A Close to Capacity Double Iterative Based Precoder Design for MU-MIMO Broadcast Channel with Multi-Streams Support

Mustapha Amara, Yi Yuan-Wu and Dirk Slock

Abstract—Many algorithms have been proposed for precoders design in a multiuser MIMO system (MU-MIMO). Nevertheless, the proposed solutions showed to have better results for some SNRs (signal to noise ratio) regions and degrade in some other parts. This paper proposes a new double iterative procedure for sum-rate maximization. The proposed algorithm is based on jointly optimizing the precoders and decoders using two different decoding schemes. The solution here is supporting multi-streams per user. The algorithm is based on a WMMSE (weighted minimum mean square error) precoder combined with two iterative receivers namely the MF (matched filter) and MMSE (minimum mean square error) decoders. The resulting precoding matrices from the first algorithm (WMMSE/MF) are used as an initialization for the second one (WMMSE/MMSE). The choice of these decoders and their combination has been done according to their properties. Another crucial point in this proposal is the decision on the switching point between these two algorithms. A dynamic algorithm introducing very low extra complexity is proposed here.

To validate our proposed solution we compare it with an existing MMSE and WMMSE based iterative optimization algorithms. The obtained results demonstrate significant gains without introducing supplementary complexity. Comparison with DPC (dirty paper coding) performances shows how close our proposed solution is to the BC (broadcast channel) channel capacity.

Index Terms—Multi-user, MIMO, broadcast channel, capacity, iterative, double iterative, WMMSE, MMSE, matched filter.

I. INTRODUCTION

Multiuser MIMO (MU-MIMO) downlink system known in the information theory as the broadcast channel system represents today one of the most important research fields in wireless communications because of the high potential it offers in improving both reliability and capacity of the system. Some theoretical analysis of the capacity demonstrated that the capacity of a broadcast MU-MIMO channel can be achieved by applying a dirty-paper coding (DPC) [1]–[3] algorithm as a precoder. Nevertheless, a DPC precoding is difficult to implement and is high resource consuming. Some suboptimal linear algorithms with lower implementation costs exist and can be divided into two main families: the iterative [4]–[8] and the closed form solutions [9]–[12].

Another level of classification of precoders is the number of streams that might be offered to each user. In fact, there are precoders that can only support at maximum one stream per user even if the system is not fully charged. Such precoders have been proposed and widely studied in [6], [7], [9]–[13]. Some multi-stream precoding solutions have nevertheless been proposed such as in [14], [15] for a closed form solution and in [4], [5], [8], [16] for an iterative solution. Closed form solutions are known to be limited ones as the optimization process is in the best cases done in a recursive way and limits thus the optimization process. On the other hand, iterative solutions are able theoretically to get very close to the optimal solution; but they present two major problems namely initialization and convergence. In fact, the initialization of the iterative algorithm is crucial for the quality of the obtained precoders and even for convergence sometimes.

In this paper we are going to focus on the iterative solution for precoder and decoder design for a MU-MIMO system considering multiple streams per user.

The multiple streams can be allocated to the same user respecting two main constraints $Q_k \leq \min(N_{R_k}, N_T)$ representing the maximum number of streams per user and $Q = \sum_{k=1}^{K} Q_k \leq \min(\sum_{k=1}^{K} N_{R_k}, N_T)$ representing the total number of streams allocated by the base station (BS). The allocation of these streams is done such as it maximizes the total sum-rate (SR). A further crucial point in SR maximization is defining the best power distribution over the selected streams. These two problems have been partially solved by applying SR-optimizing-weights to the streams [16].

We propose here a novel precoder design strategy and we compare the performances based on the total achieved sum-rate.

In next section, the model for the considered system is presented. In Section III a detailed description of the best existing WMMSE multi-stream precoding is given; the proposed algorithm is explained and its structure is detailed. Section IV contains some simulation results. They demonstrate the performances obtained by our proposed algorithm and compares it to the one presented in [16].

II. SYSTEM MODEL

Let us consider in our study a multi-user MIMO communication system with $N_T$ transmission antennas at the base station and $K$ users with $N_{R_k}$ receiving antennas for user $k$.

We assume that the base station has a perfect knowledge of the channel state information (CSI) of all $K$ users. Let $s_k$ a $Q_k \times 1$
vector representing the transmitted data symbols for user \( k \)
where \( Q_k \) is the number of transmitted streams for the same user.
In our paper we are interested in the case of multiple streams
per user \( Q_k \leq \min(N_{R_k}, N_T) \). The total number of streams
must not exceed the maximum number that can be supported
by the system and defined as \( Q \leq \min(\sum_{k=1}^{K}N_{R_k}, N_T) \).

