Statistical Multimode Transmit Antenna Selection for Limited Feedback MIMO Systems

Chang Soon Park, Member, IEEE, and Kwang Bok Lee, Senior Member, IEEE

Abstract—In a wireless multiple-input multiple-output (MIMO) system, transmit antenna selection is an effective means of achieving good performance with low complexity. We consider spatial multiplexing with linear receivers, and equal power and equal rate allocation over different selected transmit antennas in order to reduce feedback overhead. Under these constraints, we address the problem of statistical multimode transmit antenna subset selection to improve the capacity of spatially correlated MIMO fading channels. In particular, we first derive an analytical closed-form expression for the expectation of the lower bound on the capacity using the smallest eigenvalue distribution of a Wishart matrix. Then, we propose a transmit antenna subset selection criterion of maximizing this average lower-bound capacity.

Index Terms—Antenna selection, limited feedback systems, linear receiver, MIMO systems.

I. INTRODUCTION

RECENT information-theoretic results have shown that multiple-input multiple-output (MIMO) systems, which employ multiple antennas at both the transmitter and receiver, provide considerable capacity enhancement of wireless channels [1], [2]. In practical systems, however, hardware cost of RF chains for multiple antennas may limit the number of antennas. In order to overcome this limitation while achieving satisfactory MIMO gain, various (receive or transmit) antenna selection techniques have been investigated in [3]–[8], where the antenna subset is adapted to the instantaneous channel conditions. While the number of selected antennas is assumed to be fixed in [3]–[7], that of the transmitter is varied in [8], and the increased degrees of freedom in the number of transmitted substreams have been shown to improve the performance. However, the transmit antenna selection based on the instantaneous channel conditions may cause the performance degradation due to the limited feedback channel bandwidth, limiting its application to practical systems over fast fading channels. Furthermore, it requires the channel estimation process for all transmit antennas at every transmission time interval. In [9]–[12], the transmit antenna selection is based not on the instantaneous channel conditions but on the long-term channel statistics. However, the feedback overhead for different rates for all selected transmit antennas still remains and the number of selected antennas is not variable in [9]. The works in [10] and [11], respectively, use the selection criteria of maximizing the minimum signal-to-noise ratio (SNR) margin and of minimizing the average probability of error, and those still leave the need to investigate the antenna selection for capacity improvement. The optimal antenna selection based on capacity maximization is proposed for the optimal (not linear) detector in [12], where the number of selected antennas is assumed to be fixed.

In this paper, we address the problem of statistical (not instantaneous) multimode transmit antenna subset selection to improve the capacity (not error-rate performance) of spatially correlated MIMO fading channels with small feedback overhead. Multimode antenna selection allows all possible numbers of transmitted antennas to be selected, as in [8] and [10]. While our work investigates multimode selection based on the long-term channel statistics, the work in [8] investigates that based on the instantaneous channel conditions, and it is extended to a more general case using quantized precoding in [16]. The number of transmit antennas in [10] is selected to maximize the minimum SNR margin, while we select it to improve the capacity. In other works in [17]–[19], spatial mode selection techniques have been studied to adjust the number of substreams for transmission. The switching architecture in [17] presents the selection between space-time coding and spatial multiplexing based on the instantaneous channel conditions. The work in [18] proposes spatial mode selection using the joint transmit and receive minimum mean square error (MMSE) design, however, it assumes full knowledge of channel conditions at the transmitter, while we assume a limited feedback scenario. The adaptive transmission in [19] statistically switches between several spatial modes, although those modes are not transmit antenna subset selection.

We consider the linear receiver with low complexity which uses the zero-forcing (ZF) or MMSE detection scheme. In order to reduce feedback overhead and system complexity, we allocate equal power and equal rate to all selected transmit antennas, from which independent data substreams are transmitted. To solve the problem, we first derive an analytical closed-form expression for the expectation of the lower bound on the capacity using the smallest eigenvalue distribution of a Wishart matrix. Then, we propose a selection criterion of maximizing this average lower-bound capacity.

