Stair Matrix and its Applications to Massive MIMO Uplink Data Detection

Fan Jiang, Student Member, IEEE, Cheng Li, Senior Member, IEEE, Zijun Gong, Senior Member, IEEE, and Ruoyu Su, Student Member, IEEE

Abstract—In this paper, we propose a new approach of using the stair matrix for uplink data detection in massive MIMO systems. We first demonstrate the applicability of the proposed method by showing that the probability (that the convergence conditions are met) will approach one as long as sufficient large number of antennas are equipped at the base station. We then propose an iterative method to perform data detection and show that a much improved performance can be achieved with the computational complexity remaining at the same level of existing iterative methods where the diagonal matrix is adopted. Performance evaluation is conducted in terms of the probability that the convergence conditions are met, the normalized mean-square error of the Neumann series expansion to approach the matrix inverse, the residual estimation error to approach the linear ZF/MMSE detection, and the system bit error rate. Numerical simulations show significant performance enhancement of using the stair matrix over the diagonal matrix in all performance aspects.

Index Terms—Massive MIMO; Stair Matrix; Iterative Method; Convergence Condition.

I. INTRODUCTION

The development and successful applications of multiple-input multiple-output (MIMO) systems in modern wireless communications have brought the bright prospective of massive MIMO techniques in future 5G mobile communication systems [1]–[3]. It is foreseeable that massive MIMO, together with the millimeter wave frequency band [4], has been a promising candidate to meet the high rate, low latency 5G system requirements. Due to the huge potential multiplexing and diversity gain over the small-scale MIMO and single-antenna systems, massive MIMO can boom the system spectrum and energy efficiency [1]. [5]–[7]. Along with the benefits of massive MIMO, however, the cost of high computational complexity required in signal processing (data detection, precoding, etc.) increases, which prohibits the application of the optimal detection methods, such as the maximum likelihood (ML), and maximum a posteriori (MAP) detection, in realization.

To achieve a good tradeoff between the system performance and the computational complexity, linear detection (and precoding) methods, such as zero-forcing (ZF) and minimum mean-square error (MMSE), have been considered in realization [8]–[15]. It has also been demonstrated that with these linear detection methods, the near-optimal performance can be achieved in massive MIMO systems, especially when the number of antennas at base station (denoted by $N_B$) is much greater than the number of user equipment (denoted by $N_U$) in service. However, as we know, ZF/MMSE based data detection schemes experience matrix inversion, which is computational costly (almost $O(\frac{N^3}{r})$, where $N$ is the matrix size) in implementation. Therefore, the investigation of reducing computational complexity but still maintaining near-optimal system performance of ZF/MMSE based data detection schemes has emerged recently [8]–[15]. Generally, all those schemes can be summarized in two categories: the first one is to approach matrix inversion, and the other is to solve linear equations with iterative methods.

The first category is to approach the matrix inversion [8]–[10]. For example, in [8], the authors attempt to introduce Neumann series expansion to avoid the matrix inversion in linear MMSE detection. It has been shown that when the number of antennas at base station is much greater than the number of user equipment, the orders required for Neumann series expansion can be as few as 3 (for example, $r = \frac{N_B}{N_U} \geq 16$). In [9], the probability of the convergence condition that using the diagonal matrix in Neumann series expansion has been comprehensively discussed. However, Neumann series expansion suffers from matrix multiplications, and the computational complexity is comparable to the matrix inversion algorithm when the expansion order is more than two. In order to speed up the convergence rate, diagonal banded Newton iteration based matrix inversion approach is studied in [10], where the Newton iteration structure is used. Actually, the results after $P$ iterations in Newton iteration can be seen as the Neumann series expansion of the order $2^P - 1$ [10]. Inevitably, matrix multiplications are involved in diagonal banded Newton iteration based matrix inversion approach, and the iterations are limited to 2 for computational complexity consideration. In summary, the methods that are to approach matrix inversion suffer from high computational complexity due to the matrix multiplications and the slow convergence rate when the ratio $r$ is not sufficiently large.

The second category is to solve linear equations with iterative methods [11]–[15]. The basic idea of these methods is to transform the matrix inversion problem into solving linear equations. To solve the linear equations, an initial estimation is provided. Then following an iterative structure to converge, the final output is provided as the solutions to linear equations. For example, in [11], the Jacobi method is adopted, and by following the Jacobi iterative structure, the estimation eventually approaches the MMSE estimation. The Richardson
iteration in massive MIMO uplink data detection has been studied in [12], and the authors have demonstrated that the iterative structure can converge even with zero initialization. However, as pointed out in [13], [14], the convergence rate for both Jacobi method and Richardson iteration is slow, hence quite a few iterations are required for convergence. The application of Gauss-Seidel method to massive MIMO uplink data detection is studied in [13], and the convergence performance can be greatly improved. By providing an initial estimation that is close to the MMSE estimation, the joint steepest-decent and Jacobi method based data detection is proposed in [14], and the iterations are greatly reduced. In [15], the authors formulate the MMSE estimation as a minimization problem, and use the conjugate gradient to calibrate the next estimation. However, conjugate gradient-based data detection scheme involves many division operations, which is also computationally costly. Compared to the first category which is to approach matrix inversion, solving linear equations with iterative methods is of less complexity due to the replacement of matrix multiplications with matrix-vector products. However, as summarized in [14], the overall computational complexity of the iterative methods, including the computations in both the initialization and iteration, is still high. It is worth pointing out that the convergence rate of the existing iterative methods can be speeded up by using preconditioning [16]. The potential out that the convergence rate of the existing iterative methods, including the computations in both initialization and less iterations for convergence [17].

In both the previous mentioned two categories of data detection schemes, we note that most proposals in existing literatures mainly utilize the diagonal matrix in the development. In [8], [9], the applicability of using diagonal matrix to massive MIMO uplink data detection has been demonstrated. However, as we will show later, we find some limitations for using diagonal matrix. First of all, in the massive MIMO system configuration with small ratio of \( r = N_B/N_U \), the convergence rate of using the diagonal matrix is slow. Alternatively, a few iterations (or orders in Neumann series expansion) are required to provide near-optimal system performance. Besides, the convergence conditions, which are critical for the both data detection schemes mentioned above, are met with a low probability when \( r \) is small. That is to say, in some cases, the diagonal matrix may not be used to converge.

The motivation of this paper originates from achieving a better tradeoff between computational complexity and system performance in massive MIMO uplink data detection. We propose to use the stair matrix in the development. As far as we know, the applications of stair matrix in massive MIMO systems have not been studied. The contributions of this paper are summarized as follows:

- We show that when \( N_B \) grows to infinite, the probability that the convergence conditions are met approaches 1. As the antennas at base station in Massive MIMO systems can be hundreds, this conclusion demonstrates the applicability of the stair matrix in massive MIMO systems;
- We demonstrate the proposed iterative method with the use of the stair matrix has the same level of the computational complexity compared to the existing iterative methods where the diagonal matrix is applied;
- We show that by using the stair matrix, the probability that the convergence conditions are met can be greatly improved in a comparatively low \( r \) region, and the cumulative distribution function of the maximum eigenvalue of the convergence matrix indicates that the convergence rate can be speeded up by using the stair matrix;
- We demonstrate that by using the stair matrix, the mean-square error of the truncated Neumann series expansion to approach matrix inverse, can be greatly reduced;
- We show that the residual estimation error of the proposed iterative method using the stair matrix is much less than that of the Jacobi method where the diagonal matrix is applied;
- We compare the system BER performance with the proposed iterative method, and show that the performance improvement over the use of the diagonal matrix is significant.

