Transceiver Optimization of Tomlinson-Harashima Precoded MIMO systems: Based on Multiplicative Schur Convexity

Vignesh S*
*IIT Madras
ee16b127@ee.iitm.ac.in

***Abstract*—This report explains the findings and reproduces the results of the paper *Tomlinson-Harashima Precoding in MIMO Systems: A Unified Approach to Transceiver Optimization Based on Multiplicative Schur-Convexity* published by Alberto D'Amico. We consider the design of a multi-carrier multiple input multiple output(MIMO) system which has a non-linear Tomlinson-Harashima precoding (THP) and a linear equalizer. We jointly optimize the operations at the transmitter and receiver with an aim to minimize the Mean-Squared Errors pertaining to each carrier. We examine functions which are Multiplicative Schur convex or Schur concave in these mean-squared errors and obtain closed form solutions for the process matrices. The results suggest that for Multiplicative Schur Concave functions the optimized solution reduces to linear precoding at the Transmitter. As for Multiplicative Schur convex functions we observe that THP provides better results than linear prefiltering.**

***Index Terms*—MIMO, Multiplicative Schur Convex, transceiver optimization, Tomlinson - Harashima Precoding**

I. Introduction

The aim of any communication system is to transfer information error-free from the sender to the receiver within a certain minimum time. The rate of transmission is capped by the Capacity of the channel and the reliability is determined by the Error probability. One has to face a trade-off between having a higher reliability and rate. Multiple Input and Multiple Output (MIMO) systems help us tackle this problem to a certain degree by providing spatial diversity.

The performances of MIMO systems depends significantly on the Multi-stream Interference (MSI) caused due to parallel transmission of data over the same the same frequency band. But this issue can be mitigated if Channel State Information (CSI) is available at the receiver. Signal Processing techniques such as Zero-Forcing (ZF), MMSE estimation which are linear and certain non-linear ones such as V-BLAST are used to remove MSI. Channel Diagonalization which involves processing at both transmitter and receiver, is optimal if the objective functions are Schur convex or Schur concave. THP [1] [2], a non-linear technique used at the transmitter, is employed to remove the interference caused due to the channel. The approach in THP is similar to V-BLAST, but differs from the fact that it is used at the transmitter instead of the receiver and that there is no error propagation involved like in V-BLAST.

Here we consider a multi-carrier communication with a frequency selective MIMO channel. We use THP at the transmitter and linear equalizer at the receiver. We assume an overall power constraint across all the carriers along with perfect CSI at transmitter and receiver. The goal is to jointly optimize the process matrices. In section II we discuss the Signal model and transmitter and the receiver structure, in section III we talk about the optimization problem at hand, in section IV we show the closed form solutions of the optimization problems, in section V we explain multiplicative Schur convexity using some examples, in section VI we briefly discuss the Carrier Cooperative system, section VII concludes the what has done by putting the results together and in the last section (VIII) a critique on the paper.

II. System and Signal Model

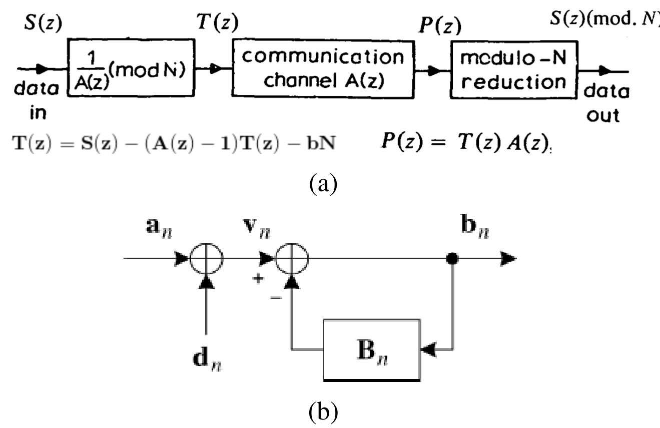


Fig. 1. a) The block diagram of the communication model with signals represented [1]. (b) THP block diagram used in [3]

Figure 1(a) from [1] shows the working of a Tomilinson-Harashima Precoder in a communication channel. The THP removes all the interference caused by the channel due to the previous symbols (using CSI at transmitter). Upon this subtraction there is chance $T(z)$ can go out of power bounds. To prevent this, the subtracted $T(z)$ is passed through a modulo operator before transmitting. The expression for the modulo-M operator used in [3] is given by

$$MOD_M(x) = x - 2\sqrt{M}\left\lfloor \frac{x+\sqrt{M}}{2\sqrt{M}} \right\rfloor$$