II. SYSTEM AND CHANNEL MODELS

A point-to-point MIMO system with \( N_t \) transmit antennas and \( N_r \) receive antennas is considered. When the number
of selected transmit antennas is denoted by \( M_t \), \( M_t \) can be taken from 1 to \( \min(N_t, N_r) \) and it is a “variable” to be selected. The subset, which consists of indices for selected transmit antennas, is denoted by \( S \). For example, when all available transmit antennas are selected, we obtain \( M_t = N_t \) and \( S = \{ 1, 2, \ldots, N_t \} \). At the receiver, all \( N_r \) receive antennas are assumed to be used to detect the transmitted data signals. Throughout this paper, the superscripts \([\cdot]^T\) and \([\cdot]^H\) denote the transpose and conjugate transpose, respectively. Moreover, \([A]_{i,j}\) denotes the element in the \( i \)th row and \( j \)th column of a matrix \( A \), and \( E[\cdot] \) represents the expectation operator.

The \( N_r \times N_t \) column matrix is modeled as [13]

\[
H = R_{Rx}^{1/2} \tilde{H}_w(S) R_{Tx}^{1/2}
\]

where \( R_{Rx} = R_{Rx}^{1/2} R_{Rx}^{1/2} \) and \( R_{Tx} = R_{Tx}^{1/2} R_{Tx}^{1/2} \) are \( N_r \times N_r \) and \( N_t \times N_t \) transmit correlation matrices, respectively. The elements of \( \tilde{H}_w \) in (1) are assumed to be independent and identically distributed (i.i.d.) circularly symmetric complex Gaussian random variables with zero mean and unit variance. The \( N_r \times M_t \) channel matrix for transmit antenna subset \( S \) may be written as

\[
\tilde{H}(S) = R_{Rx}^{1/2} \tilde{H}_w(S) R_{Tx}^{1/2}(S)
\]

where \( \tilde{H}_w(S) \) is an \( M_t \times M_t \) principal submatrix of \( \tilde{H}_w \), and \( \tilde{H}_w(S) \) is an \( N_r \times M_t \) matrix whose elements have the same distributions as those of \( \tilde{H}_w \). It is assumed that \( \tilde{H}_w(S) \) has full rank. Here, note that \( \tilde{H}_w(S) \) may not be a submatrix of \( \tilde{H}_w \), which is contrary to the case of \( \tilde{H}_w(S) \) and \( \tilde{H}_w(S) \).

The received signal vector for a given \( S \) may be expressed as

\[
y = \tilde{H}(S)x + n
\]

where \( x \) is an \( M_t \times 1 \) transmit symbol vector. The covariance matrix of \( x \) is given by \( E[xx^H] = (E_x/M_t) I_{M_t} \), where \( E_x \) is the total average transmit energy and \( I_{M_t} \) is the \( M_t \times M_t \) identity matrix. This implies that independent data symbols are transmitted from all selected transmit antennas. The vector \( n \) denotes an \( N_r \times 1 \) circularly symmetric complex Gaussian noise vector with zero mean and covariance matrix \( E[nn^H] = \sigma_n^2 I_{N_r} \).

We assume that the receiver has perfect knowledge of the instantaneous and statistical channel state information (CSI). Based on the statistical CSI, such as the correlation matrices and average SNR, the transmit antenna subset is selected at the receiver and its information is fed back to the transmitter. This feedback information is not necessary to adapt to the varying channel conditions, if the channel statistics are not changed. Only a single information for equal rate adaptation based on the instantaneous CSI is fed back to the transmitter every channel instance. Here, we assume that the instantaneous rate adaptation is performed every channel instance to achieve the instantaneous (not ergodic) capacity with short-term (not long-term) delay constraint.

The ergodic capacity requires long-term delay constraint since we construct codewords long enough to cover all channel states according to the channel’s statistical distribution. In practical systems, however, it may be more feasible to construct codewords according to the given instantaneous channel condition and change the data rate every channel instance. In this case, the instantaneous channel capacity can be achieved every channel instance and it may be treated as a random variable. We consider this case and the information for instantaneous rate should be fed back to the transmitter every time the instantaneous channel state changes. Therefore, if different rates are allocated to different transmit antennas, the feedback overhead is proportional to the number of selected transmit antennas. As a result, compared to the different rate allocation, the equal rate allocation reduces the feedback overhead by a factor of \( M_t \). As an example, in the case of a practical system using 5 bits for representing different coding rates and modulation levels, only 5 bits are required for equal rate allocation while 20 bits are required for different rate allocation for \( M_t = 4 \).