The rest of this paper is organized as follows. Section II provides the system model, including the massive MIMO structure and the preliminary work of linear ZF/MMSE detection. In section III, the introduction to stair matrix and its applicability in massive MIMO will be presented. The implementation of stair matrix in massive MIMO data detection with iterative method is presented in section IV. In section V, we conduct the numerical simulations and present the results and discussion. Finally, the conclusions are drawn in section VI.

Notations: Throughout the paper, the lowercase and upercase bold symbols denote the column vector and matrix, respectively. \( \cdot^T \), \( \cdot^H \), and \( \cdot^{-1} \) are reserved for matrix transpose, conjugate transpose, and inverse, respectively. \( \mathbb{C} \) and \( \mathbb{N} \) are reserved for the sets of the complex and natural numbers, respectively. \( |A|_F \) and \( \|A\|_2 \) are the Frobenius-norm of a matrix \( A \) and the \( \ell_2 \)-norm of a vector \( a \). \( E \{ \cdot \} \) and \( \text{cov} \{ \cdot, \cdot \} \) denote the expectation, and covariance operation. \( \exp \{ \cdot \} \) and \( \ln \{ \cdot \} \) denote the exponential and natural logarithmic functions, respectively. \( \mathbf{I}_L \) is reserved for the size \( L \) identity matrix and \( e_i \) represents the \( i \)-th column of \( \mathbf{I}_L \); \( \text{diag} \{ a \} \) converts a column vector \( a \) to a diagonal matrix and \( \text{diag} \{ A \} \) obtains the diagonal elements in a matrix \( A \) to form a column vector. \( \rho(A) \) is the spectral radius of the matrix \( A \).

II. SYSTEM MODEL

We consider the massive MIMO uplink with \( N_B \) antennas at base station to simultaneously serve \( N_U \) single-antenna user equipment. The \( N_B \) bitstream from each user is first encoded, then interleaved, and fed into digital modulator. The modulated symbols are transmitted into massive MIMO channel, and the received signal vector at base station can be expressed as

\[
y = \mathbf{H}x + z,
\]

where \( y = [y_1, y_2, \cdots, y_{N_B}]^T \) is a complex-valued \( N_B \times 1 \) vector, with \( y_m \) denoting the received signal from the \( m \)-th receiving antenna. \( x = [x_1, x_2, \cdots, x_{N_U}]^T \) with the transmitted symbol of user \( u \) denoted by \( x_u \). \( \mathbf{H} = [h_{11}, h_{12}, \cdots, h_{N_U}] \) denotes the channel matrix with \( h_{uu} \in \mathbb{C}^{N_B \times 1} \) where each entry is independent and identically distributed (i.i.d.), modeled as the flat
Rayleigh fading channel \([1], [5], [13]\). \(z = [z_1, z_2, \cdots, z_{N_B}]^T\) is the noise vector, satisfying \(E\{zz^H\} = \sigma_z^2 I_{N_B}\) with each entry modeled as zero-mean complex Gaussian circularly symmetric (ZMCGS) random variable. It is worth noting that in frequency selective fading channels, by applying the OFDM/SC-FDMA techniques, the signal model expressed in \([1]\) is applied to each subcarrier.

\[ x = (H^HH + \sigma_z^2 I_{N_B})^{-1}H^Hy = W^{-1}y^{MF}, \quad (2) \]

where \(y^{MF} = H^Hy\) can be seen as the matched-filter output, and the MMSE equalization matrix \(W\) can be expressed as

\[ W = G + \sigma_z^2 I_{N_B}, \quad (3) \]

where \(G = H^HH\) is the Gram matrix. It is worth noting that in high signal-to-noise ratio (SNR) region, Equation (2) can be reduced to

\[ \hat{x} = G^{-1}y^{MF}, \quad (4) \]

which is the linear ZF data detection scheme, where the noise component is not considered in the equalization process.

To obtain the a posteriori LLR of the bits associated with the modulated symbols, we write the estimation in Equation (2) as

\[ \hat{x}_u = e_u^H\hat{x} = \rho_u x_u + \xi_u, \quad (5) \]

where the equivalent channel gain \(\rho_u\) and the a posteriori noise-plus-interference (NPI) \(\xi_u\) can be given by

\[ \rho_u = e_u^H W^{-2} G e_u, \quad (6) \]

\[ \xi_u = e_u^H W^{-2} G (x - x_u e_u) + e_u^H W^{-1} H^Hz. \quad (7) \]

The covariance of the NPI \(\nu_u^2 = \text{cov}(\xi_u, \xi_u)\) is given by

\[ \nu_u^2 = e_u^H W^{-2} G G W^{-1} e_u + \sigma_z^2 e_u^H W^{-1} G W^{-1} e_u - \rho_u^2 \]

\[ = \rho_u^2 - \rho_u^2. \quad (8) \]

Given Equation (5), (6), and (8), we derive the max-log approximated LLR of the bits associated with \(x_u\), given by

\[ L(b_{u,k}) = \gamma_u \left( \min_{s \in \chi} \left| \frac{\hat{x}_u}{\rho_u} - s \right|^2 - \min_{s \in \chi} \left| \frac{\hat{x}_u}{\rho_u} - s \right|^2 \right), \quad (9) \]

where \(b_{u,k}\) is the \(k\)-th mapping bit associated with \(x_u\); \(\gamma_u = \rho_u^2/\nu_u^2\) is the a posteriori signal-to-noise-plus-interference ratio (SINR); \(\chi^b_s \triangleq \{s|s \in \chi, y_k = b\}\) denotes the subset of \(\chi\), where the \(k\)-th mapping bit associated with the constellation symbol \(s\), i.e. \(y_k\), is \(b\); \(\chi\) is the constellation symbols set. After data detection of all users, the LLRs are fed into the soft-input channel decoder for decoding process.

\[ L = \sum_{l=0}^{\infty} (X^1 (X - W))^1 X^{-1}, \quad (10) \]

with the following condition satisfied:

\[ \lim_{l \to \infty} (I - X^{-1} W)^l = 0. \quad (11) \]

When the high order is ignored, the truncated Neumann series expansion can be expressed as

\[ L_{L_1} = \sum_{l=0}^{L-1} (X^{-1} (X - W))^1 X^{-1}. \quad (12) \]

Generally, when we select the matrix \(X\) that is close to \(W\), the \(L\) order expansion \(L_{L_1}\) in Equation (12) can be close to \(W^{-1}\). Fortunately, in massive MIMO systems, the gram matrix \(G\) is diagonally dominant; hence the diagonal matrix, i.e., \(D = \text{diag}\{W\}\) can be selected as \(X\), then the approximation of \(W^{-1}\) is given by

\[ L_{L_1} = \sum_{l=0}^{L-1} (D^{-1} (D - W))^1 D^{-1}. \quad (13) \]