Any value x (complex number) upon passing through the modulo operator is made to fall inside a square of side $2\sqrt{M}$, centred at the origin. For example if $x = 1.99\sqrt{M}; MOD_M(x) =$

$-0.01\sqrt{M}$. Hence we see that this is equivalent to adding a $\pm 2n\sqrt{M}$ to the input x, where n is an integer. Figure 1(b) shows us this idea by replacing the modulo operator entirely by a summing operation($+\mathbf{d_n}$) with the input ($\mathbf{a_n}$). $\mathbf{B_n}$ is backward block matrix which subtracts the interference from $\mathbf{b_n}$ which is then transmitted.

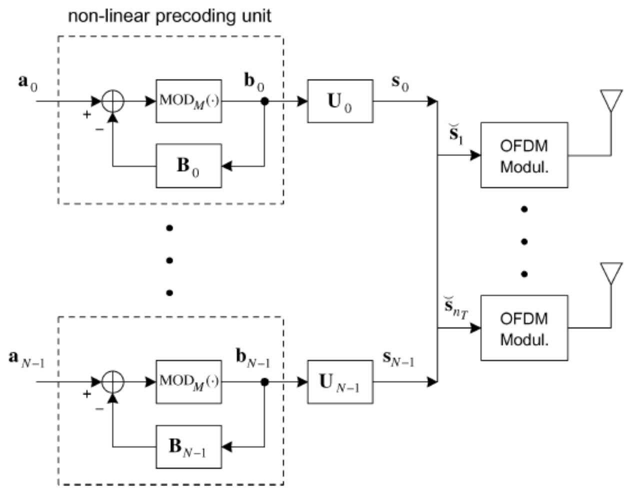


Fig. 2. Transmitter block diagram for Non-cooperative scheme in [3]

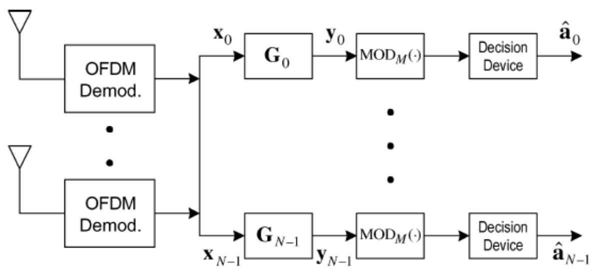


Fig. 3. Receiver block diagram for Non-cooperative scheme in [3]

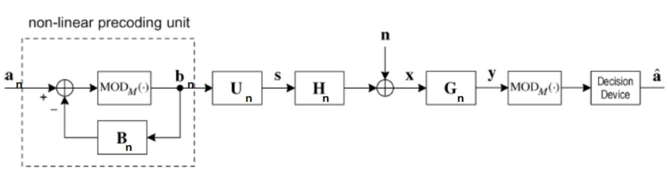


Fig. 4. Block diagram of complete signal model in [3]

The transmitter and the receiver have η_T, η_R antennas respectively. We transmit $K(< min(\eta_T, \eta_R))$ symbols across N sub-carriers each (KN symbols in total). Each of the K symbols of the input vector $\mathbf{a_n} = [a_n(1), a_n(2), ..., a_n(K)]$ are taken from M-QAM constellation with $\sigma_a^2 = \frac{2(M-1)}{3}$. The transmitter structure is as shown in Figure 2. As explained above the role of $\mathbf{B_n}$ is to remove interference from previous symbols, and hence it is evidently a strictly lower triangular matrix. Using the model in Figure 1(b) we write,

$$\mathbf{b_n} = \mathbf{a_n} + \mathbf{d_n} - \mathbf{B_n bn}$$

$$\mathbf{b_n} = \mathbf{C_n^{-1} v_n}$$

where $\mathbf{v_n} = \mathbf{a_n} + \mathbf{d_n}$ and $\mathbf{C_n} = \mathbf{B_n} + \mathbf{I_K}$. Observe that $\mathbf{C_n}$ is a lower triangular matrix with unit diagonal entries. The transmit matrix $\mathbf{U_n}$ converts the sub-carrier input vectors into suitable input vectors for transmit antennas. We use OFDM modulators to transmit across the frequency selective channel and OFDM demodulators to receive the information. The receiver structure is as shown in Figure 3. The observations at the receiving antennas are passed through a receive matrix $\mathbf{G_n}$ which is then passed through the modulo operator to remove the $\mathbf{d_n}$ added at the transmitter. This is then passed through a detector which gives the estimated input. This means that $\mathbf{y_n}$ (as shown in Figure 3) is an estimate of $\mathbf{v_n}$. From the signal model in Figure 4 we write the following,

