Report

Tianpei Chen

Department of Electrical and Computer Engineering

July 15, 2014

1 System Model

Consider a MIMO system with N_r receive antennas and N_t transmit antennas, where N_r is large, for simplification, we assume $N_r \to \infty$, the corresponding discrete time model is given:

$$y = \mathbf{H}s + n \tag{1}$$

$s \in \mathbb{C}^{N_t \times 1}$ is the transmit symbol vector, $y \in \mathbb{C}^{N_r \times 1}$ is the receive symbol vector, $\mathbf{H} \in \mathbb{C}^{N_r \times N_t}$ is the channel propagation matrix with independent identity distributed(i.i.d) Circularly Symmetric Complex Gaussian(CSCG) zero mean elements of unit variance (Rayleigh fading), $n \in \mathbb{C}^{N_r \times 1}$ is the additive white Gaussian noise (AWGN) with zero mean elements of variance σ^2, $\varepsilon(ss^H) = E_s\mathbf{I}_{Nt}$, E_s denotes the average symbol energy, ε denotes the expectation operation and M denotes the constellation size.

2 Diversity Maximization Antenna Selection Criterion

First we consider a general case to illustrate how the diversity maximization antenna selection scheme works, $N_t - N$ denotes the number of selected antennas and $\mathbf{H_j} \in \mathbb{C}^{N_r \times (N_t - N)}$ denotes the associated propagation matrix, $j \in \left[1, \binom{N_t}{N}\right]$, $s_d \in \mathbb{C}^{(Nt-N)\times 1}$ denotes the the transmit symbol vector, $y_d \in \mathbb{C}^{Nr\times 1}$ denotes the received symbol vector, $n \in \mathbb{C}^{Nr\times 1}$ denotes the Additive White Gaussian Noise (AWGN) with zero mean and variance σ^2 hence

$$y_d = \mathbf{H}_j s_d + n \tag{2}$$

matrix $\mathbf{G}_j \in \mathbb{C}^{(Nt-N)\times Nr}$ denote the linear equalizer(ZF/MMSE), the linear estimates after equalization $\hat{s_d}$ is expressed by:

$$\hat{s_d} = \mathbf{G}_j y_d = \mathbf{G}_j \mathbf{H}_j s_d + \mathbf{G}_j n \tag{3}$$

From (3) we can see, the equalizer colors the noise, the off-diagonal elements of $\mathbf{G}_j\mathbf{H}_j$ contributes to spatial inter symbol interference, vector $\mathbf{G}_j n$, contributes to additional noise, from (3) we have:

$$[\hat{s_d}]_m = (gh)_m s_d + g_m n = g_m h_m [s_d]_m + \sum_{f \neq m} g_m h_f [s_d]_f + g_m n \tag{4}$$

where $[\hat{s_d}]_m$ denotes the mth element of $\hat{s_d}$, $(gh)_m$ denotes the mth row of $\mathbf{GH}_j$, g_m denotes the mth row of $\mathbf{G}_j$ and h_m denotes the mth column of $\mathbf{H}_j$, The post processing SINR of mth sub datastream and jth selected antenna subset is defined as [1]:

$$\rho_m^j = \frac{E_s |g_m h_m|^2}{\sigma^2 \parallel g_m \parallel^2 + E_s \sum_{j \neq m} |g_m h_j|^2} \tag{5}$$

E_s and σ are defined in section 1, $\sigma^2 ||g_m||^2$ is the energy of noise after equalization, which can be derived as follow:

$$\varepsilon[|g_m n|^2] = \varepsilon[(g_m n)(g_m n)^H] = \varepsilon[g_m n n^H g_m^H] = g_m \varepsilon[n n^H] g_m^H = g_m \sigma^2 I g_m^H = \sigma^2 \parallel g_m \parallel^2 \tag{6}$$

as to $(gh)_m$, we have:

$$\varepsilon[(gh)_m s_d s_d^H (gh)_m^H] = (gh)_m \varepsilon[s_d s_d^H](gh)_m^H = (gh)_m E_s I (gh)_m^H = E_s ||(gh)_m||^2 \tag{7}$$

In (7), $E_s||(gh)_m||^2 = E_s|g_m h_m|^2 + E_s \sum_{j\neq m} |g_m h_j|^2$, $E_s|g_m h_m|^2$ is the post processing energy of detected symbol, $E_s \sum_{j\neq m} |g_m h_j|^2$ is the energy of spatial inter symbol interference.