The total transmit power at the base station is supposed to
be constant and equal to \( P_T \). The noise variance is noted \( \sigma^2 \).

For the channel part, \( H_k \) denotes the MIMO channel for user
\( k \) which is a \( N_{R_k} \times N_T \) matrix. Each element composing
the channel matrix is considered to be a complex Gaussian random
variable with unit variance and zero mean.

In this paper \( X^H \) stands for the transpose conjugate of \( X \),
\( tr(X) \) for the trace of \( X \) and \( diag \{ X_1, \ldots, X_n \} \) for a zero-filled
matrix with matrix elements \( \{ X_i \}_{1 \leq i \leq n} \) on the diagonal.

III. PROPOSED ALGORITHM

The objective is to design the precoding matrices \( T_k \) under
the total transmit power constraint \( \sum_{k=1}^{K} P_k = P_T \). Here \( P_k =
tr\left(T_kT_k^H\right) \) denotes the transmitted power aimed to user \( k \).

Therefore, we consider a MMSE precoder and two decoders
an MF (matched filter) and a MMSE receiver.

A. WMMSE Precoder

Let’s consider the MMSE precoder minimizing the mean square error given in (1)

\[
T_k = \alpha \left( \sum_{j=1}^{K} H_j^D H_j^T D_j H_j + \frac{tr(D_k D_k^H)I_{N_T}}{P_T} \right)^{-1} H_k^H D_k^H \tag{1}
\]

where \( \alpha \) is a scalar factor. And we apply a stream distribution
among the available ones. The distribution of streams is done
by assigning a weight matrix \( W_k \) to each user \( k \).

Applying the distribution to the precoders gives the expression of the new MMSE precoder named WMMSE (weighted MMSE)
according to (2)

\[
T_k = \beta \frac{H_k^H D_k^H W_k}{\sum_{j=1}^{K} H_j^H D_j^H W_j D_j H_j + \frac{tr(W_k W_k^H)I_{N_T}}{P_T}} \tag{2}
\]

Where \( W_k \) is the weight matrix given to streams of user \( k \)
and \( \beta \) is a scalar factor to respect the total power constraint
\( \sum_{i=1}^{K} tr\left(T_i T_i^H\right) = P_T \).

A condensed expression for all users is given in (3)

\[
T = \beta \left( H^H D^H W D H + \frac{tr(W D D^H)I_{N_T}}{P_T} \right)^{-1} H^H D^H W \tag{3}
\]

Here

\[
\begin{cases}
T = diag \{ T_1, \ldots, T_K \} \\
D = diag \{ D_1, \ldots, D_K \} \\
W = diag \{ W_1, \ldots, W_K \}
\end{cases}
\]

B. Receiver Design

Different structures have been proposed in the literature for
the receiver design for MIMO systems. Among the existing
proposed solutions there is the matched filter (MF). We propose
the multi-stream matched filter of equation (4)

\[
D_{MF,k} = \frac{T_k^H H_k^H}{\| H_k T_k \|} \tag{4}
\]

where \( \| X \| \) is the Frobenius norm of matrix \( X \).

Another receiving structure minimizing the mean square error
is the MMSE receiver given in (5)

\[
D_{MMSE,k} = T_k^H H_k^H \left( I_{N_R} + \sum_{i=1}^{K} H_k T_i T_i^H H_k^H \right)^{-1} \tag{5}
\]

C. Iterative Algorithm

Based on the previous precoder and decoders, an iterative
algorithm can be defined to optimize the precoder and decoder
design. The algorithm given in Algorithm 1 has been originally
proposed in [16] with an MMSE decoder. It can nevertheless be
applied to different receiving structures.

Algorithm 1 Iterative WMMSE

1) Initialize the precoders \( T_k, k \in \{ 1, \ldots, K \} \).
2) Compute the selected type of receiver corresponding to
the precoders \( T_k \) for all the users \( k \in \{ 1, \ldots, K \} \).
3) Compute the weights according to equation (6).
4) Compute the new precoders \( T_k, k \in \{ 1, \ldots, K \} \) according
to equation (3).
5) Repeat steps 2) to 4) until convergence.