$$\mathbf{x_n} = \mathbf{H_n s_n} + \mathbf{w_n}$$
$$\mathbf{y_n} = \mathbf{G_n H_n U_n b_n} + \mathbf{G_n w_n}$$

$$\begin{aligned} \mathbf{MSE_{n,k}} &= \mathbf{E}[|(\mathbf{y_n} - \mathbf{v_n})_\mathbf{k}|^\mathbf{2}] \\ &= \mathbf{E}[(\mathbf{y_n} - \mathbf{C_n b_n})(\mathbf{y_n} - \mathbf{C_n b_n})^\mathbf{H}]_\mathbf{k,k} \end{aligned} \quad (1)$$

The noise $\mathbf{w_n}$ is AWGN with $\mathbf{E}[\mathbf{w_n}] = \mathbf{0}$ and $\mathbf{E}[\mathbf{w_n w_n^H}] = \sigma_\mathbf{w}^\mathbf{2} \mathbf{I_{n_R}}$. Also define $\rho = \sigma_w^2/\sigma_a^2$ which will be used later. $\mathbf{b_n}, \mathbf{w_n}$ are independent of each other. The mean-squared error as represented in Equation 1 is defined for each symbol of every sub-carrier. There is independent processing of signals and optimization for each sub-carrier. Such a system is called as Carrier Non-Cooperative scheme.

III. Optimization Problem for Non-Cooperative Scheme

The Non-Cooperative Optimal Power Allocation (NC-OPA) problem to be solved by any multi-carrier communication system is stated below,

$$\min_{P_n} \tilde{f}(P_0, P_1, ...P_{N-1})$$
$$\text{s.t.} \sum_{n=0}^{N-1} P_n = KN \ ; \ P_n \geqslant 0$$

But the above stated problem is very convoluted and does not give the insight required to solve it. We convert it to a form involving the mean-squared errors so that we can use our signal model and simplify the problem. We state the same optimization problem in a different manner as follows,

$$\tilde{f}(P_0, P_1, ...P_{N-1}) = f(\hat{\alpha_0}(P_0), \hat{\alpha_1}(P_1), ..., \hat{\alpha_{N-1}}(P_{N-1}))$$
where $\hat{\alpha_n}(P_n)$ is the optimum solution
of the below problem,
$$\min_{\mathbf{C_n}, \mathbf{U_n}, \mathbf{G_n}} \alpha_n = f_n(MSE_{n,1}, ...MSE_{n,K})$$
$$\text{s.t.} \ \ Tr(\mathbf{U_n U_n^H}) = P_n \ ; \ P_n \geqslant 0$$

The above sub-problem has per carrier power constraint and process matrices(Transmit matrix - $\mathbf{U_n}$, Precoding matrix - $\mathbf{C_n}$ and Receive matrix - $\mathbf{G_n}$)

IV. Optimization of Processing Matrices

We assume that the functions - f_n are increasing in each $MSE_{n,k}$ and hence it is enough to minimize each $MSE_{n,k}$

A. Optimization of $\mathbf{G_n}$

As explained earlier $\mathbf{y_n}$ is a linear estimate of $\mathbf{v_n}$ using the observations $\mathbf{x_n}$. Hence this is a LMMSE (linear minimum mean-squared error) estimation problem. The solution to this is nothing but the Wiener filter. The set of equations below describe the solution.

$$\mathbf{y_n} = \mathbf{\hat{v}_n} = \mathbf{G_n x_n}$$
$$[(\mathbf{G_n})^{\mathbf{H}}_{\mathbf{opt}}]_{:,\mathbf{k}} = \mathbf{R^{-1}_{x_n} R_{x_n v_n}}$$
$$[(\mathbf{G_n})^{\mathbf{H}}_{\mathbf{opt}}]_{:,\mathbf{k}} = (\mathbf{H_n U_n U_n^H H_n^H} + \rho \mathbf{I_{n_R}})^{-1}[\mathbf{H_n U_n C_n^H}]_{:,\mathbf{k}}$$

Below we use results from LMMSE and obtain the minimum error.