Considering zero forcing algorithm first, the spatial equalizer $(\mathbf{G}_j)_{ZF} = (\mathbf{H}_j^H \mathbf{H}_j)^{-1}\mathbf{H}_j^H$, therefore from (3), $\mathbf{G}_j\mathbf{H}_j = \mathbf{I}_{N_t-N}$ and

$$(\rho_{ZF}^j)_m = \frac{E_s}{\sigma^2 \parallel g_m \parallel^2} \tag{8}$$

Let ω_m denote the mth row of $(\mathbf{H_j^H H_j})^{-1}$, we have

$$\parallel g_m \parallel^2 = g_m g_m^H = \omega_m \mathbf{H}_j^H \mathbf{H}_j \omega_m^H \tag{9}$$

In (9) we have $(\mathbf{H}_j^H \mathbf{H}_j)^{-1}\mathbf{H}_j^H \mathbf{H}_j = \mathbf{I}_{N_t-N}$, hence $\omega_m \mathbf{H_j}^H \mathbf{H_j} = e_m$, where e_m deontes the row vector with mth element equals to 1, the others are all 0. Therefore (9) can be written as:

$$\parallel g_m \parallel^2 = ((\mathbf{H}_j^H \mathbf{H}_j)_{mm}^{-1})^H \tag{10}$$

$((\mathbf{H}_j^H\mathbf{H}_j)_{mm}^{-1})^H$ denotes the mth diagonal element of $((\mathbf{H}_j^H\mathbf{H}_j)^{-1})^H$. Because $\mathbf{H}_j^H\mathbf{H}_j$ is a Hermitian matrix, thus $((\mathbf{H}_j^H\mathbf{H}_j)^{-1})^H = (\mathbf{H}_j^H\mathbf{H}_j)^{-1}$ [2] and (8) can be written as:

$$(\rho_{ZF}^j)_m = \frac{E_s}{\sigma^2(\mathbf{H}_j^H\mathbf{H}_j)_{mm}^{-1}} \tag{11}$$

Consider minimum mean square error (MMSE) detection, where $\mathbf{G} = \mathbf{G}_{MMSE} = (\mathbf{H}_j^H\mathbf{H}_j + \frac{\sigma^2}{E_s}\mathbf{I})^{-1}\mathbf{H}_j^H$, from [3] [4], we have the expression of the post processing SINR of MMSE:

$$(\rho_{MMSE}^j)_m = \frac{E_s}{\sigma^2(\mathbf{H}_j^H\mathbf{H}_j + \frac{\sigma^2}{E_s}\mathbf{I})_{mm}^{-1}} - 1 \tag{12}$$

where $(\mathbf{H}_j^H\mathbf{H}_j + \frac{\sigma^2}{E_s}\mathbf{I})_{mm}^{-1}$ denotes the mth diagonal element of $(\mathbf{H}_j^H\mathbf{H}_j + \frac{\sigma^2}{E_s}\mathbf{I})^{-1}$. The criterion of the diversity maximization selection scheme is to choose the antenna subset which can maximize the worst post-processing SINR of the sub-datastream. The column elements of $\mathbf{H}_j$ are mutually independent Gaussian random variables, under favorable propagation condition, when $N_r \to \infty$ and the 2 norm of the column will become much larger than the inner product of two different columns. Hence, the columns of $\mathbf{H}_j$ become asymptotically orthogonal, a phenomenon that can be expressed as

$$\lim_{N_r \to \infty} \frac{(\mathbf{H}_j^H \mathbf{H}_j)}{N_r} = \mathbf{I}_{N_t - N} \tag{13}$$

Where $\mathbf{I}_{N_t-N}$ denotes an $(N_t - N) \times (N_t - N)$ identity matrix. Based on this property, for any m, $(\rho^j_{MMSE})_m$ and $(\rho^j_{ZF})_m$ in (11)(12) are all the same, meaning that we can select arbitrarily the $N_t - N$ antenna subset.

3 Complexity Analysis of sel-MMSE and sel-MMSE-OSIC algorithm

3.1 sel-MMSE

Define the cost of complex multiplication as 6 flops and complex addition as 2 flops, the sel-MMSE algorithm can be divided into 4 steps [5]:

1. Antenna selection

Based on diversity maximization antenna selection criterion, divide the detection task into two parts:

$$y = \mathbf{H}_1 s_1 + \mathbf{H}_2 s_2 + n \tag{14}$$

Hence there are $N_u = \binom{N_t}{N}$ possible antenna subsets, $\mathbf{H}_1 \in \mathbb{C}^{N_r \times N}$ and $\mathbf{H_2} \in C^{N_r \times (N_t - N)}$.

2. As to $s_1 \in \mathbb{C}^{N \times 1}$, we choose every possible symbol vector s_1 in the lattice and calculate:

$$y^k = y - \mathbf{H}_1 s_1^k \tag{15}$$

s_1^k denotes the kth possible $N \times 1$ symbol vector that is chosen in channel partition stage.

For ML detection:

$$x^{ML} = arg \min_{s} \parallel y - \mathbf{H}s \parallel^2 \tag{16}$$

based on (14), (16) can be changed to

$$arg \min_{s_1, s_2} \parallel y - \mathbf{H}_1 s_1 - \mathbf{H}_2 s_2 \parallel \tag{17}$$

At this stage sel-MMSE uses brute-force search, take every possible $s_1^k \in \mathbb{C}^{N \times 1}$ vector according to signal constellation into consideration, therefore there are M^N possible s_1^k

3. For every s_1^k calculate $y_k = y - \mathbf{H}_1 s_1^k$. There are M^N possible s_1^k as well as y^k. Using linear detection or OSIC to estimate s_2^k, the estimation is expressed as $\hat{s_2}^k$, then mapping $\hat{s_2}^k$ to the constellation to the nearest points, we use s_2^k to represent it, every s_1^k corresponds to one s_2^k, hence there are M^N posssible s_2^k.