In \([8]\), the authors have provided the upper bound of the residual estimation error using \(L_{L_1}\) to approach \(W^{-1}\), i.e.,

\[ \left\| (W^{-1} - L_{L_1}) y^{MF} \right\|_2 \leq \left\| I - D W \right\|_F \left\| \hat{x} \right\|_2, \quad (14) \]

where \(\|A\|_F\) and \(\|a\|_2\) are the Frobenius norm of a matrix \(A\) and the \(l_2\)-norm of a vector \(a\). From Equation (14), we can see that the upper bound of residual estimation error decreases as the increase of the expansion order and \(N_B\). In other words, if the number of antennas at the base station is sufficiently large, even with a small order expansion, the residual estimation error will be small. Particularly, when \(N_B\) is sufficient large and the expansion order \(L \leq 2\), the computation required for the Neumann series expansion will be much reduced, compared to the matrix inverse operations. These two factors provide the evidence to support the usage of the diagonal matrix in Neumann series expansion for massive MIMO systems.

\[ \text{C. Jacobi Method} \]

In Neumann series expansion, if the expansion order is greater than 2, the matrix multiplication operations are involved; hence, the computational complexity is comparable with that of the matrix inverse operations. On the other hand, as we can see in Equation (14), if \(N_B\) is not sufficiently large, with the expansion order that is less than 2, the residual estimation error is still considerable. These two factors limit the applications of diagonal matrix in Neumann series expansion.
To avoid the matrix multiplication operations, but maintain a reasonable orders of expansion, we can use the iterative methods. To be specific, we first rewrite the MMSE estimation in Equation (2) as
\[
W\hat{x} = y^{\text{MF}}.
\] (15)

By transforming the matrix inverse problem into the format of Equation (15), we can adopt the iterative methods to solve linear equations. Generally, the iterative methods follow the following process:

1. Provide an initial estimation;
2. Follow an iterative structure to obtain the next estimation;
3. When the estimation converges, output the final estimation.

In Jacobi method, we can have the initial estimation as
\[
x^{(0)} = D^{-1}y^{\text{MF}},
\] (16)
which is the common selection in most of the existing literature. The iterative structure is given by
\[
x^{(i+1)} = D^{-1}(D - W)x^{(i)} + D^{-1}y^{\text{MF}},
\] (17)
where \(x^{(i)}\) denotes the \(i\)-th estimation. According to the iterative structure in Equation (17), and use the initial estimation given by Equation (16), we can derive the \(i\)-th estimation given by
\[
x^{(i)} = \sum_{l=0}^{i} (D^{-1}(D - W))^l D^{-1}y^{\text{MF}}.
\] (18)
That is to say, by selecting the initial estimation given by (16), after \(i\) iterations following Jacobi iterative structure, we have the same estimation results as the \((i + 1)\)-th order expansion in Neumann series. Therefore, the convergence conditions, the residual estimation error, and the estimation results are the same as those in the previous subsection. However, as we can see from Equation (16) to Equation (17), only matrix-vector product operations are involved; therefore, Jacobi method has low complexity compared to the Neumann series expansion with the same iterations (or orders in Neumann series).

III. STAIR MATRIX AND ITS APPLICABILITY TO MASSIVE MIMO SYSTEMS

In this section, we will first introduce the stair matrix and its properties. And then, we will demonstrate the applicability of the stair matrix to massive MIMO systems.

A. Stair Matrix and its Properties

In an \(N \times N\) matrix \(A\), if its entry \(A_{m,m} = e_m^H A e_n, m, n = 1, 2, \ldots, N\), satisfies \(A_{m,m} = 0\) where \(m \notin [m - 1, m, m + 1]\), we then call it as a tridiagonal matrix, denoted by \(A = \text{tridiag}(A_{m,m-1}, A_{m,m}, A_{m,m+1})\). A special tridiagonal matrix is defined as a stair matrix if one of the following conditions is satisfied [20, 21]:

1. \(A_{m,m-1} = 0, A_{m,m+1} = 0\), where \(m = 2k - 1, k = 1, 2, \ldots, \lfloor (N + 1)/2 \rfloor\); (I)

Corollary 1: Let \(A\) be a stair matrix. Then \(A^H\) is also a stair matrix. In addition, if \(A\) is of type I, then \(A^H\) is of type II, and vice versa.

Proof: Using the definition, it is straightforward to obtain Corollary 1.

In accordance, a stair matrix is of type I if the condition (I) is satisfied and is of type II if the condition (II) is satisfied. For example, a 5 \(\times\) 5 stair matrix has the following forms:
\[
A = \begin{bmatrix}
\times & \times & \times & \times & \times \\
\times & \times & \times & \times & \times \\
\times & \times & \times & \times & \times \\
\times & \times & \times & \times & \times \\
\times & \times & \times & \times & \times
\end{bmatrix}
\] or \(A = \begin{bmatrix}
\times & \times & \times & \times & \times \\
\times & \times & \times & \times & \times \\
\times & \times & \times & \times & \times \\
\times & \times & \times & \times & \times \\
\times & \times & \times & \times & \times
\end{bmatrix}\).

The previous one is of type I and the latter one is of type II. Next, we provide the following properties of the stair matrix in Corollary 1 and 2.

Corollary 2: Let \(A\) be a stair matrix. \(A\) is nonsingular if and only if \(A_{m,m} = 0, m = 1, 2, \ldots, N\), is nonsingular. Furthermore, the inverse of \(A\), i.e., \(A^{-1}\) is also a stair matrix of the same type, given by \(A^{-1} = (2D - A)D^{-1}\), where \(D = \text{diag}(A)\).

Proof: Since \(\text{det}(A) = \prod_{m=1}^{N} A_{m,m}\), we can see that \(A\) is nonsingular if and only if \(A_{m,m}\), \(m = 1, 2, \ldots, N\), is nonsingular. Following the matrix multiplications, we can obtain that \(D^{-1}(2D - A)D^{-1}A = I_N\). Moreover, we can easily derive that \(A^{-1}\) is also a stair matrix and of the same type as \(A\).

Without loss of generality, we denote a stair matrix of type I as \(A = \text{stair}(A_{m,m-1}, A_{m,m}, A_{m,m+1})\). From Corollary 2 we have the Algorithm 1 to obtain \(A^{-1}\). It is clear from Algorithm 1 that the complexity to obtain the inverse of a stair matrix is \(O(N)\), which is the same order of the computation of \(D^{-1}\).

B. Using Stair Matrix in Neumann Series Expansion

We define the stair matrix \(S = \text{stair}(G_{u,u-1}, G_{u,u}, G_{u,u+1})\), derived from Gram matrix \(G\) as
\[
S_{(u,v)} = \begin{cases} 
G_{(u,v)}, & u \in U_1, v = u; \\
G_{(u,v)}, & u \in U_2, v \in \{u - 1, u, u + 1\}; \\
0, & \text{otherwise};
\end{cases}
\]
where \( U \triangleq \{ n|n \in \mathbb{N}, n \notin N_U \} \); \( U_1 \) and \( U_2 \) are subsets of \( U \), defined as \( U_1 \triangleq \{ n|n \in U, n = 2k - 1, k \in \mathbb{N} \} \), and \( U_2 \triangleq \{ n|n \in U, n = 2k, k \in \mathbb{N} \} \), respectively.