$$\mathbf{LMMSE_{n,k}} = \mathbf{E}[|\tilde{\mathbf{x}}|^2] = \mathbf{E}[|(\tilde{\mathbf{v}}_\mathbf{n})_\mathbf{k}|^2]$$
$$\mathbf{LMMSE_{n,k}} = [\mathbf{R_{v_n}}]_{\mathbf{k},\mathbf{k}} - [\mathbf{R_{v_n x_n}}]_{\mathbf{k},:}[(\mathbf{G_n})^{\mathbf{H}}_{\mathbf{opt}}]_{:,\mathbf{k}}$$
$$\mathbf{MSE_{n,k}} = \sigma^2_\mathbf{w}[\mathbf{C_n}(\mathbf{H_n U_n U_n^H H_n^H} + \rho \mathbf{I_{n_R}})^{-1}\mathbf{C_n^H}]_{\mathbf{k},\mathbf{k}}$$

B. Optimization of $\mathbf{C}_n$

We can write the $MSE_{n,k}$ after optimizing $\mathbf{G_n}$ as

$$\mathbf{MSE_{n,k}} = \sigma^2_\mathbf{w}([\mathbf{C_n^H}]_{:,\mathbf{k}})^\mathbf{H}(\mathbf{H_n U_n U_n^H H_n^H} + \rho \mathbf{I_{n_R}})^{-1}[\mathbf{C_n^H}]_{:,\mathbf{k}}$$

We see that it only depends on columns of $\mathbf{C_n^H}$ which are rows of $\mathbf{C_n}$. Observe that $(\mathbf{H_n U_n U_n^H H_n^H} + \rho \mathbf{I_{n_R}})^{-1}$ is positive definite symmetric matrix on which we can apply Cholesky factorization. Upon doing this we get,

$$\begin{aligned}\mathbf{MSE_{n,k}} &= \sigma^2_\mathbf{w}([\mathbf{C_n^H}]_{:,\mathbf{k}})^\mathbf{H}\mathbf{L_n L_n^H}[\mathbf{C_n^H}]_{:,\mathbf{k}} \\ &= \sigma^2_\mathbf{w}||[(\mathbf{C_n L_n})^\mathbf{H}]_{:,\mathbf{k}}||^2 \\ &= \sigma^2_\mathbf{w}\sum_{\mathbf{i=0}}^{\mathbf{k-1}}[\mathbf{L_n}]^2_{\mathbf{i},\mathbf{i}}\ |[(\mathbf{C_n\tilde{L}_n})^\mathbf{H}]_{\mathbf{i},\mathbf{k}}|^2\ + \sigma^2_\mathbf{w}[\mathbf{L_n}]^2_{\mathbf{k},\mathbf{k}}\end{aligned} \tag{2}$$

where $\mathbf{L_n} = \mathbf{\tilde{L}_n D_n}$; $\mathbf{D_n} = \mathbf{diag}\{[\mathbf{L_n}]_{\mathbf{1},\mathbf{1}}, ..., [\mathbf{L_n}]_{\mathbf{K},\mathbf{K}}\}$. $\mathbf{\tilde{L}_n}$ is basically a unit diagonal lower triangular matrix. The only constraint on $\mathbf{C_n}$ is that it has unit diagonal entries(along with being lower triangular). Hence we separate those diagonal product terms as shown in Equation 2. We need to minimize only the summation of the terms of vector $|[(\mathbf{C_n\tilde{L}_n})^\mathbf{H}]_{:,\mathbf{k}}|^2$ only until diagonal (since product of lower triangular matrices is also lower triangular). The value of the first term in equation 2 can be made 0 (which is minimum of the non-negative quantity) when $\mathbf{C_n\tilde{L}_n}$ is made diagonal. But since their product also has unit diagonal elements, it effectively,

$$\mathbf{C_n\tilde{L}_n} = \mathbf{I_K}$$
$$\implies \mathbf{C_n L_n} = \mathbf{D_n}$$
$$\therefore \mathbf{C_n^{opt}} = \mathbf{\tilde{L}_n^{-1}} = \mathbf{D_n L_n^{-1}}$$

If the first term in the equation 2 is 0, then the mean-squared error simplifies to,

$$MSE_{n,k} = \sigma^2_w[\mathbf{L_n}]^2_{\mathbf{k},\mathbf{k}}$$

We see that this was a simple norm optimization problem.

C. Optimization of $\mathbf{U}_n$

Before optimizing $\mathbf{U_n}$ we introduce the concepts of multiplicative Schur convexity.