4. Find the solution $[s_1^k, s_2^k]$ among M^N candidates that has the minimum total Euclidean distance.

$$\hat{s} = arg \min_{k=1,2,\ldots,M^N} \parallel y^k - \mathbf{H}_2 s_2^k \parallel^2 \tag{18}$$

According to section 2, based on diversity maximization antenna selection criterion and channel hardening phenomenon we can choose an arbitrary antenna subset, and hence the cost of antenna selection is zero. The second step calculating $y^k = y - \mathbf{H}_1 s_1^k$, including an matrix multiplication and vector substruction,

$$N_r N \tag{19}$$

complex multiplications and

$$N_r N \tag{20}$$

complex substitution are required. Then calculate the MMSE estimation $\hat{s}_2^k$:

$$\hat{s}_2^k = [\mathbf{H}_2^H \mathbf{H}_2 + \frac{\sigma^2}{E_s}\mathbf{I}]^{-1} \mathbf{H}_2^H y^k \tag{21}$$

$\mathbf{H}_2 \in \mathbb{C}^{N_r \times (N_t - N)}$, the calculation of weight matrix $\mathbf{G}_{MMSE} = [\mathbf{H}_2^H \mathbf{H}_2 + \frac{\sigma^2}{E_s}\mathbf{I}]^{-1} \mathbf{H}_2^H$, because of the channel hardening phenomenon mentioned in section 2, $\mathbf{G}_{MMSE} = [\mathbf{H}_2^H \mathbf{H}_2 +$

$\frac{\sigma^2}{E_s}\mathbf{I}]^{-1}\mathbf{H}_2^H = \frac{1}{N_r+\frac{\sigma^2}{E_s}}\mathbf{I}\mathbf{H}_2^H$, hence the calculation of $\mathbf{G}_{MMSE}$ is real-complex multiplication, which cost 2 flops, 1/3 of complex multiplication, hence get $\mathbf{G}_{MMSE}$ requires:

$$\frac{(N_t - N)N_r}{3} \tag{22}$$

complex multiplication.

In the whole process $\mathbf{G}_{MMSE}$ only need to be calculated one time. Calculation of $\hat{s}_2^k = \mathbf{G}_{MMSE}y_k$ is the complex matrix multiplication which requires:

$$(N_t - N)N_r \tag{23}$$

complex multiplications and

$$(N_t - N)(N_r - 1) \tag{24}$$

complex additions. Finally the calculation of MED in (18) requires matrix multiplication, vector substruction which cost $N_r(N_t - N)$ complex multiplications and $N_r(N_t - N)$ complex additions and the calculation of 2-norm, which includes the calculation of absolute value and real number addition, calculation of absolute value cost 3 flops, which is 1/2 of the cost of complex multiplication, real number addition cost 1 flop which is

1/2 of the cost of complex addition. Hence the calculation of 2-norm costs $\frac{N_r}{2}$ complex multiplications and $\frac{N_r-1}{2}$ complex additions. Totally, the calculation of MED costs:

$$N_r(N_t - N) + \frac{N_r}{2} \tag{25}$$

complex multiplications and

$$N_r(N_t - N) + \frac{N_r - 1}{2} \tag{26}$$

complex additions are required. All the above are the cost of single candidate $[s_1^k, s_2^k]$, Totally as mentioned before there are M^N possible solution candidates $[s_1^k, s_2^k]$, therefore the sel-MMSE requires overall adding (19)(22)(23)(25), please notice (22) only need to be calculated one time in the whole process.

$$M^N(N_rN+(N_t-N)N_r+N_r(N_t-N)+\frac{N_r}{2})+\frac{(N_t-N)N_r}{3} = M^N(2N_rN_t-N_rN+\frac{N_r}{2})+\frac{(N_t-N)N_r}{3} \tag{27}$$

complex multiplications and adding (20)(24)(26)

$$M^N(N_rN+(N_t-N)(N_r-1)+N_r(N_t-N)+\frac{N_r-1}{2}) = M^N(2N_rN_t-N_rN-N_t+N+\frac{N_r-1}{2}) \tag{28}$$

complex additions, Based on (27)(28) and with 6 flops for complex multiplication and 2 flops for complex addition, we have the total cost of sel-MMSE algorithm as

$$
\begin{aligned}
f_{sel-MMSE}(N) &= 6[M^N(2N_rN_t - N_rN + \frac{N_r}{2}) + \frac{(N_t - N)N_r}{3}] \\
&+ 2M^N(2N_rN_t - N_rN - N_t + N + \frac{N_r - 1}{2}) \qquad (29) \\
&= M^N(16N_rN_t - 8N_rN + 4N_r - 2N_t + 2N - 1) + 2(N_t - N)N_r \quad flops
\end{aligned}
$$