Applying the stair matrix in Neumann series expansion in Equation \((10)\), we have

\[
\mathbf{B}_{(u,v)} = \begin{cases} 
\frac{-\mathbf{G}_{(u,v)}}{\mathbf{G}_{(u,v)}}, & u \in U_1, v \neq u; \\
\frac{\mathbf{G}_{(u,v)}}{\mathbf{G}_{(u,v)}}, & u \in U_1, v = u; \\
\frac{\mathbf{G}_{(u,v)}}{\mathbf{G}_{(u,v)}} + \frac{\mathbf{G}_{(u,v+1)} \mathbf{G}_{(v+1,v)}}{\mathbf{G}_{(u,v)} \mathbf{G}_{(u+1,v+1)}}, & u \in U_2, v \neq u; \\
\frac{\mathbf{G}_{(u,v+1)}}{\mathbf{G}_{(u,v) + 1,v+1)}}, & u \in U_2, v = u.
\end{cases}
\]  

\((22)\)

Proof: See Appendix C.

Lemma 3: \( \mathbf{B}_{(u,v)} \) is given by Equation \((22)\). For \( u \in U_2, v \notin \{ u - 1, u, u + 1 \} \), and \( N_B > 4 \), we have

\[
\mathbb{E}\{|\mathbf{B}_{(u,v)}|^2\} \leq \frac{16A_3A_5}{B_1^3},
\]

\((28)\)

where \( A_3 \) and \( B_1 \) are given by Equations \((29)\) and \((25)\), respectively. \( A_5 \) is given by

\[
\begin{align*}
A_5 &= 576N_B + 24N_B (N_B - 1) (N_B - 2) (N_B - 3) \\
&\quad + 864N_B (N_B - 1) + 288N_B (N_B - 1) (N_B - 2).
\end{align*}
\]

\((32)\)
Proof: See Appendix E.

With the results in Lemma 7, we have

\[
E \left\{ \left\| \mathbf{B}_f^2 \right\| \right\} = \sum_{u=1}^{N_U} \sum_{v=1}^{N_U} E \left\{ \left| \mathbf{B}_{(u,v)} \right|^2 \right\}
\]

\[
\leq \frac{N^2_U - 1}{2} \frac{A_3}{B_1} + (N_U - 1) \frac{A_2}{B_1} + \left( \frac{N_U - 1}{2} \right) \sqrt{16A_4A_5}
\]

\[
+ \frac{N^2_U - 4N_U + 3}{2} \frac{12A_2A_3 + 6A_1A_3^2 + 24A_4 + 48\sqrt{A_1A_2}A_2^2}{B_1^3}
\]

(33)

Applying the Markov’s inequality, we have

\[
Pr \left\{ \left\| \mathbf{B}_f^2 \right\| < 1 \right\} = 1 - Pr \left\{ \left\| \mathbf{B}_f^2 \right\| \geq 1 \right\} \geq 1 - E \left\{ \left\| \mathbf{B}_f^2 \right\| \right\}.
\]

(35)

As \( \left\| \mathbf{B}_f^2 \right\| = \sum_{i=0}^{N^2_U - 1} |\lambda_i|^2 \), together with the inequality (35), we can see that with sufficiently large number of antennas at base station (i.e., \( N_B \rightarrow \infty \)), the probability that convergence condition in (21) is satisfied, will approach 1.

Following the similar analysis, we can also demonstrated that with sufficient large \( N_B \), using stair matrix, the probability that the convergence condition is met will also approach 1 in the approximation of the linear MMSE estimation. Hence we demonstrate the applicability of the stair matrix in massive MIMO systems.

C. Residual Estimation Error

We now investigate the residual estimation error by using the truncated Neumann series expansion. According to Equation (12), we have

\[
\mathbf{G}_{L}^{-1} = \sum_{l=0}^{L-1} \left( \mathbf{S}^{-1} (\mathbf{S} - \mathbf{G}) \right)^l \mathbf{S}^{-1}.
\]

Replacing \( \mathbf{G}^{-1} \) with \( \mathbf{G}_{L}^{-1} \) in Equation (4), we have

\[
\hat{x}^{(L)} = \mathbf{G}_{L}^{-1} \mathbf{y}^{MF}.
\]

(36)

Therefore, the residual estimation error \( J = \left\| \hat{x}^{(L)} - \hat{x} \right\|_2 \), is bounded as

\[
J = \left\| \left( \mathbf{G}^{-1} - \mathbf{G}_{L}^{-1} \right) \mathbf{y}^{MF} \right\|_2
\]

\[
= \left\| \sum_{l=L}^{\infty} \left( \mathbf{S}^{-1} (\mathbf{S} - \mathbf{G}) \right)^l \mathbf{S}^{-1} \mathbf{y}^{MF} \right\|_2
\]

\[
= \left\| \left( \mathbf{S}^{-1} (\mathbf{S} - \mathbf{G}) \right)^L \mathbf{G}^{-1} \mathbf{y}^{MF} \right\|_2
\]

\[
\leq \left\| \mathbf{B}_f^2 \right\| \left\| \hat{x} \right\|_2,
\]

(38)

where the inequality holds since \( \left\| \mathbf{A} \right\| \leq \left\| \mathbf{A} \right\| \left\| \mathbf{x} \right\|_2 \). As \( N_B \rightarrow \infty \), \( Pr \left\{ \left\| \mathbf{B}_f^2 \right\| < 1 \right\} \rightarrow 1 \). That is to say, the residual estimation error will approach 0 as indicated by inequality (38). Inequality (38) also indicates that increasing the truncation order in Neumann series expansion, the upper bound of the residual estimation error can be reduced. This evidence, together with high probability with the convergence condition to be met, supports the applications of the stair matrix to massive MIMO systems.

IV. IMPLEMENTATION OF THE STAIR MATRIX IN ITERATIVE METHOD

Due to the involvement of matrix multiplications, the truncation order in Neumann series expansion is limited to three; otherwise, the computational complexity is comparable with matrix inversion algorithm. Besides, we note that in existing work, the computation of the LLR is obtained by utilizing the NPI after the first truncation order in Neumann series expansion (or first iteration in iterative method). This implementation, however, causes significant performance loss when \( N_B \) is not sufficiently large (or \( r = N_B/N_U \) is not large, for example, \( r < 8 \)). In this section, we address these issues in the application of stair matrix in iterative method.

A. Stair Matrix in Iterative Method

Compared to the linear ZF detection, linear MMSE detection achieves a better balance in consideration of interference and noise. Therefore, we will introduce an iterative method to approach the linear MMSE detection.