Definition : Given a vector **a** = $(a_1, ..., a_p)^T \in \mathbb{R}^p_{++}$ we define an order among its terms as follows $a_{[1]} \geqslant a_{[2]} \geqslant .... \geqslant a_{[p]}$

Definition : Consider vectors **a**,**b** $\in \mathbb{R}^p_{++}$, we say $\mathbf{a}$ is multiplicatively majorized by $\mathbf{b}$ written as $\mathbf{a} \prec_{\mathbf{\Pi}} \mathbf{b}$ if

$$\prod_{i=1}^{k} a_{[i]} \leqslant \prod_{i=1}^{k} b_{[i]} \quad \text{for } 1 \leqslant k \leqslant p-1$$
$$\prod_{i=1}^{p} a_{[i]} = \prod_{i=1}^{p} b_{[i]}$$

Consider an example where **a** = $(2, 3, 2)^T$ and **b** = $(1, 6, 2)^T$. It can be observed that $\mathbf{a} \prec_{\mathbf{\Pi}} \mathbf{b}$

Definition : A function $f : \mathbb{R}^p_{++} \rightarrow \mathbb{R}$ is multiplicatively Schur convex if for every **x**,**y** $\in \mathbb{R}^p_{++}$

$$\mathbf{x} \prec_{\mathbf{\Pi}} \mathbf{y} \implies f(\mathbf{x}) \leqslant f(\mathbf{y})$$

and multiplicatively Schur concave if,

$$\mathbf{x} \prec_{\mathbf{\Pi}} \mathbf{y} \implies f(\mathbf{x}) \geqslant f(\mathbf{y})$$

Keep in mind that we want to find a lower bound for both these functions in our optimization problems. Hence it is necessary to find a vector which is multiplicatively majorized by our input vector (to the function) in case of multiplicatively Schur convex f and the other way round in case of multiplicatively Schur concave f

Let $\mathbf{A} = (\mathbf{X^H R X} + \rho\mathbf{I_K})^{-1} = \mathbf{L L^H}$, $\mathbf{H_n^H H_n} = \mathbf{R}$, $P_n = T$, $\mathbf{U_n} = \mathbf{X}$, $\mathbf{L_n} = \mathbf{L}$ where $\mathbf{R}$ is positive semi-definite and $\mathbf{A}$ is positive definite. The optimization problem at hand is,

$$\min_{\mathbf{X}} \phi(\mathbf{d^2_L}(\mathbf{X}))$$
$$\text{st} \ \ \text{Tr}(\mathbf{X}\mathbf{X}^H) = T$$

where $\mathbf{d^2_L}(\mathbf{X}) = ([\mathbf{L}]^2_{\mathbf{1},\mathbf{1}}, ..., [\mathbf{L}]^2_{\mathbf{K},\mathbf{K}})$

(i) Multiplicative Schur Concave

Suppose ϕ is multiplicative Schur concave. Then the optimal solution is given by,

$$\mathbf{X_{opt}} = \mathbf{E\Gamma}$$
$$\mathbf{U_n^{opt}} = \mathbf{E_n\Gamma_n}$$
$$\mathbf{C_n} = \mathbf{I_K}\ ;\ \mathbf{B_n} = \mathbf{0} \quad \because \mathbf{L_n} \sim \mathbf{diag}(..)$$

where columns of $\mathbf{E}$ are the eigenvectors of $\mathbf{R}$ associated with K largest eigenvalues in increasing order and $\mathbf{\Gamma} = diag(\gamma_1, ..., \gamma_K)$; $\gamma_i \geqslant 0$ (whose values are function dependent). This tells us that the precoding matrix has its columns along the beam-forming directions of the channel. We see that for optimal values this reduces to linear precoding (since $\mathbf{B_n} = \mathbf{0}$) with channel diagonalization [4].

Proof Outline

It involves 2 steps. In the first step we obtain a condition on $\mathbf{X}$ using a vector which provides a lower bound for our desired function. In the second step we obtain the closed form solution for $\mathbf{X}$ using this condition. We define $\lambda_\mathbf{A}(\mathbf{X}) = (\lambda_1, ..., \lambda_q)$

which are the set of eigen values of $\mathbf{A}$. Using Weyl's theorem it can be shown that

$$\mathbf{d}_{\mathbf{L}}^{\mathbf{2}}(\mathbf{X}) \prec_{\Pi} \lambda_{\mathbf{A}}(\mathbf{X})$$
$$\implies \phi(\lambda_{\mathbf{A}}(\mathbf{X})) \leqslant \phi(\mathbf{d}_{\mathbf{L}}^{\mathbf{2}}(\mathbf{X}))$$