According to [5], the complexity of sel-MMSE algorithm without making use of channel hardening phenomenon is

$$
\begin{aligned}
f_{sel-MMSE_O}(N) &= N_u[4(N_t - N)^3 + (N_t - N)^2(4N_r + 1) + (N_t - N)(4N_r - 6) + 6] + N_r(N_t - N) \\
&\quad (8N_t - 8N - 2) + M^N(16N_rN_t - 8N_rN + 4N_r - 1 - 2N_t + 2N) \quad flops \qquad (30)
\end{aligned}
$$

where $N_u = \binom{N_t}{N}$. Table 1 shows the complexity comparison between $f_{sel-MMSE}$and$f_{sel-MMSE_O}$, at channel partition stage they all use diversity maximization scheme in [5] to determine N. From table 1 we can see, as the size of array increases, the complexity abbreviation gain increases, as the size of signal constellation increases, the complexity abbreviation gain decreases.

$\frac{f_{sel-MMSE}}{f_{sel-MMSE_O}}$ $N_r \times N_t$ / M	10×10	32×32	64×64	128×128
BPSK	0.0249	1.3746e-05	2.3406e-09	1.5672e-14
QPSK	0.1683	4.3817e-04	5.9893e-07	3.2095e-11
16QAM	0.9282	0.3098	0.0378	1.3460e-04
64QAM	0.9988	0.9978	0.9996	0.9982

Table 1: complexity comparison between $f_{sel-MMSE}$ and $f_{sel-MMSE_O}$

3.2 sel-MMSE-OSIC

First we use a general case to illustrate how MMSE-OSIC works. Consider a MIMO system as mentioned in section 1:

$$y = \mathbf{H}s + n \tag{31}$$

With $\mathbf{G}_{MMSE} = (\mathbf{H}^H\mathbf{H} + \frac{\sigma^2}{E_s}\mathbf{I})^{-1}\mathbf{H}^H$, after equalization using $\mathbf{G}_{MMSE}$, (31) can be changed to

$$\hat{s} = \mathbf{G}_{MMSE}y = \mathbf{G}_{MMSE}Hs + \mathbf{G}_{MMSE}n \tag{32}$$

where $\hat{s}$ denote the unconstrained estimate of s. From (32) we can express mth estimation of sub-datastream $\hat{s}_m$:

$$\hat{s}_m = g_m y \tag{33}$$

where g_m denote the mth row of $\mathbf{G}_{MMSE}$, unlike MMSE, the MMSE-OSIC algorithm detects the s_m symbol by symbol. The detection order at every iteration is based on the following principle: $\hat{s_m}$ is detected based on their post-processing SINR at each iteration, where $\hat{s}_m$ with the highest SNR will be detected. According to (12), the post processing SINR is associated to the diagonal elements of $(\mathbf{H}^H\mathbf{H} + \frac{\sigma^2}{E_s}\mathbf{I})^{-1}$. The smaller is the mth diagonal element of $(\mathbf{H}^H\mathbf{H} + \frac{\sigma^2}{E_s}\mathbf{I})^{-1}$, the higher is the post-processing SINR of $\hat{s}_m$. Because there are N_t symbols in vector s, hence MMSE-OSIC algorithm will detect s_m for N_t times, in the ith ($i \in [0, N_t - 1]$) iteration, it majorly has 3 steps:

1.Ordering the channel based on the principle mentioned above, choose the sub-datastream with highest SNR(smallest diagonal element of $((\mathbf{H}_j^i)^H\mathbf{H}_j^i + \frac{\sigma^2}{E_s}\mathbf{I})^{-1}$ to detect, $\mathbf{H}_j^i \in \mathbb{C}^{Nr\times(Nt-N-i)}$ denotes the ith step sub-matrix in MMSE-OSIC iteration from the jth propagation matrix after channel partition as mentioned in section 2, $\mathbf{H}_j^0 = \mathbf{H}_j$

2.Find the unconstrained estimate

$$\hat{s}_{i+1} = g_{i+1}y_i \tag{34}$$

and slice $\widehat{s_{i+1}}$ to $\bar{s}_{i+1}$ corresponds to signal constellation.

3.Cancel the information correlated to $\bar{s}_i$ from received vector y_i, $h_m \quad m \in [1, N_t]$ denotes the mth column of $\mathbf{H}_j^i$, s_{2m} denotes the mth symbol of s_2. We have

$$y_k = \sum_{m=1,2,\ldots Nt-N} h_m s_{2m} + n \tag{35}$$

Therefore the cancellation step is removing $\bar{s}_{i+1}$ and corresponding channel column h_{i+1}

$$y_{i+1} = y_i - h_{i+1}\bar{s}_{i+1} \tag{36}$$

Then update propagation matrix by removing the column h_{i+1} from $\mathbf{H}_j^i$.