### III. Statistical Transmit Antenna Subset Selection

In order to reduce feedback overhead and system complexity, we allocate equal rate across all selected transmit antennas. This equal rate should be the capacity value computed using the minimum SNR, since the data signal can be sent reliably with an arbitrarily small probability of error when the rate is smaller than the capacity value [14]. Thus, the instantaneous overall capacity for equal rate allocation for a given \( S \) is calculated as

\[
C_{ER}(S) = M_t \log_2(1 + \text{SNR}_{min}(S))
\]

where \( \text{SNR}_{min}(S) \) is the minimum among the SNR values for the linear receiver, for example, \( \text{SNR}_{min}(S) = \min_{1 \leq k \leq M_t} E_s/(M_t \sigma_n^2) \). Here, \( \text{SNR}_{min}(S) \) is the instantaneous rate for equal rate allocation while 20 bits are required for different rate allocation for \( M_t = 4 \).

Now, we propose a criterion for selecting transmit antenna subset based on the long-term statistics of MIMO fading channels. The distribution of \( \text{SNR}_{min}(S) \) in (4) should be derived to find the analytical expectation of \( C_{ER}(S) \) with respect to random fading channels. However, the exact solution of that problem is difficult to find. Thus, we use a lower bound on \( \text{SNR}_{min}(S) \) (for both ZF and MMSE receivers), which is given as [10]

\[
\text{SNR}_{min}(S) \geq \frac{E_s}{M_t \sigma_n^2} \lambda_{min}(R_{Rx}) \lambda_{min}(\tilde{R}_{Tx}(S))
\]

where \( \lambda_{min}(A) \) denotes the smallest (real) eigenvalue of a Hermitian matrix \( A \). Correspondingly, the \( C_{ER}(S) \) in (4) is lower bounded as

\[
C_{ER}(S) \geq C_{ER,LB}(S)
\]

with the lower bound \( C_{ER,LB}(S) \) given as

\[
C_{ER,LB}(S) = M_t \log_2 \left( 1 + \frac{E_s}{M_t \sigma_n^2} \lambda_{min}(R_{Rx}) \lambda_{min}(\tilde{R}_{Tx}(S)) \times \lambda_{min}(\tilde{H}_w(S)^H \tilde{H}_w(S)) \right).
\]
Then, the expectation of $C_{\text{ER, LB}}(S)$ with respect to random fading channels is expressed as

$$E[C_{\text{ER, LB}}(S)] = M_t \int_0^\infty \log_2(1 + \rho(S)\lambda)f_{\min}^{ \text{M}_t}(\lambda)d\lambda$$

(8)

where $\rho(S) = (E_s/(M_t\sigma^2))\Lambda_{\min}(\mathbf{R}_{\text{Rx}})\Lambda_{\min}(\mathbf{R}_{\text{Tx}}(S))$ and $f_{\min}^{ \text{M}_t}(\lambda)$ denotes the probability density function (pdf) of $\Lambda_{\min} = \Lambda_{\min}(\hat{\mathbf{H}}_w(S)^H\hat{\mathbf{H}}_w(S))$.

Note that an $M_t \times M_t$ random matrix $\hat{\mathbf{H}}_w(S)^H\hat{\mathbf{H}}_w(S)$ is called a complex Wishart matrix and $\Lambda_{\min}$ is its smallest eigenvalue. Using the results in [15] and considering that the elements of $\hat{\mathbf{H}}_w(S)$ have unit variance, the pdf of $\Lambda_{\min}$ can be given as

$$f_{\min}^{ \text{M}_t}(\lambda) = \frac{K_{M_tN_r}}{(M_t-1)!}e^{-M_t\lambda}\lambda^{N_r-M_t-1}\int \prod_{1\leq i<j\leq M_t-1}(x_i-x_j)^2$$

$$\times \prod_{i=M_t-1}^{M_t-1}x_i^2(\lambda)^{N_r-M_t}e^{-x_i^2}dx_i$$

(9)

where $K_{M_tN_r} = \left(\prod_{i=1}^{M_t}(N_r-i)!(M_t-i)\right)^{-1}$ and $R = \{x_1, x_2, \ldots, x_{M_t-1}\}$ is the integration region. From (9), it can be easily shown that $f_{\min}^{ \text{M}_t}(\lambda)$ is expressed as [15]

$$f_{\min}^{ \text{M}_t}(\lambda) = e^{-M_t\lambda}\sum_{n=0}^{D}c_n^{(M_t)}\lambda^n$$