The equality in the above equations can be achieved only if $\mathbf{A}$ is diagonal which also means $\mathbf{X^H R X}$ is also diagonal. Observe that this also means that $\mathbf{L}$ is diagonal. From [4] Lemma 12 it can be shown that if $\mathbf{R}$ is positive semi-definite Hermitian, $Tr(\mathbf{X X^H}) = T$ and $\mathbf{X^H R X}$ diagonal with increasing entries, then it is possible to find

$$\tilde{\mathbf{X}} \ s.t. \ \tilde{\mathbf{X}}^{\mathbf{H}}\mathbf{R}\tilde{\mathbf{X}} = \mathbf{X^H R X}$$
$$\tilde{\mathbf{X}} = \mathbf{E\Sigma} \ ; \ Tr(\tilde{\mathbf{X}}\tilde{\mathbf{X}}^{\mathbf{H}}) \leqslant Tr(\mathbf{X X^H})$$
$$\mathbf{X_{opt}} = \hat{\mathbf{X}} = \alpha\tilde{\mathbf{X}} = \mathbf{E}(\alpha\mathbf{\Sigma}) = \mathbf{E\Gamma}$$
$$\alpha = \sqrt{\frac{T}{Tr(\tilde{\mathbf{X}}\tilde{\mathbf{X}}^{\mathbf{H}})}} \geqslant 1$$

This $\mathbf{X_{opt}}$ is our optimal $\mathbf{U_n}$

(ii) Multiplicative Schur Convex

All the variables are similar to the previous case except ϕ which is multiplicative Schur convex here. The optimal solution is stated below,

$$\mathbf{X_{opt}} = \mathbf{E\Omega F}$$
$$\mathbf{U_n^{opt}} = \mathbf{E_n \Omega_n F_n}$$
$$\mathbf{C_n^{opt}} = \kappa_n \mathbf{L_n^{-1}}$$
$$\omega_{n,k}^2 = \left(\mu_n - \frac{\rho}{\eta_{n,k}}\right)_+ \quad ; \quad \sum_{i=1}^{q} \omega_{n,k}^2 = T$$

where $\mathbf{\Omega} = \mathbf{diag}(\omega_1, ...\omega_K)$, $(\eta_1, ...\eta_K)$ are K largest eigen-values of $\mathbf{R}$ and $\mathbf{F}$ is a unitary matrix chosen such that all the principal diagonal elements of $\mathbf{L}$ are equal. Since these diagonal elements represent the MSEs, all of them are also equal and their value is given by

$$MSE_{n,k} = \sigma_w^2 \left(\prod_{k=1}^{K} \frac{1}{\eta_{n,k}\omega_{n,k}^2 + \rho}\right)^{\frac{1}{K}}$$

Proof Outline

This also involves two steps which are the same as the previous case. We define a vector $\mathbf{x_{GM}}$ for convenience as follows,

$$\mathbf{x_{GM}} \prec_{\Pi} \mathbf{x} \ ; \ [\mathbf{x_{GM}}]_i = \left(\prod_{j=1}^{q} x_j\right)^{\frac{1}{q}} \tag{3}$$

We choose $\mathbf{F}$ such that all the diagonal elements of $\mathbf{L}$ are equal. Using certain linear algebra manipulations we can obtain the following,

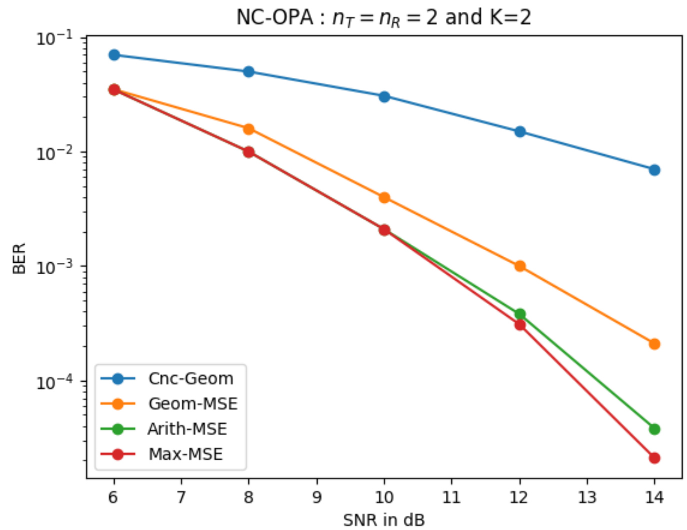
$$\mathbf{F^H A F} = \mathbf{L L^H}$$
$$\mathbf{F^H A F} = (\mathbf{F^H X^H R X F} + \rho\mathbf{I_K})^{-1}$$
$$\phi(\Delta^{1/q}, ..., \Delta^{1/q}) = \phi(\mathbf{d}_{\mathbf{L}}^{\mathbf{2}}(\mathbf{XF})) \leqslant \phi(\mathbf{d}_{\mathbf{L}}^{\mathbf{2}}(\mathbf{X}))$$