Consider now the computational complexity of sel-MMSE-OSIC. The difference between sel-MMSE-OSIC and sel-MMSE is in the linear detection step to detect s_2 in $y = \mathbf{H}_1 s_1 + \mathbf{H}_2 s_2 + n$, so that their computation cost except this step are all the same. Firstly finding y^k in (15) requires $N_r N$ complex multiplications and $N_r N$ complex additions. In MMSE-OSIC detection of s_2^k, when $N_r \to \infty$, because of channel hardening phenomenon, at ith iteration($i \in [0, \quad N_t - N - 1]$) in step 1, the diagonal elements of $(\mathbf{H}_j^i)^H \mathbf{H}_j^i + \frac{\sigma^2}{E_s}\mathbf{I})^{-1}$ are all the same. Therefore the post-processing SINR of $\hat{s_m}$ are all the same, detection order can be arbitrary, the step 1 in MMSE-OSIC can be ignored. in the second step, based on (33) and channel hardening phenomenon, $\mathbf{G}_{MMSE} = \frac{1}{N_r + \frac{\sigma^2}{E_s}}(\mathbf{H}_j^i)^H$, $\upsilon_m = \frac{1}{N_r + \frac{\sigma^2}{E_s}}\theta_m^H$, where θ_m^H denotes the mth row of $(\mathbf{H}_j^i)^H$, hence finding the the weight vector requires a real-complex vector multiplication, with cost of $\frac{N_r}{3}$ complex multiplications, totally $N_t - N$ weight vectors need to be calculated, therefore $(N_t - N)\frac{N_r}{3}$ complex multiplications are needed, furthermore based on (34) getting the detection statistic $\hat{s}_i$ requires N_r complex multiplications and $N_r - 1$ complex additions, then the interference cancellation step in

(36) requires N_r complex multiplications and N_r complex additions. Since the MMSE-OSIC algorithm performs $N_t - N$ iteration, the cost of getting $\hat{s}_2$ is:

$$M^N(2N_r(N_t - N)) + (N_t - N)\frac{N_r}{3} \tag{37}$$

complex multiplications and

$$M^N(2N_r - 1)(N_t - N) \tag{38}$$

complex additions.

Basing on (36), after $N_t - N$ iterations $y^k - \mathbf{H}_2 s_2^k$ has already been calculated, therefore finding MED in (18) just requires additional

$$N_r/2 \tag{39}$$

complex multiplications and

$$(N_r - 1)/2 \tag{40}$$

complex additions, the total cost:

Adding (19)(37)(39)

$$M^N(N_rN + 2N_r(N_t - N) + \frac{N_r}{2}) + (N_t - N)\frac{N_r}{3} \tag{41}$$

complex multiplications and adding (20)(38)(40)

$$M^N(N_rN + (2N_r - 1)(N_t - N) + \frac{N_r - 1}{2}) \tag{42}$$

complex addtions. The computation cost in flops, based on (41)(42) is

$$\begin{aligned}
f_{sel-MMSE-OSIC}(N) &= 6[M^N(2N_rN_t - N_rN + \frac{N_r}{2}) + \frac{(N_t - N)N_r}{3}] && (43)\\
&+ 2M^N(2N_rN_t - N_rN - N_t + N + \frac{N_r - 1}{2}) && (44)\\
&= M^N(16N_rN_t - 8N_rN + 4N_r - 2N_t + 2N - 1) + 2(N_t - N)N_r \quad flops
\end{aligned}$$

From (30)(45), $f_{sel-MMSE-OSIC}(N) = f_{sel-MMSE}(N)$.

According to [5], the complexity of sel-MMSE-OSIC algorithm without making use of

$\frac{f_{sel-MMSE-OSIC}}{f_{sel-MMSE-OSIC_O}}$ $N_r \times N_t$ / M	10×10	32×32	64×64	128×128
BPSK	0.0250	1.3746e-05	2.3406e-09	1.5672e-14
QPSK	0.1687	4.3817e-04	5.9893e-07	3.2095e-11
16QAM	0.9284	0.3098	0.0378	1.3460e-04
64QAM	0.9988	0.9978	0.9996	0.9982

Table 2: complexity comparison between $f_{sel-MMSE-OSIC}$ and $f_{sel-MMSE-OSIC_O}$

channel hardening phenomenon is

$$\begin{aligned} f_{sel-MMSE-OSIC_O}(N) &= N_u[4(N_t-N)^3+(N_t-N)^2(4N_r+1)+(4N_r-6)(N_t-N)+6] \\ &+ 4/3(N_t-N)^3+(N_t-N)^2(4N_r+3)+(N_t-N)(2N_r-13/3) \\ &+ M^N(16N_rN_t-8N_rN+4N_r-2N_t+2N-1) \quad flops \end{aligned} \tag{45}$$

Table 2 shows the complexity comparison between $f_{sel-MMSE}$and$f_{sel-MMSE_O}$, at channel partition stage they all use diversity maximization scheme in [5] to determine N.

