) m_{UE}^2$

参考文献：

- [1] Larsson, Erik G., et al. "Massive MIMO for next generation wireless systems." *IEEE Communications Magazine* 52.2(2014):186-195.
- [2] Rusek, F., et al. "Scaling Up MIMO: Opportunities and Challenges with Very Large Arrays." *Signal Processing Magazine IEEE* 30.1(2012):40-60.
- [3] Wu, M., et al. "Approximate matrix inversion for high-throughput data detection in the large-scale MIMO uplink." *IEEE International Symposium on Circuits and Systems IEEE*, 2013:2155-2158.
- [4] Liang, X., et al. "On the low-complexity, hardware-friendly tridiagonal matrix inversion for correlated massive MIMO systems." *arXiv preprint arXiv:1802.10444*, 2018.
- [5] Wu, H., et al. "Low-complexity detection algorithms based on matrix partition for massive MIMO." *IEEE International Conference on Wireless Communication and Signal Processing IEEE*, 2017:1-6.
- [6] Wei, Hao , D. Wang , and X. You . "ICD reciprocity calibration for distributed large-scale MIMO systems with BD precoding." *IEEE/CIC International Conference on Communications in China-workshops IEEE*, 2017.
- [7] Chen, Chun An , et al. "A generalized eigenvalue decomposition processor for multi-user MIMO precoding." *Circuits & Systems IEEE*, 2017.
- [8] Spencer, Q. H. , A. L. Swindlehurst , and M. Haardt . "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels." *IEEE Transactions on Signal Processing* 52.2(2004):461-471.
- [9] Godana, Bruhtesfa E. , and T. Ekman . "Parametrization Based Limited Feedback Design for Correlated MIMO Channels Using New Statistical Models." *IEEE Transactions on Wireless Communications* 12.10(2013):5172-5184.
- [10] Stewart, and W. G. . "Matrix algorithms, v.1 : basic decompositions /."