The last equation holds from Equation 3 and from the fact that all the elements of $\mathbf{L}$ are equal to $\Delta^{1/q}$, where Δ is the determinant of $\mathbf{A}$. We need to find a $\tilde{\mathbf{X}}$ which minimizes Δ

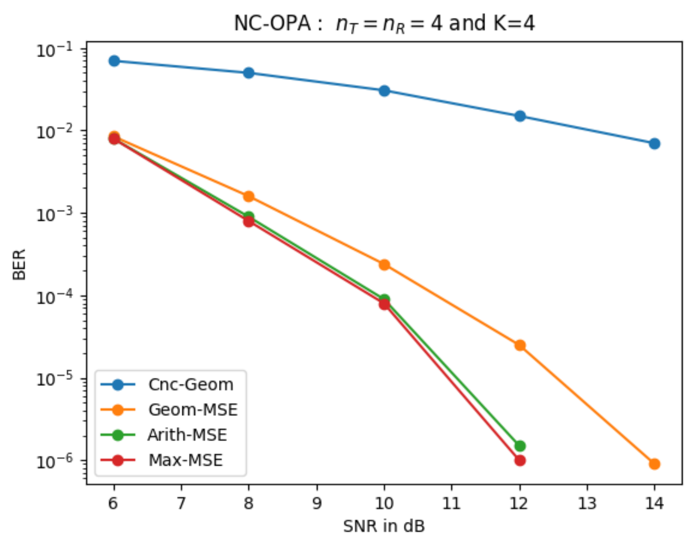
$$\min_{\mathbf{X}} \ \Delta = det((\mathbf{X^H R X} + \rho\mathbf{I_K})^{-1})$$
$$s.t. \ Tr(\mathbf{X X^H}) = T \tag{4}$$

$$\mathbf{X_{opt}} = \tilde{\mathbf{X}}\mathbf{F} = \mathbf{E\Omega F}$$

We observe that the optimization in Equation 4 is similar to the water-filling problem and hence one can observe similarity in their solutions.



(a)



(b)

Fig. 5. Block diagram of complete signal model in [3]

V. Multiplicative Schur convexity and examples

We state some example objective functions and explain their nature. Consider the following multiplicative Schur concave function(CNC-GEOM)

$$f_n(MSE_{n,k}) = \prod_{k=1}^{K} MSE_{n,k}^{\beta_k}$$
$$\text{s.t.} \quad 0 \leqslant \ \beta_1 \ \leqslant \ ..... \ \leqslant \ \beta_K$$

Let us look at the example stated above where $\mathbf{a} = (2,3,2)^T$ and $\mathbf{b} = (1,6,2)^T$ and $\mathbf{a} \prec_{\mathbf{\Pi}} \mathbf{b}$. Let these **a,b** represent a set of MSE_n for different carriers. The terms vectors arranged in descending order would look like $\tilde{\mathbf{a}} = (3,2,2)^T$ and $\tilde{\mathbf{b}} = (6,2,1)^T$. Take $\beta_1 = 1, \beta_2 = 2, \beta_3 = 3$

$$
\begin{gathered}
f(\mathbf{b}) = f(\tilde{\mathbf{b}}) = 6.2^2.1 = 24 \\
f(\mathbf{a}) = f(\tilde{\mathbf{a}}) = 3.2^2.2^3 = 96 \\
\therefore f(\mathbf{a}) \geqslant f(\mathbf{b})
\end{gathered}
$$

Hence we see that the function is multiplicative Schur concave. Now let us look at the following multiplicative Schur convex function (ARITH-MSE),

$$f(MSE_{n,k}) = \sum_{k=1}^{K} MSE_{n,k}$$

Consider the same inputs **a,b** for this function

$$
\begin{gathered}
f(\mathbf{b}) = f(\tilde{\mathbf{b}}) = 6 + 2 + 1 = 9 \\
f(\mathbf{a}) = f(\tilde{\mathbf{a}}) = 3 + 2 + 2 = 7 \\
\therefore f(\mathbf{a}) \leqslant f(\mathbf{b})
\end{gathered}
$$