From table 2 we can see, as the size of array increases the complexity abbreviation gain increases, as the size of signal constellation increases, the complexity abbreviation gain decreases.

4 Minimization of the Complexity

Consider $N_r \leq 2 \quad N_t \leq 2 \quad M \leq 2$.

4.1 sel-MMSE

Proposition 1: $f_{sel-MMSE}(N)$ increases when N increases, in $[1, N_t - 1]$

Proof: According to section(3), the computational cost of sel-MMSE algorithm is:

$$f_{sel-MMSE}(N) = M^N(16N_rN_t - 8N_rN + 4N_r - 2N_t + 2N - 1) + 2(N_t - N)N_r \quad flops \tag{46}$$

In this section a discussion of the relation between N and $f(N)$ is presented. Firstly we assume N is a real number, take the derivative of $f_{sel-MMSE}(N)$,

$$\frac{\partial f_{sel-MMSE}(N)}{\partial N} = M^N[(16N_rN_t + 4N_r - 2N_t - 1)\ln M + (2 - 8N_r)(N \ln M + 1)] - 2N_r \tag{47}$$

according to (47), obviously when $N \geq 0$, $\frac{\partial f_{sel-MMSE}(N)}{\partial N}$ increases with N increasing, consider $N = 0$, we have:

$$\begin{aligned}
\frac{\partial f_{sel-MMSE}(0)}{\partial N} &= \ln M(16N_rN_t + 4N_r - 2N_t - 1) + 2 - 10N_r \\
&= \ln M[4(N_t + 1)N_r + 2(N_r - 1)N_t] + 10N_r(N_t \ln M - 1) + 2 - \ln M
\end{aligned}$$

because $N_r \geq 2$ $N_t \geq 2$ $M \geq 2$, therefore $\ln M[4(N_t - 1)N_r + 2(N_r - 1)N_t] > 0$, $\ln M \geq 0.6931$, $N_t \ln M \geq 1.3863$, hence $10N_r(N_t \ln M - 1) + 2 > 9.7259$, if $M < 1.6746 \times 10^4$,

$10N_r(N_t \ln M - 1) + 2 - \ln M > 0$, for the signal constellation size we consider, we can conclude that $10N_r(N_t \ln M - 1) + 2 - \ln M > 0$ holds, therefore $\frac{\partial f_{sel-MMSE(0)}}{\partial N} > 0$.

In conclusion when $N \geq 0$, $\frac{\partial f_{sel-MMSE(N)}}{\partial N} > 0$, proposition 1 is proved.

4.2 sel-MMSE-OSIC

Proposition 2: $f_{sel-MMSE-OSIC}(N)$ increases when N increases, when $N \in [1, N_t - 1]$

Proof: From section (3), the complexity expressions of sel-MMSE and sel-MMSE-OSIC are the same, therefore, proposition 2 holds.

5 Trade off between Complexity and diversity

Sel-MMSE and sel-MMSE-OSIC can achieve optimal diversity gain N_r asymptotically [4] [6] with $SNR \to \infty$ in a $N_r \times N_t$ spatial multiplexing system [5], however the most of practical communication systems are working under a moderate SNR, further more the optimal diversity gain of massive MIMO system is extensively large which is unnecessary, on the other hand, achieving the asymptotically optimal diversity gain will lead to extremely high complexity, therefore we consider sacrifice redundant diversity gain and abbreviate complexity.

let P_e denotes the average probability of detection error of sel-MMSE and sel-MMSE-

OSIC, P_e^{ML} denotes the average probability of detection error of Maximum-Likelihood detection, $(P_e)_2$ denotes the average probability of detection error of s_2^k after channel partition, P_e is bounded by [5]:

$$\max[(P_e)_2, P_e^{ML}] \leq P_e \leq P_e^{ML} + (P_e)_2 \tag{48}$$

The diversity gain of sel-MMSE and its OSIC counterpart are expressed by d_{MMSE} and $d_{MMSE-OSIC}$ diversity maximization selection, they can be expressed as [5]

$$d_{MMSE} = (N+1)(N_r - N_t + N + 1) \tag{49}$$

$$d_{MMSE-OSIC} = (N+1)(N_r - N_t + N + 1) \tag{50}$$

(49) is proven in [6], in [5], $d_{MMSE-OSIC} \geq (N+1)(N_r - N_t + N + 1)$ is proven however strictly proven of (50) is still lacking. therefore we take (50) as a conjecture here.

$$\lim_{SNR \to \infty} \frac{\log P_e^{ML}}{\log SNR} = -N_r \tag{51}$$

$$\lim_{SNR \to \infty} \frac{\log (P_e)_2}{\log SNR} = -d \tag{52}$$

let $(P_e)_2^{LD}$ denote the average probability of detection error of s_2^k using linear detection. If $d \geq N_r$, sel-MMSE can achieve optimal performance asymptotically $\lim_{SNR\to\infty} P_e = P_e^{ML}$. If $d < N_r$, $\lim_{SNR\to\infty} \frac{\log P_e}{\log SNR} = d$ [5]. $0 \leq x \leq N_r$ denote the minimum required diversity.