The algorithm requires computing the weights for the different
streams required for stream selection and power distribution.
[16] proposes (6)

\[
W_k = I_k + T_k^H H_k^H \left( I_{N_R} + \sum_{i=1, i \neq k}^{K} H_k T_i T_i^H H_k^H \right)^{-1} H_k T_k \tag{6}
\]

corresponding to the inverse of the mean square error.

For performance analysis we estimate the total sum-rate of the
MU-MIMO system. The expression of the throughput is the sum
over all selected streams of the individual achieved throughputs
for each user and can be given by equation (7)

\[
SR = \sum_{k=1}^{K} log_2 det \left( I + \frac{H_k T_k T_k^H H_k^H}{\gamma_k + N_0 I} \right) \tag{7}
\]

here, \( \gamma_k = \sum_{j=1, j \neq k}^{K} H_k T_j T_j^H H_k^H \) represents the interference part
received by user \( k \).

D. Dynamic Flip Procedure

Analyzing the performances of the iterative algorithm using
different decoders to optimize the precoders shows various
throughput levels. Some algorithms present higher sum-rates at
high SNRs like the WMMSE/MF proposed in the last subsection
and some other have better performances at lower SNRs like the
WMMSE/MMSE proposed in [16].

One naive and direct solution to get high performances in the entire SNR range would be to run in parallel these two algorithms and then choose the best among them. This would solve the problem but requires twice the resources used by the iterative algorithms.

Moreover, iterative algorithms are sensitive to initialization and known to suffer from convergence problems as mentioned in [10]. This major problem remains unsolved by adopting this strategy.

The main idea of the double iterative procedure proposed in this paper is to combine two versions of the iterative algorithm derived from Algorithm 1 using different receiving structures to be able to cover the largest part of the space containing the possible transmitters. Therefore, a first algorithm WMMSE/MF is performed sweeping the space of possible precoders trying to maximize the received powers. The second algorithm WMMSE/MMSE presenting an increasing sum-rate behavior refines the solution towards the maximum. This will minimize the probability of entering a local maximum, one of the main limiting factor for iterative algorithms.

But combining two versions of the algorithm, implies a flipping point where the used algorithm (receiver) is changed. Furthermore, some statistical analysis of the throughputs given by the cascade of the two versions WMMSE/MF and WMMSE/MMSE described in Algorithm 1 demonstrated that the optimal flipping point is not only a function of the SNR (Signal to Noise Ratio), the system configuration (Number of transmitting and receiving antennas) but also of the channel realizations namely the matrices \( H \) and \( \epsilon \)

A solution would be to perform some lookup tables in function of the system configuration. But the dimensions of these tables are exponential and can rapidly explode.

To get rid of these constraints and still be able to get a significant gain, we propose a dynamic selection procedure based on an instantaneous convergence analysis.

The selection procedure is then based on the monitoring of the obtained throughputs over a fixed number of iterations that we are going to call sliding window. The number of iterations considered in this window is noted \( WIN_{MF} \) as the first considered receiver is the MF one.

To be able to run the selection procedure, a minimum of \( WIN_{MF} \) observations of the iterative algorithm must be available. Therefore, in the first phase, the Algorithm 1 is run for \( WIN_{MF} \) iterations. Starting from this point, the monitoring procedure is launched: at each iteration \( \text{iter} \geq WIN_{MF} \) the variance of the obtained sum-rate over the last \( \text{WIN}_{MF} \) considered iterations is computed according to equation (8). This quantity is noted \( V_{SR} \).

\[
V_{SR} = \text{Var}\left(\left[SR^{\text{iter}-WIN_{MF}+1} \ldots SR^{\text{iter}}\right]\right) \\
= \frac{1}{WIN_{MF}} \sum_{i=0}^{WIN_{MF}-1} \left(\frac{1}{WIN_{MF}} \sum_{i=0}^{WIN_{MF}-1} SR^{\text{iter}-i}\right)^{2}
\]

This variance is compared to a prefixed threshold \( \varepsilon_{MF} \) defining the convergence of the algorithm.

So the average evolution of the SR over the last \( WIN_{MF} \) iterations is observed. If the \( V_{SR} \) is stabilizing or if increase is below a prefixed threshold, the MF receiver giving the best \( SR \) in this window is retained.