We can write the optimum value of f_n for this function. It is the sum of all the $MSE_{n,k}$, all of which are equal at the optimum.

$$\hat{\alpha}_n(P_n) = \sigma_w^2 K \left(\prod_{k=1}^{K} \frac{1}{\eta_{n,k}\omega_{n,k}^2 + \rho} \right)^{\frac{1}{K}}$$

Shown in Figure 5 are the simulation results of BER vs SNR for different objective functions. We see that the Concave Geometric function has the worst performance amongst all of them. This can attributed to the fact that the optimal precoding reduces to linear in this case and there is no removal of interference using THP as in case of the other Schur convex functions. GEOM-MSE is the objective function which takes product of all the $MSE_{n,k}$, and it is multiplicative Scur convex. The maximum MSE minimizes the maximum $MSE_{n,k}$ over all K and hence has the best performance. Since ARITH-MSE is L_1 norm and max-MSE is l_∞ norm we can expect their performances to be similar. It can be shown that the carrier cooperative scheme is better than max-MSE when THP is used.

VI. Carrier Cooperative Scheme

The multi-carrier cooperative scheme is similar to a single carrier scheme from mathematical perspective. Here there is cooperation allowed among the carriers and we do joint processing and optimization. The independent processing we did in NC-OPA was an approach similar to greedy algorithm where we minimize the $MSE_{n,k}$ of each carrier. The carrier cooperative scheme does not have imposition of separate carrier processing and hence will have performance which is at least as good as NC-OPA if not better. This scheme will have joint process matrices for all carriers. The signal model for the above scheme will be,

$$\mathbf{y} = \mathbf{GHUb} + \mathbf{Gw} \ ; \ \mathbf{H} = \mathbf{diag}(\mathbf{H_0}, ...\mathbf{H_{N-1}})$$

VII. Conclusion

We used THP precoding to remove the interference caused by the channel. NC-OPA scheme is used to process and optimize each carrier independently. The optimization of $\mathbf{G_n}$ was a LMMSE estimation problem and that of $\mathbf{C_n}$ was a norm minimization problem. For multiplicative Schur concave objective functions the solution boiled down to linear precoding and channel diagonalization. In case of multiplicative Schur convex objective function the effective optimization was similar to the water-filling solution. In both of the above cases the precoding matrix $\mathbf{U_n}$ is aligned along the beam-forming directions of the channel. The simulations show that the non-linear precoding provides good performance for multiplicative Schur convex objective functions. We assessed that the carrier cooperative scheme has better performance when compared to NC-OPA.

VIII. Critique

The paper is a good development from [4] where he analyses non-linear precoding instead of linear. The clarity of information, from the structure and signal model to the rigor in the proof of the optimizations to the quantities in the simulation, is very good all around the paper. He has used figures in appropriate places to strengthen the idea he wants explain. But for any reader who looks for a more intuitive idea, amidst the intricate linear algebra and optimization, the paper appears slightly dry. For example, the motivation for using certain vectors to compare with the vector under consideration (in the proof for multiplicative Schur convex functions) or the result that $\mathbf{U_n}$ is the set of right singular vectors of the channel matrix are not explained clearly. Even while explaining the simulation results, there has been no mention as to why one particular function is better than the others or as to why the multiplicative Schur concave function performs so poorly compared to the others. A mention as to why these multiplicative Schur functions are special and why such properties hold for them would have been insightful. Apart from these minor issues the paper is a good read and we should commend the author for putting in a lot of thought in explaining his research.

References

1 Tomlinson, M., "New automatic equaliser employing modulo arithmetic," *Electronics letters*, vol. 7, no. 5, pp. 138–139, 1971.

2 Harashima, H. and Miyakawa, H., "Matched-transmission technique for channels with intersymbol interference," *IEEE Transactions on Communications*, vol. 20, no. 4, pp. 774–780, 1972.

3 D'Amico, A. A., "Tomlinson-harashima precoding in mimo systems: A unified approach to transceiver optimization based on multiplicative schur-convexity," *IEEE Transactions on Signal Processing*, vol. 56, no. 8, pp. 3662–3677, 2008.

4 Palomar, D. P., Cioffi, J. M., and Lagunas, M. A., "Joint tx-rx beamforming design for multicarrier mimo channels: A unified framework for convex optimization," *IEEE Transactions on Signal Processing*, vol. 51, no. 9, pp. 2381–2401, 2003.