Define

$$g(N) = d - x = (N+1)(N_r - N_t + N + 1) - x \quad 1 \leq N \leq N_t - 1 \tag{53}$$

Based on the results in section 4, we seek minimum N (N_{min}) satisfying $g(N) \geq 0$. We see that $g(N)$ is a quadratic function of N with the two zero points:

$$N_1 = \sqrt{x + 1/4(N_r - N_t)^2} - 1/2(N_r - N_t) - 1 \tag{54}$$

$$N_2 = -\sqrt{x + 1/4(N_r - N_t)^2} - 1/2(N_r - N_t) - 1 \tag{55}$$

Assume $N_r \geq 2$ $N_t \geq 2$

Consider the following conditions:

1. $N_1 > 0 \quad N_2 > 0$

It is obvious that $N_1 > N_2$, thus when $N_2 > 0$ this condition is satisfied.

$$\begin{aligned} N_2 > 0 \;\; &\rightarrow \;\; -\sqrt{x + \frac{1}{4}(N_r - N_t)^2} - \frac{1}{2}(N_r - N_t) - 1 > 0 \\ &\rightarrow \;\; \sqrt{x + \frac{1}{4}(N_r - N_t)^2} < \frac{1}{2}(N_t - N_r) - 1 \end{aligned} \tag{56}$$

(56) holds when the following 3 conditions are satisfied:

$$\begin{cases} \frac{1}{2}(N_t - N_r) - 1 > 0 & \rightarrow N_t > N_r + 2 \\ 0 < x < N_r - N_t + 1 & \rightarrow N_t < N_r + 1 \end{cases}$$

The first and last conditions are contradictory thus $N_1 > 0 \quad N_2 > 0$ does not exit.

2. $N_1 > 0 \quad N_2 < 0$

$$\begin{aligned} N_1 > 0 \;\; &\rightarrow \;\; \sqrt{x + \frac{1}{4}(N_r - N_t)^2} - \frac{1}{2}(N_r - N_t) - 1 > 0 \\ &\rightarrow \;\; x > N_r - N_t + 1 \end{aligned} \tag{57}$$

$$
\begin{aligned}
N_2 < 0 \;\; &\rightarrow \;\; -\sqrt{x+\frac{1}{4}(N_r-N_t)^2}-\frac{1}{2}(N_r-N_t)-1<0 \\
&\rightarrow \;\; x > N_r - N_t + 1
\end{aligned}
\tag{58}
$$

based on (57)(58), when $x > N_r - N_t + 1$, $N_{min} = \lceil N_1 + 1 \rceil = \lceil \sqrt{x + \frac{1}{4}(N_r - N_t)^2} - \frac{1}{2}(N_r - N_t) \rceil$, where $\lceil n \rceil$ denote the largest integer smaller than n.

3. $N_1 \le 0 \quad N_2 \le 0$

$$
\begin{aligned}
N_1 \le 0 \;\; &\rightarrow \;\; \sqrt{x+\frac{1}{4}(N_r-N_t)^2} \le \frac{1}{2}(N_r-N_t)+1 \\
&\rightarrow \;\; x \le N_r - N_t + 1
\end{aligned}
\tag{59}
$$

$$
\begin{aligned}
N_2 \le 0 \;\; &\rightarrow \;\; \sqrt{x+\frac{1}{4}(N_r-N_t)^2} \ge \frac{1}{2}(N_t-N_r)-1 \\
&\rightarrow \;\; x \ge N_r - N_t + 1
\end{aligned}
\tag{60}
$$

Therefore based on (59)(60), if and only if $N_1 = N_2 = 0 \quad x = N_r - N_t + 1$, this condition holds, $N_{min} = 1$.

Conclusion

$$N_{min} = \begin{cases} \lceil \sqrt{x + \frac{1}{4}(N_r - N_t)^2} - \frac{1}{2}(N_r - N_t) \rceil & x > N_r - N_t + 1 \\ 1 & x = N_r - N_t + 1 \end{cases}$$

6 Future Work

if the channel hardening phenomenon is applied directly to sel-MMSE and sel-MMSE-OSIC algorithm. The performance lost are extremely severe, Figure.1 presents the simulated BER performance of sel-MMSE and sel-MMSE-OSIC algorithms with channel hardening phenomenon directly applied, we consider uncoded system with 16QAM modulation, $N_t = N_r = 40$, the transmitted bit sequence is randomly generated and mutually independent, the elements of propagation matrix H is i.i.d Circularly Symmetric Complex Gaussian(CSCG) random variable with zero means and unit variance, magnitude Rayleigh fading. the minimum diversity requirement x is 8, for each SNR point there is a minimum 10^5 symbols transmitted and 500 symbol errors are collected. As Figure.1 shows, there is no diversity gain for sel-MMSE and sel-MMSE-OSIC algorithms with channel hardening phenomenon, however the performance tends to be better when the ratio $\frac{N_r}{N_t}$ is large that is several tens.