A last control parameter is introduced to avoid the divergence problem previously mentioned. It consists in limiting the number of possible iterations for the WMMSE/MF algorithm to \( \text{iter} = N_{\text{iter}}^{\text{max}} - \Delta \) where \( \Delta \in \mathbb{N}^{*} \) and \( 1 \leq \Delta \leq N_{\text{iter}}^{\text{max}} - WIN_{MF} \).

Here \( N_{\text{iter}}^{\text{max}} \) is the total number of iterations allowed for the processing of a given transmission and \( \mathbb{N}^{*} = \{1, 2, \ldots\} \). The goal of this limitation is the number of the total iterations is to avoid that the algorithm gets blocked in case of divergence or of non convergence.

### E. Double Iterative Procedure

In this last subsection, the entire double iterative procedure is presented. In a first phase, the WMMSE/MF algorithm given in Algorithm 1 with an MF decoder is executed followed by the WMMSE/MMSE given in Algorithm 1 with an MMSE decoder. The decision is taken based on the DFP (dynamic flip procedure) presented in the past subsection. The evolution of the receiver and the precoder through the iterations is performed thanks to the weights distribution given by (6). The overall iterative algorithm is then given in Algorithm 2.

### IV. Simulations And Results

In all our simulations, we consider that the number of receiving antennas is the same for all users \( N_R = N_R \). We suppose a Rayleigh fading channel \( H_k = (h_{i,j}^k)_{1 \leq i \leq N_R, 1 \leq j \leq N_T} \) such as \( E[\|h_{i,j}^k\|^2] = 1 \). The simulation generates 10000 independent channel realizations for each user. To generate the total throughput of the system, we perform an average over all channel realizations on the quantity \( SR \) given in equation (7). The two convergence control parameters for both algorithms \( \varepsilon_{MF} \) and \( \varepsilon \) are fixed and equal to to \( 10^{-3} \). In all the following, the maximal number of iterations \( N_{\text{iter}}^{\text{max}} \) is fixed to 50. The number of iterations is the sum of iterations performed by each of the two iterative optimization procedures. We consider \( WIN_{MF} = 5 \) and \( \Delta = 5 \).

Figure 1 represents simulation results for a MU-MIMO system with \( N_T = 4 \) transmitting antennas, \( N_R = 4 \) receiving antennas per user and \( K = 4 \) users. The analysis done below remains true for any system configurations (especially the LTE defined ones). The curves WMMSE/MMSE/MMSED obtained with Algorithm 2 is compared to the WMMSE/MMSE proposed in [16] and WMMSE/MMF of Algorithm 1 with an MF decoder. We add also two curves representing a single stream MMSE/MMSE algorithm. The first MMSE/MMSE algorithm is the original version proposed in [4] and the corresponding curve is entitled MMSE/MMSEOriginal. The second MMSE/MMSE curve named MMSE/NormalizedMMSE is a modified version of the algorithm proposed in [4] where the considered receiver is a normalized MMSE. The SNR/MSR represents the performances obtained with algorithm in [6]. For the DPC curve, we consider the algorithms given in [3].

Comparing WMMSE/MMSE and the WMMSE/MMF curves shows that the WMMSE/MMF gives better performances especially at high SNRs. This behavior can be explained by the fact that at high SNRs, the streams can be well separated...
Algorithm 2 Double Iterative WMMSE

