Resource Allocation for QoS Multiuser MIMO with Zero Forcing and MMSE Beamforming

Zvi Rosberg 1, Antonio Cantoni 2 and Ren P. Liu 3

Abstract—To enable ubiquitous end-to-end quality of service over IP networks expanding to rural areas, we examine resource allocation problems of a MAC layer for an OFDM MIMO wireless base station (BS) using multiuser beamforming. We compare two linear beamforming techniques, zero-forcing (ZF-BF) and minimum mean square error (MMSE-BF). To guarantee minimum bandwidth and low packet delay in an environment where the number of users is much larger than the number of BS antennas, one needs to partition the users into several sets and combine space division multiple access (SDMA) with time division multiple access (TDMA). We study the impact on the optimal transmission power resulting from selecting the following design parameters: (i) ZF-BF vs. MMSE-BF; (ii) the number of user sets multiplexed by TDMA; and (iii) the number of BS antennas. Two notable results are observed: (1) power wise, ZF-BF is far more superior to MMSE-ZF; (2) substantial transmission power can be saved by increasing the number of BS antennas; however, only by large increments, e.g., from 6 to 9 and further to 21.

Index Terms—QoS, SDMA, TDMA, MIMO, Beamforming, ZF, MMSE.

I. INTRODUCTION

Forthcoming services over the Internet require adaptive and ubiquitous end-to-end guarantee of quality of service (QoS). An architecture, referred to as QoS-RMP (QoS with a Rate Management Protocol), its implementation and its demonstration enabling these QoS requirements is presented in [7] [8]. The reader is also referred to a video of our prototype demonstration at:


QoS-RMP is built on top of the native DiffServ and comprises edge and core modules enhancing the edge and core routers, respectively. It guarantees that each flow receives at least its minimum required bandwidth and that the end-to-end delay of every packet is below its required maximum. It also guarantees that the packet loss rate of every flow is below its required maximum loss rate.

QoS-RMP requires that all core routers/switches support the single-hop services of minimum bandwidth guarantee at their attached links, packet time scheduler and rate policing at the edge switches/routers. It worth noting that with current deployment of IP routers and the IP protocol suite, QoS-RMP can guarantee end-to-end adaptive QoS. However, to guarantee end-to-end QoS ubiquitously, the single-hop services must be supported by the wireless base stations (BS) as well.

Following IEEE 802.22 initiative of Wireless Regional Area Network (WRAN) to bring broadband access to rural areas, more research attention is given to the design and the algorithms in the physical (PHY) and media access control (MAC) layers. Akin to a classical cellular base station, each WRAN system comprises a single base station providing wireless up and down links access to stationary premises within a vicinity of about 100 km. The back-haul and the gateway connection to the Internet are beyond the scope of this paper.

For the PHY layer, orthogonal frequency multiplexing (OFDM) with multi-input multi-output (MIMO) antenna arrays is a mature and most spectral efficient technology. Particularly, linear beamforming (BF) using either zero forcing (ZF-BF) or minimum mean square error (MMSE-BF) have efficient implementations adaptable to time varying QoS requirements.

We address the MAC layer resource allocation problems of how to meet the QoS requirements of bandwidth and packet delay per end user when the BS broadcast/receive data to/from downlink/uplink channels using OFDM MIMO with multiuser beamforming [2] [3] [9] [13]. These problems include: (i) space division multiple access (SDMA); (ii) time division multiplexing (TDMA); and (iii) transmission power optimization.

For a typical WRAN, we also study the impact on the transmission power resulting from: selecting ZF or MMSE; the number of user sets; and the number of BS antennas. Two notable observations of this paper are the following. One is that transmission power wise, ZF is far more superior to MMSE. The other is that substantial transmission power can be saved by increasing the number of BS antennas; however, only by large increments, e.g., from 6 to 9 and further to 21.

The system model is presented in Section II. The optimal minimum transmission powers subject to minimum bandwidth requirement per user under ZF-BF and MMSE-BF are derived in Section III. In Section IV we present the combined SDMA-TDMA scheduler, and in Section V we evaluate the impact of several MAC design parameters on the optimal transmission powers. Conclusions are drawn in Section VI.

II. SYSTEM MODEL

Consider a WRAN comprising a base station (BS) with M antenna elements and \( K > M \) user terminals. We assume the worst case, which is also the most common one, where each user terminal has a single element antenna. Demonstrators of “broadband-to-the-bush” system being developed by CSIRO,
Australia, will use a 12-element antenna MIMO BS and will typically cover about 100 terminal users [10].