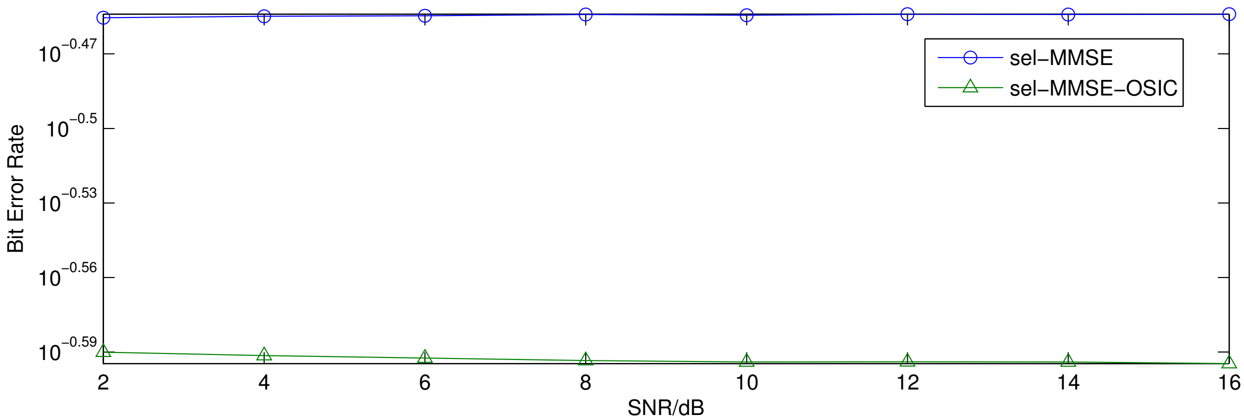

Figure 1: BER performance of sel-MMSE and sel-MMSE-OSIC with channel hardening phenomenon

Consider MMSE algorithm:

$$\hat{s} = \mathbf{G}y = \mathbf{GH}s + \mathbf{G}n \tag{61}$$

the definitions of $\mathbf{H}$ y s n $\hat{s}$ are similar to the definitions in section 1. $G = (\mathbf{H}^H\mathbf{H} + SNR^{-1})^{-1}\mathbf{H}^H$ is the weight matrix, define $\mathbf{C} = (\mathbf{H}^H\mathbf{H} + SNR^{-1})^{-1}$. From (61) we have:

$$\hat{s}_k = c_{kk}\mathbf{H}^H\mathbf{H}s + c_{kk}\mathbf{H}^H n + \sum_{j \neq k} c_{kj}\mathbf{H}^H\mathbf{H}s + \sum_{j \neq k} c_{kj}\mathbf{H}^H n \tag{62}$$

$k \in [1, N_t]$, $\hat{s}_k$ is the kth element of $\hat{s}$, c_{kj} is the element of $\mathbf{C}$ at kth row and jth column.

with channel hardening phenomenon (62) is changed to:

$$\hat{s}_k^{hardening} = \frac{1}{N_r + SNR^{-1}}\mathbf{H}^H\mathbf{H}s + \frac{1}{N_r + SNR^{-1}}\mathbf{H}^H n \tag{63}$$

$\hat{s}_k^{hardening}$ is the MMSE estimation with channel hardening phenomenon applied. comparing (62) and (63), the off diagonal elements of $\mathbf{C}$ is ignored in (63), with the fixed N_r if N_t increases, the number of off diagonal elements of $\mathbf{C}$ increases and performance get worse, so it is not reasonable to ignore them. This explain the result in Figure.1, this kind of inter user interference(IUI) causes performance lost. Therefore, the next step is to find a strategy to eliminate this IUI without increasing too much complexity.

References

[1] R. Heath and A. Paulraj, "Antenna selection for spatial multiplexing systems based on minimum error rate," in *Communications, 2001. ICC 2001. IEEE International Conference on*, vol. 7, 2001, pp. 2276–2280 vol.7.

[2] F. R. Gantmacher, "Matrix theory, vol. i," *IIChelsea, New York*, 1959.

[3] S. Verdu, *Multiuser detection.* Cambridge university press, 1998.

[4] A. Paulraj, R. Nabar, and D. Gore, *Introduction to space-time wireless communications.* Cambridge university press, 2003.

[5] D. Radji and H. Leib, "Interference cancellation based detection for v-blast with diversity maximizing channel partition," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 3, no. 6, pp. 1000–1015, Dec 2009.

[6] J. Jaldén and B. Ottersten, "On the maximal diversity order of spatial multiplexing with transmit antenna selection," *Information Theory, IEEE Transactions on*, vol. 53, no. 11, pp. 4273–4276, 2007.