1) Initialize $N_{\text{max}}^{\text{iter}}, W_{\text{MF}}, \epsilon, \epsilon_{\text{MF}}, \Delta$ and $\text{iter} = 0$
2) Initialize $T_k^{\text{iter}} = \beta H_k^H, k \in \{1, \ldots, K\}$ where $\beta$ is a scalar factor to respect the power constraint $\sum_{k=1}^{K} \text{tr}(T_k T_k^H) = P_T$
3) $\text{iter} = \text{iter} + 1$. Compute $D_k^{\text{iter}}, k \in \{1, \ldots, K\}$ using $T_{\text{iter}-1}$ with (4), $W_k^{\text{iter}}$ using $T_{\text{iter}-1}$ as in (6) and $T_k^{\text{iter}}$ using $D_k^{\text{iter}}$ and $W_k^{\text{iter}}$ with (3).
4) Compute the $SR_{\text{iter}}$ by using $T_k^{\text{iter}}$ with (7).
   - if $\text{iter} < W_{\text{MF}}$ then
     - jump to step 3
   - end if
5) Verify convergence of Algorithm 1 with the MF decoder:
   - if $\text{Var}(\{SR_{\text{iter}}-W_{\text{MF}}+\ldots-SR_{\text{iter}}\}) \leq \epsilon_{\text{MF}}$ then
     - jump to 6
   - else if $\text{iter} \leq N_{\text{iter}}^{\text{max}} - \Delta$ then
     - jump to step 3
   - else
     - jump to 6.
   - end if
6) Consider the precoder giving the best $SR$ over the last $W_{\text{MF}}$ iterations.
   - $\text{iter}_{SR_{\text{max}}} = \sum_{i \in \{\text{iter} \times W_{\text{MF}}+\ldots, \text{iter}\}} \max_i SR_i$
   - $T_k^{\text{iter}} = T_k^{\text{iter}_{SR_{\text{max}}}}$
7) $\text{iter} = \text{iter} + 1$. Compute $D_k^{\text{iter}}$ using $T_{\text{iter}-1}$ with (5), $W_k^{\text{iter}}$ using $T_{\text{iter}-1}$ using (6) and $T_k^{\text{iter}}$ applying (3).
8) Compute $SR_{\text{iter}}$ with (7) and verify convergence of Algorithm 1 with the MMSE decoder:
   - if $|SR_{\text{iter}} - SR_{\text{iter}-1}| < \epsilon$ then
     - jump to 9
   - else if $\text{iter} \leq N_{\text{iter}}^{\text{max}}$ then
     - jump to step 7
   - else
     - jump to 9.
   - end if
9) Stop the algorithm and consider the last computed precoders $T_k^{\text{iter}}$ and decoders $D_k^{\text{iter}}, k \in \{1, \ldots, K\}$.

just by using a matched filter at the reception and through the iterative procedure, the optimal precoder is calculated to maximize the received power for each user. At low SNRs, on the other hand, the MF filter fails to recover the streams in an optimal way and thus induces suboptimal precoder derivations. But, the MMSE receiver is capable of providing a better separation of the users and reorients iteratively aided by W the search towards the least interfering users.

The proposed algorithm gives a curve presenting better throughput in all the considered SNR range. The obtained throughputs are even higher than the maximum obtainable by selecting the best among the two considered algorithms. In fact, analyzing figures 1.b shows that at high SNRs, the proposed algorithm gives better throughput performances. At low SNRs,
as shown on the curves in 1.c the proposed procedure is capable not only of recovering the best of the two used algorithms but to generate even a slightly better throughput. These results show the stability of our algorithm and its convergence. These performances are obtained just by introducing a dynamic flipping procedure that does not introduce any supplementary computational complexity, any extra delay or any increase in the number of iterations thus no extra processing latency.

Comparing the performances obtained with the existing single stream solutions (MMSE/MMSE and SJNR/MSR) demonstrates much higher performances and shows the importance of optimizing the weights affected to the different possible streams. Moreover, comparing the curve of the proposed double iterative solution with dynamic selection procedure with the DPC curve shows that the solution is getting very close to the optimal precoding. The two curves are parallel for all SNR range presenting an offset of less than 1 bit/s/Hz. Despite the fact that a DPC is in general a non linear precoding technique able of perfectly canceling out all the interference parts, our proposed linear precoder offers almost the same performances in all the SNR range. This shows the stability and the good convergence of the algorithm containing a quick search using the WMMSE/MF algorithm followed by a refinement procedure with the WMMSE/MMSE.

Figure 2 gives the performances obtained in a system with $N_T = N_R = 4$ and $K = 2$ users. In this case the single stream solutions are under exploiting the system capabilities as only 2 streams are scheduled whereas the multi stream tries to exploit the full diversity by distributing the four available streams on the $K = 2$ users. This explains the difference in slope of the two solutions. This figure also confirms the previous comments.

V. CONCLUSION

In this paper, a novel iterative joint optimization procedure for sum-rate maximization is proposed. We introduce a new iterative procedure which combines two iterative sum-rate maximization algorithms based on joint precoder and receiver optimization namely the WMMSE/MMSE described in Algorithm 1 and the WMMSE/MMSE proposed in [16]. A dynamic switching solution has been proposed to cascade these two algorithms allowing us to extract the best of them without introducing further complexity. This solves the burden of finding the optimal flipping point. We also showed throughout the realized simulations that the presented algorithm is not only achieving the best throughput given by the two used algorithms but even gives further gains getting closer to the system capacity represented by the DPC. Comparisons done with some existing MMSE/MMSE iterative solution given in [4], [8], showed much better performances with the same complexity levels.

REFERENCES