(10)

where $D = M_t(N_r-M_t)$ is the degree of the polynomial part of $f_{\min}^{ \text{M}_t}(\lambda)$ and $c_n^{(M_t)}$ is the constant coefficient of $\lambda^n$ for a given $M_t$. In [15], the polynomial part in (10) still remains unclosed except for the very simple case of $M_t = N_r$. The closed-form expressions for $f_{\min}^{ \text{M}_t}(\lambda)$, and thus $c_n^{(M_t)}$, in (10) are derived in Appendix A. In this Appendix, the solutions for $M_t = 1, 2, 3, 4$ and with arbitrary $N_r$ are derived, since $M_t$ seems to be seldom more than four in practical systems. Particularly, we provide the exact pdf results for $M_t = 2, 3, 4$, and $N_r \leq 6$ in Table I. To the best of authors’ knowledge, these simple and explicit expressions in Appendix A, which do not require complicated integration and consist of a single exponential function multiplied by a polynomial of degree $D$, have not been found in the previous works. Furthermore, the expression of (10) can lead to the derivation of a closed-form average lower-bound capacity as described in the following.

Using (10), $E[C_{\text{ER, LB}}(S)]$ in (8) is given as

$$E[C_{\text{ER, LB}}(S)] = M_t \sum_{n=0}^{D}c_n^{(M_t)}\int_0^\infty \log_2(1 + \rho(S)\lambda)\lambda^n e^{-M_t\lambda}d\lambda.$$  

(11)

**Table I**

<table>
<thead>
<tr>
<th>$M_t$</th>
<th>$N_r$</th>
<th>$\rho(S)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>2</td>
<td>$2e^{-2\lambda}$</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>$e^{-2\lambda}(2 + 3\lambda)$</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>$e^{-2\lambda}(4\lambda + 6\lambda^2 + 12\lambda^3)/6$</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>$e^{-2\lambda}(5\lambda^2 + 9\lambda^3 + 36\lambda^4 + 60\lambda^5)/72$</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>$e^{-2\lambda}(6\lambda^3 + 12\lambda^2 + 72\lambda^4 + 240\lambda^5 + 360\lambda^6)/1440$</td>
</tr>
</tbody>
</table>

The integral term in (11) is derived in Appendix B. Using this result,

$$E[C_{\text{ER, LB}}(S)] = \sum_{n=0}^{D}c_n^{(M_t)}\int_0^\infty \frac{(-1)^ne^{\rho(S)\lambda}}{(\ln 2)^n}$$

$$\times \sum_{k=0}^{n}\frac{(-1)^k}{(n-k)!}\frac{1}{\mu(S)^k}\sum_{i=0}^{k}\frac{\Gamma(i, \mu(S))}{i!}$$

(12)

where $\mu(S) = M_t/\rho(S)$ and $\Gamma(\alpha, x) = \int_x^\infty e^{-t}\Gamma^{\alpha-1}dt$ denotes the incomplete gamma function [23].

Now, we propose the selection criterion, which is to select the transmit antenna subset $S^*$ that maximizes the closed-form average lower-bound capacity in (12). The straightforward application of this criterion requires at most $\min(N_t, N_r)$ transmit antenna subsets, each one of which is the subset corresponding to maximum $\rho(S)$ for each $M_t$ and is denoted by $S^*(M_t)$ ($M_t = 1, 2, \ldots, \min(N_t, N_r)$). Then, we select the one subset $S^*$ among $S^*(M_t)$’s that maximizes (12) (i.e., the number of all possible $S$’s) computations of $E[C_{\text{ER, LB}}(S)]$ in (12). From (11), however, it can be seen that more $\rho(S)$ (depending on $\Lambda_{\min}(\hat{\mathbf{R}}_{\text{Tx}}(S))$) leads to more $E[C_{\text{ER, LB}}(S)]$ when a particular value of $M_t$ is given. This implies that a simple comparison of $\rho(S)$’s can provide the best transmit antenna subset for a “particular” $M_t$ without computation of $E[C_{\text{ER, LB}}(S)]$. Thus, we first choose $\min(N_t, N_r)$ transmit antenna subsets, each one of which is the subset corresponding to maximum $\rho(S)$ for each $M_t$ and is denoted by $S^*(M_t)$ ($M_t = 1, 2, \ldots, \min(N_t, N_r)$). Then, we select the one subset $S^*$ among $S^*(M_t)$’s that maximizes (12). This two-step selection procedure is described as follows.