Although the optimal sum-capacity MIMO broadcast channel is attained by dirty paper coding (DPC), [4] [14], its computational complexity give rise to interest in suboptimal linear BF techniques such as zero forcing (ZF-BF) and linear minimum mean square error (MMSE-BF) [2] [3] [9] [13].

The baseband channel of subcarrier \( n \) (out of \( N \) OFDM subcarriers) is modeled by the downlink gain matrix \( H \in \mathbb{C}^{K \times M} \). The signal received at user \( k \) on subcarrier \( n \) is

\[
y_{k,n} = h_{k,n}^* x_n + z_{k,n}, \quad k = 1, \ldots, K,
\]

where \( h_{k,n}^* \) denotes the conjugate transpose of an array \( A \), \( h_{k,n} = [h_{k,n,1}, \ldots, h_{k,n,M}]^T \) is the \( k \)-th user gain vector at subcarrier \( n \), \( x_n = [x_{n,1}, \ldots, x_{n,M}]^T \) is the transmitted symbol vector of the BS on subcarrier \( n \), and \( z_{k,n} \) is the additive complex Gaussian noise variance \( \sigma_{k,n}^2 \) on subcarrier \( n \) at the \( k \)-th user.

Without loss of generality, the noise power is normalized to \( \sigma_{k,n}^2 = 1 \), and \( h_{k,n} \) is given in noise magnitude units. Note that \( x_n \) is the output for all the users receiving the broadcast signal on subcarrier \( n \); hence it is not subscripted with \( k \).

A linear beamforming transmit spatial filter is characterized by a weight vector \( w_{k,n} = [w_{k,n,1}, \ldots, w_{k,n,M}]^T \) mapping the \( k \)-th user data symbol on subcarrier \( n \), \( b_{k,n} \), to the BS antenna array outputs. That is, given that beams are formed concurrently for any subset of \( K' \) users, the output signal is

\[
x_n = \sum_{k=1}^{K'} b_{k,n} w_{k,n}, \quad \forall \ n.
\]

If the receiver detects the transmitted symbols by assuming that the interference terms are distributed as additive Gaussian noise, the information theoretic rate of user \( k \) of this scheme, given that the transmission powers are \( \{P_{k,n}\} \), is [11]

\[
R_k = \sum_{n=1}^{N} \frac{BW}{N} \log_2 \left( 1 + \frac{P_{k,n}|h_{k,n}^* w_{k,n}|^2}{\sigma_{k,n}^2 + \sum_{j \neq k} P_{j,n}|h_{j,n}^* w_{j,n}|^2} \right),
\]

where \( BW \) is the channel bandwidth (Hz).

A. Zero-Forcing Beamforming

The objective of ZF-BF is to “zero” the multiple access interference (MAI). From vector orthogonality consideration, this can only be done with users having orthogonal received signals, i.e., with at most \( M \) users (as the BS antenna elements). Thus, the output signal for each user \( k \) must be determined by a weight vector, \( w_{k,n} \), which is orthogonal to all other users, i.e., satisfying

\[
h_{j,n}^* w_{k,n} = 0, \quad \forall \ j \neq k.
\]

Note that the weights may depend on the subcarrier. To comply with this restriction of ZF-BF broadcasting, we need to partition the \( K \) users into subsets, each of which is of at most \( M \) users. Then, we have to time-multiplex the BS broadcasting between those user sets. A separate user partitioning process may be required for each subcarrier.

For every \( n \), let \( S_n \) denote a generic set of \( K', K' \leq M \) users relabeled by \( k = 1, \ldots, K' \). Using a matrix representation, the received signals on subcarrier \( n \) are given by

\[
y_n = H_n^*(S_n)x_n + z_n,
\]

where \( y_n = [y_{1,n}, \ldots, y_{K',n}]^T \) represent the received symbols of the users; \( H_n^*(S_n) \in \mathbb{C}^{K' \times M} \) is the channel complex gain matrix of subcarrier \( n \) corresponding to user set \( S_n \) (with a row per user and a column per antenna element); \( x_n \) is the complex antenna array output on subcarrier \( n \); and \( z_n \) is the additive complex Gaussian noise on subcarrier \( n \) at the users. The noise variance is normalized to one. ZF-BF can be achieved by using the following weight matrix

\[
W_n(S_n) = H_n(S_n)(H_n^*(S_n)H_n(S_n))^{-1},
\]

where \( M^\dagger \) denotes the pseudo inverse of the matrix \( M \).