To start with, we define the stair matrix \( \mathbf{S} = \text{stair}(\mathbf{W}_{(u,u-1)}, \mathbf{W}_{(u,u)}, \mathbf{W}_{(u,u+1)}) \). It is worth noting that compared to the stair matrix we discussed in previous section, the diagonal elements in the new stair matrix has increased by \( \sigma^2 \) according to Equation (3), which brings negligible computational cost. According to Equation (17), we have

\[
\mathbf{x}^{(i+1)} = \mathbf{S}^{-1} \left( (\mathbf{S} - \mathbf{W}) \mathbf{x}^{(i)} + \mathbf{y}^{MF} \right)
\]

\[
= \mathbf{x}^{(i)} - \mathbf{S}^{-1} \mathbf{W} \mathbf{x}^{(i)} + \mathbf{S}^{-1} \mathbf{y}^{MF},
\]

(39)

where \( \mathbf{x}^{(i)} \) is the \( i \)-th estimation.

In accordance, if the initial estimation \( \mathbf{x}^{(0)} \) is selected as

\[
\mathbf{x}^{(0)} = \mathbf{S}^{-1} \mathbf{y}^{MF},
\]

(40)

following the iterative process in Equation (39), we can derive

\[
\mathbf{x}^{(i)} = \sum_{l=0}^{i} (\mathbf{S}^{-1} (\mathbf{S} - \mathbf{W}))^l \mathbf{S}^{-1} \mathbf{y}^{MF},
\]

(41)

which indicates that the iterative method in Equation (39) can be seen as truncated Neumann series expansion method. However, in Equation (39), only matrix-vector product is involved, hence the overall computational complexity is of the order \( O(KN^2_U) \), where \( K \) denotes the iteration numbers.

B. Computation of the LLR

After the estimation of transmitted vector \( \mathbf{x} \), we need to compute the LLRs of the associated bits for the soft-input channel decoder. After \( K \) iterations, the equivalent channel gain \( \rho_u^{(K)} \) and the covariance of the NPI \( \left| \nu_u^{(K)} \right|^2 \) can be respectively given by

\[
\rho_u^{(K)} = e_u^H \mathbf{W}^{-1} \mathbf{K} \mathbf{G} e_u,
\]

(42)

\[
\left| \nu_u^{(K)} \right|^2 = e_u^H \mathbf{W}^{-1} \mathbf{K} \mathbf{G} \mathbf{G}^H \mathbf{W}^{-1} \mathbf{K} e_u + \sigma^2 e_u^H \mathbf{W}^{-1} \mathbf{K} \mathbf{G} \mathbf{G}^H \mathbf{W}^{-1} \mathbf{K} e_u - \left| \nu_u^{(K)} \right|^2
\]

(43)
Apparently, Equations 42 and 43 requires matrix multiplications if $K \geq 2$. Therefore, in [8], [13]–[15], $D^{-1}$, which is the first truncation order, is considered for the simplification. This approximation, however, as we will show in the next section, has caused a significant performance loss.

As we can see from Equation (8), the exact a posteriori covariance of the NPI in linear MMSE estimation can be derived if the equivalent channel gain is obtained. However, in [8], the authors have claimed that this relationship is not supported in truncated Neumann series expansion. The main reason for that claim is attributed to the fact that $W_{K}^{-1}$ is far away from $W_{-1}$ with small $K$. In previous section, we introduce the iterative method for detection, and the iteration numbers can be sufficiently large since the computational complexity in one iteration is of the order $O\left(N_{U}^{2}\right)$. With sufficiently large iterations, $W_{K}^{-1}$ can be close to $W_{-1}$ (we will show this in the next section); hence, we can used Equation (8) to derive the covariance of the NPI. The next question is how to maintain low computational complexity to obtain the equivalent channel gain.

We rewrite the equivalent channel gain in Equation (8) as

\[ \rho_{u} = e_{i}^{H}W_{-1}G\sigma_{e}^{2} = 1 - \sigma_{e}^{2}e_{i}^{H}W_{-1}e_{u}. \]

In addition, we replace $W_{-1}$ with $W_{K}^{-1}$, leading to

\[ \rho_{u}^{(K)} = 1 - \sigma_{e}^{2}e_{i}^{H}W_{K}^{-1}e_{u}. \]

That is to say, we need obtain the diagonal elements in $W_{K}^{-1}$ to compute $\rho_{u}^{(K)}$.

If $N_{B}$ and $r$ are sufficiently large, the Gram matrix $G$ and $W$ will become diagonal dominant; therefore, $D^{-1}$ can be a good approximation of $W_{-1}$, and we have the approximation to $\rho_{u}^{(K)}$ given by

\[ \rho_{u}^{(K)} \approx 1 - \sigma_{e}^{2}\sigma_{z}^{2}D_{(u,u)}^{-1}, \]

and $\left|\nu_{u}^{(K)}\right|^{2}$ is approximated as

\[ \left|\nu_{u}^{(K)}\right|^{2} \approx \rho_{u}^{(K)} \left(1 - \rho_{u}^{(K)}\right). \]

As a consequence, the a posteriori SINR is approximated as

\[ \gamma_{u}^{(K)} \approx \frac{\rho_{u}^{(K)}^{2}}{\left|\nu_{u}^{(K)}\right|^{2}} = \frac{\rho_{u}^{(K)}{\rho_{u}^{(K)}}}{1 - \rho_{u}^{(K)}}. \]

$\rho_{u}^{(K)}$ and $\gamma_{u}^{(K)}$ are used in Equation 9 to compute $L(b_{u,k})$. It is worth pointing out that although we utilize the diagonal matrix to estimate the equivalent channel gain, the computation of $\gamma_{u}^{(K)}$ in Equation 47 indicates that we try to approach the SINR in linear MMSE detection to derive the LLRs of the associated bits. This is quite different from the existing work [8], [13]–[15], where the SINR after the first iteration (or the first truncation order in Neumann series expansion method) is adopted. In fact, as the iterations increase, the covariance of the NPI will decrease, and our proposed approximation method is more efficient and accurate. In numerical simulations, we also validate that our approximation in (45) and (47) outperforms the approximation in existing work.

To summarize, we present Algorithm 2 for the proposed iterative method using stair matrix.

**Algorithm 2: Proposed Iterative Method Using Stair Matrix**

**Input:** $y$, $H$, $\sigma_{e}^{2}$, and Iteration number $K$.

**Output:** LLRs of the associated bits $L(b_{u,k})$.

**Initialization:**

1. $G = H^{H}H$, $W = G + \sigma_{e}^{2}I_{N_{U}}$, $y^{\text{MF}} = H^{H}y$.
2. $S = \text{stair}(\{W_{(u,u+1)}, W_{(u,u)}, W_{(u,u+1)}\})$.
3. Compute $S^{-1}$ through Algorithm 1, and $D^{-1} = \text{diag}(S^{-1})$.
4. Initial estimation: $x^{(0)} = S^{-1}y^{\text{MF}}$.