**Step 1** Select $S^*(M_t)$’s for $M_t = 1, 2, \ldots, \min(N_t, N_r)$ as

$$S^*(M_t) = \arg\max_{S \subset \{1, 2, \ldots, N_r\}, |S| = M_t} \rho(S)$$

$$= \arg\max_{S \subset \{1, 2, \ldots, N_r\}, |S| = M_t} \Lambda_{\min}(\hat{\mathbf{R}}_{\text{Tx}}(S))$$

(13)

where $|S|$ denotes the number of elements of a set $S$.
Step 2) Select $S^*$ among $S^*(M_t)$’s obtained from Step 1 as

$$
S^* = \arg \max_{S: S = S^*(M_t), 1 \leq M_t \leq \min(N_t, N_r)} E[C_{ER, LB}(S)]
$$

$$
= \arg \max_{S: S = S^*(M_t), 1 \leq M_t \leq \min(N_t, N_r)} \sum_{n=0}^{D(M_t)} \frac{(-1)^n e^{\mu(S)}}{(\ln 2)^{p(S)^n}} \times \sum_{k=0}^{n} \left( -1 \right)^k \frac{n!}{(n-k)!} \frac{1}{\mu(S)^k} \sum_{i=0}^{k} \frac{\Gamma(i, \mu(S))}{i!}
$$

(14)

where $c_n^{(M_t)}$’s are given by (A.11).

As shown in (13), the selection of $S^*(M_t)$ in Step 1 is dependent on the smallest eigenvalue of the corresponding transmit “correlation submatrix”. On the other hand, the selection of $S^*$ (and thus “$M_t^*$”, i.e., the number of independent data substreams) in Step 2 is dependent on the “average SNR” ranges, which will be shown in the following section. It should be noted that this two-step procedure provides the same result as the exhaustive comparison of (12) for all possible transmit antenna subsets, as described previously.

IV. NUMERICAL RESULTS

In this section, the performance of the proposed antenna selection is provided. The average SNR is defined to be $E_x / \sigma^2$. The average capacity $E[C_{ER}(S^*)]$ is calculated by averaging the capacity $C_{ER}(S^*)$ in (4), where $S^*$ is selected by the criterion in (13) and (14), over 10,000 independently generated channel realizations (i.e., by Monte Carlo simulations). For performance comparisons, we also provide $E[C_{ER}(S)]$ obtained by Monte Carlo simulations for the transmissions with fixed $S$, which is called the non-adaptive technique in this section. Furthermore, we assume no fading correlation among receive antennas (i.e., $R_{Rx} = I_{N_r}$), and the ZF detection is assumed to be used.

Prior to the performance comparisons, we first compare $E[C_{ER}(S)]$ and its lower bound $E[C_{ER, LB}(S)]$ for fixed $S$’s in Fig. 1, when $N_t = N_r = 3$ and $R_{Tx} = I_{N_t}$ (i.e., i.i.d. fading channels). The values of $E[C_{ER, LB}(S)]$ are illustrated using (12), and can be shown to be equal to the values obtained by Monte Carlo simulations, although those are not included in Fig. 1. The lower bound is observed to become tighter for smaller $M_t$ (note that $E[C_{ER, LB}(S)]$ becomes equal to $E[C_{ER}(S)]$ in the case of $M_t = 1$). Fig. 1 also shows that the performance for smaller $M_t$ tends to be superior to that for larger $M_t$ at low SNR range, while the reverse trend is observed at high SNR range. This may be explained by using the property of SNR (in “linear” receiver), which tends to increase with $M_t$ decreasing, and the property of log function, which increases logarithmically with SNR increasing. Although the SNR can grow by decreasing $M_t$, the effect of increased SNR on the capacity may be negligible at “high” SNR range due to the property of log function. Thus, in order to increase the overall capacity in (7), it may be more effective to increase $M_t$ (rather than to decrease $M_t$) at high SNR range. On the contrary, at low SNR range, the log function increases more rapidly with the SNR,4 and thus, the effect of increased SNR on the capacity becomes great. Hence, the opposite behavior may be observed at low SNR range.