One form of ZF-BF channel gains is obtained by using the singular value decomposition (SVD) of \( H_n(S_n) \)

\[
H_n(S_n) = U \Sigma V^*,
\]

where \( U \) is a \( K' \) by \( K' \) unitary matrix, \( \Sigma \) is an \( K' \) by \( M \) diagonal matrix with nonnegative real numbers on the main diagonal, and \( V^* \) denotes the conjugate transpose of an \( M \) by \( M \) matrix \( V \).

The diagonal entries of \( \Sigma \), denoted by \( \sigma_i, 1 \leq i \leq \min\{K', M\} \), are called the singular values of \( S_n \). There are at most \( \min\{K', M\} \) positive singular values.

By (3), given the transmission powers \( \{P_{k,n}\}, (2) \) becomes

\[
R_k = \sum_{n=1}^{N} \frac{BW}{N} \log_2 \left( 1 + \frac{P_{k,n}|h_{k,n}^* w_{k,n}|^2}{\sigma_{k,n}^2} \right),
\]

and \( \sigma_{k,n} \) are the singular values of \( H_n^*(S_n)H_n(S_n) \).

Note that \( \{\gamma_{k,n}(S_n)\} \) depend on the entire selected set \( S_n \), emphasizing the importance of the user set partitioning.

It is also noteworthy that the SVD-based ZF-BF (Moore-Penrose ZF) is not the optimal max-min ZF. That is, a ZF-BF that maximizes the minimum rate in the network subject to per-antenna power constraints [5]. Indeed, with SVD-based ZF-BF some user \( k \) may have \( \gamma_{k,n}(S_n) = 0 \) (zero rate).

B. Minimum Mean Square Error Beamforming

When each user has a single antenna element, the weights of the MMSE-BF is given by (e.g., [3])

\[
w_{k,n} = \frac{h_{k,n}}{\sigma_{k,n}^2 + h_{k,n}^* h_{k,n}}.
\]

In a matrix form, the weight matrix is given by

\[
W_n = H_n \left[ \text{diag}(\sigma_{k,n}^2) + \text{diag}(H_n^*H_n) \right]^{-1},
\]

where \( H_n \) is the gain matrix of subcarrier \( n \).
By (8), given the transmission powers \( \{P_{k,n}\} \), (2) becomes

\[
R_k = \frac{N}{N} \log_2 \left( 1 + \frac{P_{k,n} \rho_{k,k,n}}{\sigma_{k,n}^2 + \sum_{j \neq k} P_{j,n} \rho_{j,k,n}} \right), \quad (9)
\]

where

\[
\rho_{k,j,n} = \left| [H_n^* W_n]_{k,j} \right|^2. \quad (10)
\]

In the next section we derive the optimal power allocation subject to minimum bandwidth constraint per user.

III. MINIMUM POWERS WITH BANDWIDTH GUARANTEE

Transmission power control is an efficient control device for addressing minimum user bandwidth. Note that in our WRAN system, all users are stationary and the channel gains can be estimated with very high accuracy. This gives rise to a power control problem with perfect channel state information (CSI).

The relevant optimization criterion for power control with bandwidth guarantee is minimizing the powers guaranteeing the minimum bandwidth requirements of all users. The optimization problem with ZF-BF is very distinctive from the problem with MMSE-ZF. In the ZF-BF case, MAI is not present and the optimization problem boils down into selecting the best subcarriers for each user. With MMSE-BF, the presence of MAI defines an optimization where signal and interference must be overcome. Indeed, by (9),

\[
\min \big\{P_{n} \geq 0 \mid I(P_{n}) \big\} \quad (18)
\]

where \( P_n = (P_{1,n}, \ldots, P_{K,n})^T \) denotes the BS transmitter power vector for each user and \( I(P_n) = (I_1(P_n), \ldots, I_K(P_n))^T \) denotes the effective interference vector of the combined noise and MAI each user power must overcome. Indeed, by (9),

\[
I_k(P) = \mathbf{\gamma} \left( \rho_{k,k,n}^2 + \sum_{j \neq k} P_{j,n} \rho_{j,k,n} \right). \quad (19)
\]

Problem (18) is a classical power allocation problem, extensively studied since 1983, for DS-CDMA and multi-cell TDMA systems trading off between signal and interference powers. It can be solved centrally [15] using the perfect CSI. There is also a distributed iterative solution based only on each user local information [12].

With perfect CSI, the minimum powers are given by solving the linear equations

\[
P_n = I(P_n).
\]

By [12], a unique \( P_n \) exists if there a \( P_n \geq I(P_n) \). This solution, which is also the minimum power vector (component-wise), is given by inverting the following corresponding coefficient matrix.