**Iteration:**

5. for $i = 1 : K$
6. \[ x^{(i)} = S^{-1}(S - W)x^{(i)} + y^{\text{MF}}; \]
7. end

**LLR Computation:**

8. $\rho_{u}^{(K)} = 1 - \sigma_{e}^{2}D_{(u,u)}^{-1}$, $\gamma_{u}^{(K)} = \frac{\rho_{u}^{(K)}}{1 - \rho_{u}^{(K)}}$.
9. $L(b_{u,k}) = \gamma_{u}^{(K)}\left(\min_{s \epsilon B_{k}} \frac{z_{u}^{(K)}(s)}{\rho_{u}^{(K)}} - s - \min_{s \epsilon B_{k}} \frac{z_{u}^{(K)}(s)}{\rho_{u}^{(K)}} - s^{2}\right)$.

Return $L(b_{u,k})$.

![Fig. 1. Cumulative distribution function of the maximum eigenvalue $N_{U} = 25$.](image)
next section, the stair matrix outperforms the diagonal matrix at all round.

V. NUMERICAL SIMULATIONS AND PERFORMANCE EVALUATION

A. Convergence Conditions

We first investigate the convergence condition using the stair matrix. Using Monte-Carlo method, we generate 2e7 random channel matrix H. For each H, we extract the diagonal matrix D and the stair matrix S, and compute the maximum eigenvalues of the matrix \( I - D^{-1}G \) and \( I - S^{-1}G \), respectively. Using numerical simulations, we eventually obtain the cumulative distribution function (CDF) of the maximum eigenvalues, given by Figure 1. In Figure 1, we evaluate the scenario that 25 users are in service and we increase the number of antennas at base station from 100 to 200. The following observations can be found:

- With the increase of antennas at base station, the probability that the convergence conditions are met, i.e., \( \Pr \{ \rho (I - S^{-1}G) < 1 \} \) and \( \Pr \{ \rho (I - D^{-1}G) < 1 \} \), will increase. Specifically, for the usage of the diagonal matrix, the probability that the convergence conditions are met, is increase from 0.29 when \( N_B = 100 \), to 1 when \( N_B = 200 \). In accordance, for the usage of the stair matrix, \( \Pr \{ \rho (I - S^{-1}G) < 1 \} \) is increased from 0.74 when \( N_B = 100 \), to 1 when \( N_B = 200 \); 
- In low \( r = N_B/N_U \leq 5 \) region, the usage of the stair matrix can increase the probability that the convergence conditions are met. For example, when \( N_B = 100 \), \( \Pr \{ \rho (I - D^{-1}G) < 1 \} \) is only 0.29, while \( \Pr \{ \rho (I - S^{-1}G) < 1 \} \) becomes 0.76. This indicates that in some low r region, the diagonal matrix is not applicable while the stair matrix can be used; 
- In any system configuration, \( \Pr \{ \rho (I - S^{-1}G) < a \} \geq \Pr \{ \rho (I - D^{-1}G) < a \}, a \in (0, 1) \). As the maximum eigenvalue determines the convergence rate, we can conclude that by using the stair matrix, the convergence rate is more likely faster compared to the usage of the diagonal matrix. 

The above observations validate the applicability of the usage of the stair matrix and diagonal matrix in massive MIMO systems. Besides, the results reveal that by using stair matrix, we can increase the probability that the convergence conditions are met in low r region compared to the usage of the diagonal matrix. Furthermore, we also find that by using the stair matrix, the convergence rate is more likely accelerated than the usage of the stair matrix.

B. Matrix Inverse

We now investigate the performance of the stair matrix in Neumann series expansion to approach the matrix inverse. We define \( \Delta(S) = \frac{1}{N_U} \left( I - \sum_{l=0}^{L-1} (I - S^{-1}G)S^{-1}G \right) \). In implementation, we propose the iterative method as shown in section IV. However, the results of the iterative method can be seen as the Neumann series expansion.

C. Residual Estimation Error

In iterative method, the estimation is to approach the estimation vector in linear ZF/MMSE method. In section IV.C, an upper bound of the residual estimation error for the use of the stair matrix in approaching linear ZF detection is presented. In order to differentiate the residual estimation error for the use of stair matrix and the diagonal matrix in linear ZF and
MMSE detection, we define the following terms:

\[
J(D_1) = \left\| \left(D_1^{-1} (D_1 - G)\right)^L G^{-1} y_{MF}\right\|^2_2,
\]
\[
J(D_2) = \left\| \left(D_2^{-1} (D_2 - W)\right)^L W^{-1} y_{MF}\right\|^2_2,
\]
\[
J(S_1) = \left\| \left(S_1^{-1} (S_1 - G)\right)^L G^{-1} y_{MF}\right\|^2_2,
\]
\[
J(S_2) = \left\| \left(S_2^{-1} (S_2 - W)\right)^L W^{-1} y_{MF}\right\|^2_2,
\]

where \( W = G + \sigma^2 I_{N_U} \), \( D_1 = \text{diag}\{G\} \), \( D_2 = \text{diag}\{W\} \), \( S_1 = \text{stair}\{G_{m,m-1}, G_{m,m}, G_{m,m+1}\} \), and \( S_2 = \text{stair}\{W_{m,m-1}, W_{m,m}, W_{m,m+1}\} \). According to Equation (38), we can see that \( J(D_1) \) and \( J(D_2) \) denote the residual estimation error for the use of the diagonal matrix in approaching linear ZF and MMSE detection, respectively. \( J(S_1) \) and \( J(S_2) \) denote the residual estimation error for the use of the stair matrix in approaching linear ZF and MMSE detection, respectively. For a given system configuration and average receiving SNR, we present the residual estimation error performance in Figure 3. The following observations are found:

- From Figure 3(a) and 3(b), \( J(S_1) \) is always less than \( J(D_1) \), and \( J(S_2) \) is always less than \( J(D_2) \) after the same iteration numbers. These results reflect that after the same iterations, using the stair matrix in iterative method can approach both the linear ZF and MMSE estimation more closely compared to the use of the diagonal matrix;
- In Figure 3(a) we note that, for the use of the diagonal matrix, the residual estimation error decreases slowly and remains a comparatively high level even with large iteration numbers. However, by using the stair matrix, we can speed up the decreasing rate and achieve a comparatively lower estimation error level. These results are consistent with the previous numerical results where we demonstrate that the use of the diagonal matrix may not be applicable in low \( r \) ratio.
- From Figure 3(a) and Figure 3(b) we can see that, with the increase of the receiving antennas at base station, the performance gain of the use of the stair matrix over the use of the diagonal matrix becomes small. These results are reasonable as \( N_B \) increases large, \( G \) and \( W \) both become diagonal dominant. However, we can also achieve comparatively lower residual estimation error by using the stair matrix in iterative method.

To summarize, we conclude that the use of the stair matrix outperforms the use of the diagonal matrix in terms of the residual estimation error. The performance gain is more significant in low \( r \) ratio, but still obvious in high \( r \) ratio.

\[\text{Fig. 3. Residual Estimation Error: (a) } N_B = 150, N_U = 25, \text{ average SNR=5dB}; \text{ (b) } N_B = 200, N_U = 25, \text{ average SNR=3.5dB}\]

\[\text{Fig. 4. BER performance: (a) } N_B = 150, N_U = 25; \text{ (b) } N_B = 250, N_U = 25.\]

D. BER Performance

We now evaluate the system BER performance. In the system, the base station is simultaneously serving \( N_U = 25 \) users. For each user, a LDPC code with code length 64800, code rate 1/2 is considered for channel code scheme. We con-
sider 64QAM modulation, and a block independent channel is considered for the evaluation.