Fig. 2 compares the average capacity performance, when $N_t = 6$, $N_r = 4$, and $R_{Tx} = I_{N_t}$. Note that all fixed $S$’s for a given $M_t$ provide the same performance in this case of i.i.d. fading channels. Thus, the statistical selection of $S$ may be performed by only Step 2 in (14) (among arbitrary $S^*(M_t)$’s) without passing through Step 1 in (13). It can be seen that the proposed technique outperforms the non-adaptive techniques over almost all the SNR ranges. Slight performance degradation is observed in the SNR range close to the crossing point of capacity curves for fixed $S = \{1, 2\}$ and $\{1, 2, 3\}$. This is caused by the difference between $C_{ER}(S)$ in (4) and $C_{ER, LB}(S)$ in (7), as shown in Fig 1. Fig. 2 also shows that, in the proposed algorithm, the statistically selected $M_t$ increases with average SNR increasing, as expected from the results of Fig. 1.

Fig. 3 shows the average capacity of the proposed technique

3This can be shown by using (5) and pdf of $\Lambda_{\text{min}}$ in Table I.

4Approximately, the log function grows linearly with the SNR at low SNR range, i.e., $\log_2(1+x) \approx x \log_2 e$ for $0 < x \ll 1$. 

Authorized licensed use limited to: Seoul National University. Downloaded on January 28, 2009 at 22:43 from IEEE Xplore. Restrictions apply.
in the presence of transmit antenna correlation, when \( N_t = 6 \) and \( N_r = 4 \). We also provide the performance for the worst cases of fixed \( S \)'s, each one of which shows the worst performance for each given \( M_t \). Using the fading correlation model in [13], the correlation matrix for transmit antennas spaced equally along an axis is given by \( \mathbf{R}_{Tx} \), where \( J_0(\cdot) \) is the zeroth order Bessel function of the first kind, \( \theta \) is the angle spread, \( \lambda_w \) is the wavelength, and \( d_{Tx} \) denotes the distance between adjacent transmit antennas. It is assumed that \( \theta = \theta^* \) and \( d_{Tx} = 3\lambda_w \).

The proposed technique is observed to outperform the non-adaptive techniques over all the SNR ranges. The SNR gain of the proposed scheme is about 3.5 dB over the non-adaptive scheme of fixed \( S = \{1, 2\} \), 9.4 dB over that of \( S = \{1\} \), and 11 dB over that of \( S = \{1, 2, 3\} \), at average capacity of 10 bps/Hz.

Fig. 4 compares the average capacity of the proposed technique with that of the non-adaptive techniques for the best cases of fixed \( S \)'s for each \( M_t \). Note that these best cases show the best performance when each particular value of \( M_t \) is fixed regardless of SNR ranges. The simulation configurations are the same as those in Fig. 3. We first observe that the proposed algorithm selects the best \( S^*(M_t) \)'s for any given value of \( M_t \) (by the procedure of Step 1 in (13)), although it does not hold necessarily. Next, although the proposed technique outperforms the non-adaptive techniques over almost all the SNR ranges, it may not select best \( S^* \) among \( S^*(M_t) \)'s for all \( M_t \) (by the procedure of Step 2 in (14)) at some values of average SNR. This is caused by the difference between \( C_{ER}(S) \) in (4) and \( C_{ER,LB}(S) \) in (7), as in Fig. 2. This difference may increase in a correlated fading environment, which results in more performance degradation compared to that of i.i.d. fading channels in Fig. 2. In order to see the performance loss resulting from the antenna selection, we also present the performance of the transmit antenna selection based on the instantaneous channel conditions for different rate allocation (i.e., selection of \( \arg \max \sum_{k=1}^{M_t} \log_2(1 + E_s/(M_t\sigma^2[|\mathbf{H}(S)^H\mathbf{H}(S)|^{-1}]_{k,k})) \)).
[20] shows the best performance among the methods in Fig. 5, however, it requires high computational complexity. In addition, note that the approaches in [19] and [20] require the statistical feedback information of correlation matrices (not simple transmit antenna indices as in our work).

V. CONCLUSIONS

We have derived a low complexity statistical multimode transmit antenna selection algorithm to improve the capacity of a limited feedback MIMO system with linear receivers. We have considered the equal power and equal rate data transmission, which reduces feedback overhead. The proposed technique has been shown to outperform the non-adaptive transmission, which reduces feedback overhead. The proposed transmit antenna selection algorithm to improve the capacity becomes

\[ N \geq M \]

Expression of (10), when \( M = 1, \) it can be easily shown that \( a(1) \) becomes

\[ a(k) = 1 \]

For \( M = 1, \) it can be easily shown that \( f_{\Lambda_{\min}}^{(1)}(\lambda) \) in (9) becomes

\[ f_{\Lambda_{\min}}^{(1)}(\lambda) = e^{-\lambda N_r - 1} a(1) \]  

where \( a(1) = \frac{(N_r - 1)!}{(N_r - 1)!}. \) Prior to the derivation of results for \( M \geq 2, \) we define and obtain the following integral as

\[ I_n(m) = \int_{0}^{\infty} x^{m}(x + \lambda)^{n} e^{-x} \, dx \]

where the binomial theorem is used. Then, \( f_{\Lambda_{\min}}^{(M)}(\lambda) \) for \( M = 2 \) can be calculated as

\[ f_{\Lambda_{\min}}^{(2)}(\lambda) = K_{2,N_r} e^{-2\lambda N_r - 2} I_{N_r - 2}(2) \]

\[ = e^{-2\lambda N_r - 2} \sum_{k=0}^{N_r - 2} a(k) \lambda^k \]  

where

\[ a(k) = \frac{1}{(N_r - 1)!(N_r - 2)!} \binom{N_r - 2}{k} (N_r - k)! \]  

Next, using the symmetry in the integral evaluation, \( f_{\Lambda_{\min}}^{(M)}(\lambda) \) for \( M = 3 \) can be calculated as

\[ f_{\Lambda_{\min}}^{(3)}(\lambda) = \frac{e^{-3\lambda N_r - 3}}{2(N_r - 1)!(N_r - 2)!(N_r - 3)!} \times \left[ I_{N_r - 3}(4) I_{N_r - 3}(2) - (I_{N_r - 3}(3))^2 \right] \]

\[ = \frac{e^{-3\lambda N_r - 3}}{2(N_r - 1)!(N_r - 2)!(N_r - 3)!} \times \left[ \left( \sum_{k=0}^{N_r - 3} \binom{N_r - 3}{k} (N_r - k + 1)! \lambda^k \right) \right. \]

\[ \left. \times \left( \sum_{k=0}^{N_r - 3} \binom{N_r - 3}{k} (N_r - k - 1)! \lambda^k \right) - \left( \sum_{k=0}^{N_r - 3} \binom{N_r - 3}{k} (N_r - k)! \lambda^k \right)^2 \right] . \]  

(A.5)

Then, calculating the product and sum of polynomials in (A.5) yields

\[ f_{\Lambda_{\min}}^{(3)}(\lambda) = e^{-3\lambda N_r - 3} \sum_{k=0}^{2(N_r - 3)} a_k(3) \lambda^k \]  

(A.6)

where

\[ a_k(3) = \frac{1}{2(N_r - 1)!(N_r - 2)!(N_r - 3)!} \times \left[ \min_{i = \max(0,k-N_r+3)}^{\min(k,N_r-3)} \binom{N_r - 3}{k - i} \binom{N_r - 3}{k} \times (N_r - i)!(N_r - (k - i) - 1)!(k - 2i + 1)! \right] . \]  

(A.7)

In (A.6), the equality \[ \sum_{k=0}^{2n} \sum_{i=\max(0,k-n)}^{\min(n,k)} p_kq_{k-i} \lambda^k \]

is used. For \( M = 4, \) after some manipulations, \( f_{\Lambda_{\min}}^{(M)}(\lambda) \) can be calculated as

\[ f_{\Lambda_{\min}}^{(4)}(\lambda) = \frac{e^{-4\lambda N_r - 4}}{12(N_r - 1)!(N_r - 2)!(N_r - 3)!(N_r - 4)!} \times \left[ I_{N_r - 4}(6) I_{N_r - 4}(4) I_{N_r - 4}(2) - I_{N_r - 4}(6) I_{N_r - 4}(3)^2 + 2 I_{N_r - 4}(5) I_{N_r - 4}(4) I_{N_r - 4}(3) - (I_{N_r - 4}(5))^2 I_{N_r - 4}(2) - (I_{N_r - 4}(4))^3 \right] . \]  