To begin with, we investigate the proposed LLR computation given by (47), and the equivalent channel gain $\rho_u$ and the covariance of the NPI $v_u$ are approximated by (45) and (46). For comparison, we provide the linear MMSE detection as a benchmark, where the LLR computation is given by Equation (9) with $\rho_u$ and $v_u$ given by Equation (6) and (8), respectively. The LLR computation in existing work such as [8], [13], [14] is to compute the covariance of the NPI after the first iteration. It is worth pointing out that the iterative methods in [13], [14] requires less iterations to approach the linear MMSE detection; however, the LLR computation used in MMSE detection is not computed from the exact NPI of the MMSE detection, but the NPI after the first iteration. In Figure 4, we can see that the BER performance of the Jacobi method with the LLR computation in [8], [13], [14] is far away from the BER performance of the MMSE detection with the exact LLR computation. This is consistent with our previous analysis, where we pointed out that the covariance of the NPI will decrease with iterations. However, we note that the proposed LLR computation can greatly improve the BER performance of the Jacobi method by approximating the covariance of the NPI of the MMSE detection. Hereafter, we only utilize the proposed LLR computation for the BER performance comparison.

We now present the results with low $r = N_B/N_U$ region, and the results are presented in Figure 5. The following observations are found.

- From Figure 5(a), we note that the BER performance improvement with the proposed stair matrix compared to the diagonal matrix is obvious. However, the system performance is still far away from the MMSE detection even with sufficient large iterations. Specially, for the use of the diagonal matrix, the performance is level off after 9 iterations; for the use of the stair matrix, the performance is greatly improved, but a level off performance still appears. These are attributed to the slow convergence rate and not a 100 percent convergence conditions satisfied;

- From Figure 5(b) and Figure 5(c), we can see that the BER performance eventually converges to the performance of the MMSE detection. Specifically, in the system configuration $N_B = 150$, $N_U = 25$, at SNR= 5dB, the BER performance of the proposed iterative method after 13 iterations is almost the same as the performance of the MMSE detection. In the system configuration $N_B = 175$, $N_U = 25$, at SNR= 4dB, the BER performance of the Jacobi method after 9 iterations approaches the performance of the MMSE detection;

- From Figure 5(a) to Figure 5(c), we can see that the convergence rate of the proposed iterative method is faster than that of the Jacobi method. These results are consistent with the previous analysis. With faster convergence rate, fewer iterations are required for the proposed iterative method, hence reducing the overall system computational complexity.

Next, we evaluate the BER performance in the system

$$
\text{Fig. 5. BER performance: (a) } N_B = 125, N_U = 25; \text{ (b) } N_B = 150, N_U = 25; \text{ (c) } N_B = 175, N_U = 25.
$$

$$
\text{Fig. 6. BER performance: } N_B = 200, N_U = 25.
$$
configuration with high $r = N_B/N_U$ region, and the results are shown in Figure 6. It is clear that both the uses of the diagonal matrix and stair matrix require few iterations to converge. However, as indicated by the cumulative distribution function of the maximum eigenvalue, Pr $\{\rho(I - S^{-1}G) < a\} \geq Pr \{\rho(I - D^{-1}G) < a\}$, $a \in (0, 1)$, we can conclude that the convergence rate of the proposed iterative method using the stair matrix is faster than that of the Jacobi method using the diagonal matrix. The results validate these conclusions.

VI. CONCLUSIONS

In this paper, we propose the application of the stair matrix in massive MIMO systems. To begin with, we demonstrate that with sufficient large number of antennas at base station, the probability that the convergence conditions are met with the use of the stair matrix approaches 1. We then propose an iterative method to reduce the computational complexity and show that the overall computational complexity is of the same level as the existing iterative methods where the diagonal matrix is applied. Furthermore, we evaluate the performance of the stair matrix in terms of the probability that the convergence conditions are met, the normalized mean-square error of in Neumann series expansion to approach the matrix inverse, the residual estimation error of the iterative method to approach the linear ZF/MMSE estimation, and the system BER performance. Numerical simulations show that performance enhancement by using the stair matrix over the diagonal matrix is presented in all performance metrics.

APPENDIX

A. Preliminaries

We first present the preliminary lemmas.

Lemma 5: Let $a_k \sim CN(0, 1)$, we then have

$$E\{a_k^2\} = 1,$$

$$E\{a_k^4\} = 2,$$

$$E\{a_k^6\} = 6,$$

$$E\{a_k^8\} = 24,$$

where $A_1$ and $A_2$ are given by Equations (24) and (32).

Lemma 8: Let $A = a^H b b^H a$, where $a = [a_1, a_2, \ldots, a_{N_B}]^T$, $b = [b_1, b_2, \ldots, b_{N_B}]^T$, and $c = [c_1, c_2, \ldots, c_{N_B}]^T$, with each entry $a_k \sim CN(0, 1)$, $b_k \sim CN(0, 1)$, and $c_k \sim CN(0, 1)$, and i.i.d., we then have

$$E\{|A|^2\} = A_2,$$

where $A_2$ is given by Equation (27).

Lemma 9: Let $A = a^T b b^T c b^T d a^T c d a$, where $a = [a_1, a_2, \ldots, a_{N_B}]^T$, $b = [b_1, b_2, \ldots, b_{N_B}]^T$, $c = [c_1, c_2, \ldots, c_{N_B}]^T$, and $d = [d_1, d_2, \ldots, d_{N_B}]^T$, with each entry $a_k \sim CN(0, 1)$, $b_k \sim CN(0, 1)$, $c_k \sim CN(0, 1)$, and $d_k \sim CN(0, 1)$, and i.i.d., we then have

$$E\{|A|^2\} = A_4,$$

where $A_4$ is given by Equation (30).

B. Proof of Lemma 7

For $u \in \mathbb{U}_1, v \neq u$, from Equation (22), we have

$$E\{|B_{(u,v)}|^2\} = E\left\{\frac{|W_{(u,v)}|^2}{|W_{(u,u)}|^2}\right\} \leq \sqrt{E\{|W_{(u,v)}|^4\} \cdot E\{|W_{(u,u)}|^{-4}\}},$$

where the Cauchy-Schwarz inequality is applied [8]. From Lemma 7 and Lemma 8, we have

$$E\{|W_{(u,v)}|^4\} = A_1,$$

$$E\{|W_{(u,u)}|^{-4}\} = \frac{1}{B_1^4}.$$

Hence we complete the proof of Lemma 7.

C. Proof of Lemma 2

For $u \in \mathbb{U}_2, v = u - 1$, from Equation (22), we have

$$B_{(u,u-1)} = G_{(u,u-1)} G_{(u+1,u-1)}^{-1}.$$

Applying the Cauchy-Schwarz inequality, we have

$$E\{|B_{(u,u-1)}|^2\} \leq \sqrt{E\{|G_{(u,u+1)} G_{(u+1,u-1)}|^4\}} \cdot \sqrt{E\{|G_{(u,u)} G_{(u,u-1)}|^{-4}\}}.$$

According to Lemma 8 and Lemma 7, we have

$$E\{|G_{(u,u+1)} G_{(u+1,u-1)}|^4\} = A_2,$$

$$E\{|G_{(u,u)} G_{(u,u-1)}|^{-4}\} = \frac{1}{B_1^4}.$$

For $u \in \mathbb{U}_2, v = u + 1$, following the similar process, we have the same result above. Therefore, we complete the proof of Lemma 2.
we have \( E \) expectation in the right hand side of the inequality (65) given

\[
\begin{align*}
E \left\{ B_{(u,v)}^2 \right\} &= E \left\{ \left| G_{(u+1,u+1)} G_{(u,u-1)} G_{(u-1,v)} + G_{(u-1,u-1)} G_{(u,u+1)} G_{(u+1,v)} - G_{(u-1,u-1)} G_{(u+1,u+1)} G_{(u,v)} \right|^2 \right\} \\
&\leq \sqrt{E \left\{ \left| G_{(u+1,u+1)} G_{(u,u-1)} G_{(u-1,v)} + G_{(u-1,u-1)} G_{(u,u+1)} G_{(u+1,v)} - G_{(u-1,u-1)} G_{(u+1,u+1)} G_{(u,v)} \right|^4 \right\}} \\
&\cdot \sqrt{E \left\{ \left| G_{(u-1,u-1)} G_{(u,u)} G_{(u+1,u+1)} \right|^4 \right\}}
\end{align*}
\]  

(65)

\[
\begin{align*}
E \left\{ \left| G_{(u+1,u+1)} G_{(u,u-1)} G_{(u-1,v)} + G_{(u-1,u-1)} G_{(u,u+1)} G_{(u+1,v)} - G_{(u-1,u-1)} G_{(u+1,u+1)} G_{(u,v)} \right|^4 \right\} \\
= E \left\{ (A + B + C + D + E + F)^2 \right\} \\
\leq 6E \left\{ A^2 + B^2 + C^2 + D^2 + E^2 + F^2 \right\}
\end{align*}
\]  

(66)

\[
E \left\{ D^2 \right\} \leq 4E \left\{ \left| G_{(u+1,u+1)} G_{(u,u-1)} G_{(u-1,v)} G^{*}_{(u,u+1)} G^{*}_{(u+1,v)} \right|^2 \right\} = 4A_4
\]  

(69)

\[
\begin{align*}
E \left\{ B_{(u,u)}^2 \right\} &= E \left\{ \left\| G_{(u+1,u+1)} G_{(u,u-1)} G_{(u-1,v)} + G_{(u-1,u-1)} G_{(u,u+1)} G_{(u+1,v)} - G_{(u-1,u-1)} G_{(u+1,u+1)} G_{(u,v)} \right\|^2 \right\} \\
&\leq \sqrt{E \left\{ \left\| G_{(u+1,u+1)} G_{(u,u-1)} G_{(u-1,v)} + G_{(u-1,u-1)} G_{(u,u+1)} G_{(u+1,v)} - G_{(u-1,u-1)} G_{(u+1,u+1)} G_{(u,v)} \right\|^4 \right\}} \\
&\cdot \sqrt{E \left\{ \left\| G_{(u-1,u-1)} G_{(u,u)} G_{(u+1,u+1)} \right\|^4 \right\}}
\end{align*}
\]  

(73)

\[
\begin{align*}
E \left\{ C^2 \right\} \text{ is given by}
E \left\{ C^2 \right\} &= A_1 A_3^2. 
\end{align*}
\]  

(68)

\[
\begin{align*}
\text{where the results in Lemma 9 and Lemma 7 are applied.}
\end{align*}
\]

By using \((\Re(a))^2 \leq |a|^2\), we derive the result of \( E \left\{ D^2 \right\} \), given by (69), where \( A_4 \) is obtained through Lemma 9.

Applying the Cauchy-Schwarz inequality, we have

\[
\begin{align*}
E \left\{ E^2 \right\} &\leq 4E \left\{ \left| G_{(u-1,u-1)} G_{(u,u-1)} G_{(u-1,v)} G^{*}_{(u,u+1)} \right|^4 \right\} \\
&\cdot E \left\{ \left\| G_{(u+1,u+1)} \right\|^4 \right\} \\
&\leq 4E \left\{ \left| G_{(u+1,u+1)} \right|^4 \right\} \sqrt{E \left\{ \left| G_{(u,u-1)} G_{(u-1,v)} \right|^4 \right\}} \\
&\cdot \sqrt{E \left\{ \left| G_{(u-1,u-1)} G_{(u,u)} \right|^4 \right\}}
\end{align*}
\]  

(70)

With the results in Lemma 6, Lemma 7 and Lemma 8 we derive the result of \( E \left\{ E^2 \right\} = E \left\{ F^2 \right\} \leq 4A_3 \sqrt{A_1 A_2 A_3} \).

(71)

Therefore, we derive

\[
E \left\{ B_{(u,v)}^2 \right\} \leq \sqrt{\frac{12A_2 A_3 + 6A_1 A_2^2 + 24A_4 + 48\sqrt{A_1 A_2 A_3}}{B_1^4}}
\]  

(72)

Hence, we complete the proof of Lemma 3.

\[\text{E. Proof of Lemma 4}\]

For \( u \in \mathbb{U}_2, \ v = u \), from Equation (22), we have

\[
E \left\{ B_{(u,v)}^2 \right\} \text{ given by (73), where the Cauchy-Schwarz in-}
\]
equality is applied. By using \( |a + b|^2 \leq 2 \left( |a|^2 + |b|^2 \right) \), we have
\[
\begin{align*}
|G_{(u+1,u+1)}|G_{(u,u-1)}|^2 + G_{(u-1,u-1)}G_{(u,u+1)}|^2 & \\
& \leq 2 \left( |G_{(u+1,u+1)}|G_{(u,u-1)}|^4 + |G_{(u-1,u-1)}|^2 |G_{(u,u+1)}|^4 \right), \\
& \leq 8 \left( |G_{(u+1,u+1)}|^4 |G_{(u,u-1)}|^8 + |G_{(u-1,u-1)}|^4 |G_{(u,u+1)}|^8 \right),
\end{align*}
\]
(74)

Therefore, we derive
\[
\begin{align*}
E \left\{ |G_{(u+1,u+1)}|G_{(u,u-1)}|^2 + G_{(u-1,u-1)}G_{(u,u+1)}|^2 \right\} & \\
& \leq 8E \left( |G_{(u+1,u+1)}|^4 \right)^2 E \left( |G_{(u,u-1)}|^8 \right) \\
& + 8E \left( |G_{(u-1,u-1)}|^4 \right)^2 E \left( |G_{(u,u+1)}|^8 \right).
\end{align*}
\]
(75)

With the results in Lemma 6 and 7 we have
\[
E \left\{ |B_{(u,u)}|^2 \right\} \leq \sqrt{\frac{16A_3A_5}{B_1^2}}. 
\]
(76)

Hence we complete the proof of Lemma 4.

REFERENCES