(A.8)

Then, after some tedious calculations of the polynomial part in (A.8), (A.8) can be simplified to

\[ f_{\Lambda_{\min}}^{(4)}(\lambda) = e^{-4\lambda N_r - 4} \sum_{k=0}^{3(N_r - 4)} a_k(4) \lambda^k \]  

(A.9)

where

\[ a_k(4) = \frac{1}{12(N_r - 1)!(N_r - 2)!(N_r - 3)!(N_r - 4)!} \times \left[ \min_{j=\max(0,k-N_r+4)}^{\min(2(N_r-4),4)} \binom{N_r - 4}{k - j} (N_r - k + j)! \times \min_{i=\max(0,j-N_r+4)}^{\min(N_r-4)} \binom{N_r - 4}{i - j} (N_r - 4 - i)! \right] . \]  

(A.10)
\[ \times (N_r - i - 1)! (N_r - (j - i) - 2)! A(i, j, k, N_r) \]

(A.10)

and \( A(i, j, k, N_r) = \left[ (N_r + 1 - k + j)(j + i - k + 2) + (N_r - i)(N_r - j + i - 1) \right] (j - 2i + 1) - (N_r + 1 - k + j)(N_r - i) \).

Thus, for \( M_t = 1, 2, 3, \) and \( 4 \), we obtain

\[ c^{(M_t)}_n = \begin{cases} a_{n-(N_r-M_t)}^1, & \text{for } N_r - M_t \leq n \leq M_t(N_r - M_t) \\ 0, & \text{otherwise.} \end{cases} \]

(A.11)

**APPENDIX B**

**DERIVATION OF INTEGRAL TERM IN (11)**

In this Appendix, we derive the integral term in (11). By using the substitution \( t = 1 + \rho(S)\lambda \), this integral can be calculated as

\[ \int_0^\infty \log_2 (1 + \rho(S)\lambda) e^{-M_t\lambda} \lambda d\lambda = \frac{e^{\mu(S)}}{(\ln 2)\rho(S)^{n+1}} \int_1^\infty (t - 1)^n (\ln t) e^{-\mu(S)t} dt \]

\[ = \frac{e^{\mu(S)}}{(\ln 2)\rho(S)^{n+1}} \sum_{k=0}^n \left( \begin{array}{c} n \\ k \end{array} \right) (-1)^{n-k} \int_1^\infty t^k (\ln t) e^{-\mu(S)t} dt \]

(B.1)

where \( \mu(S) = M_t/\rho(S) \). Next, the integral term of the last result of (B.1) is derived as

\[ X_k \triangleq \int_1^\infty t^k (\ln t) e^{-\mu(S)t} dt \]

\[ = \frac{1}{\mu(S)} \int_1^{\infty} t^{k-1} e^{-\mu(S)t} dt + \frac{k}{\mu(S)} \int_1^{\infty} t^{k-1} (\ln t) e^{-\mu(S)t} dt \]

\[ = \frac{1}{\mu(S)^{k+1}} \Gamma(k, \mu(S)) + \frac{k}{\mu(S)} X_{k-1} \]

(B.2)

where the integration by parts formula is used and \( \Gamma(\alpha, x) \triangleq \int_x^\infty e^{-t} t^{\alpha-1} dt \) denotes the incomplete gamma function [23]. Expanding (B.2) and using \( X_0 = \mu(S)^{-1} \Gamma(0, \mu(S)) \), it can be shown that

\[ X_k = \frac{1}{\mu(S)^{k+1}} \sum_{i=0}^k \frac{k!}{(k - i)!} \Gamma(k - i, \mu(S)). \]

Thus, by substituting (B.3) into (B.1), we get

\[ \int_0^\infty \log_2 (1 + \rho(S)\lambda) e^{-M_t\lambda} \lambda d\lambda = \frac{(-1)^n e^{\mu(S)}}{(\ln 2)\rho(S)^{n+1}} \]

\[ \times \sum_{k=0}^n \left[ (-1)^k \frac{n!}{(n-k)!} \frac{1}{\mu(S)^{k+1}} \sum_{i=0}^k \frac{\Gamma(i, \mu(S))}{i!} \right]. \]

(B.4)