The solution of (13)–(14) becomes apparent by observing that it is dual to the sum-rate maximization subject to a total power constraint. Indeed, for every given total power allocated to user \( k \), \( P(k) \) is determined by

\[
P^*_k(\alpha) = \left( \frac{C_k}{\ln(2)} - \frac{1}{\gamma_{k,n}(S(k))} \right)^+, \quad \forall n, \quad (15)
\]

where \( C_k \) is determined by

\[
\sum_{n=1}^{N} \left( \frac{C_k}{\ln(2)} - \frac{1}{\gamma_{k,n}(S(k))} \right)^+ = P(k). \quad (16)
\]

and is computed by a standard bisection search.

Note that each user is allocated to its best subcarriers. Also noteworthy, is that the water filling is the optimal solution of convex programs with a simplex constrain set [1].

B. MMSE-BF Power Control

We confine ourselves to rate requirements which are equally split between all subcarriers. Assuming information theoretic rates, see (9), a minimum rate of \( \bar{T}_k \) is the same as a minimum signal to MAI and noise ratio (SINR) of

\[
\bar{T}_k = 2\pi f - 1, \quad k = 1, \ldots, K, \quad (17)
\]

where \( BW \) is the OFDM channel bandwidth.

By (9), the minimum power optimization problem for subcarrier \( n \) can be written as

\[
\min \big\{P_{n} \geq 0 \mid I(P_{n}) \big\} \quad (18)
\]

where \( P_n = (P_{1,n}, \ldots, P_{K,n})^T \) denotes the BS transmitter power vector for each user and \( I(P_n) = (I_1(P_n), \ldots, I_K(P_n))^T \) denotes the effective interference vector of the combined noise and MAI each user power must overcome. Indeed, by (9),

\[
I_k(P) = \mathbf{\gamma} \left( \rho_{k,k,n}^2 + \sum_{j \neq k} P_{j,n} \rho_{j,k,n} \right). \quad (19)
\]

Problem (18) is a classical power allocation problem, extensively studied since 1983, for DS-CDMA and multi-cell TDMA systems trading off between signal and interference powers. It can be solved centrally [15] using the perfect CSI. There is also a distributed iterative solution based only on each user local information [12].

With perfect CSI, the minimum powers are given by solving the linear equations

\[
P_n = I(P_n).
\]

By [12], a unique \( P_n \) exists if there a \( P_n \geq I(P_n) \). This solution, which is also the minimum power vector (component-wise), is given by inverting the following corresponding coefficient matrix.
Let \( \mathbf{\eta} = (\eta_1, \ldots, \eta_K)^T \) and \( \mathbf{\eta}_n = (\eta_{1,n}, \ldots, \eta_{K,n})^T \), where \( \eta_{k,n} = \sigma^2/\rho_{k,k,n} \). Define the \( K \times K \) matrix \( \mathbf{A}_n = [a_{k,j,n}] \) by
\[
a_{k,j,n} = \begin{cases} \rho_{k,j,n}/\rho_{k,k,n}, & \text{if } k \neq j, \\ 0, & \text{if } k = j. \end{cases}
\]

The linear equations \( \mathbf{P}_n = \mathbf{I}(\mathbf{P}_n) \) can be rewritten as
\[
\mathbf{P}_n = \mathbf{\eta} \odot (\mathbf{A}_n \mathbf{P}_n + \mathbf{\eta}_n),
\]
where \( \mathbf{\eta} \odot \mathbf{w} \) denotes a component-wise product between vectors \( \mathbf{\eta} \) and \( \mathbf{w} \), and \( \mathbf{\eta} \odot \mathbf{M} \) denotes a row-wise product between a column vector \( \mathbf{\eta} \) and a matrix \( \mathbf{M} \).

If there is a feasible \( \mathbf{P}_n \geq \mathbf{\eta} \odot (\mathbf{A}_n \mathbf{P}_n + \mathbf{\eta}_n) \), the optimal minimum power solution is given by
\[
\mathbf{P}_n = (\mathbf{I} - \mathbf{\eta} \odot \mathbf{A}_n)^{-1} \mathbf{\eta} \odot \mathbf{\eta}_n.
\]

Remark 3.1: Since \( \mathbf{A}_n \) has a zero diagonal, the existence of \( (\mathbf{I} - \mathbf{\eta} \odot \mathbf{A}_n)^{-1} \) is determined by the magnitude of \( \mathbf{\eta} \).

By sufficiently reducing \( \mathbf{\eta} \) until \( (\mathbf{I} - \mathbf{\eta} \odot \mathbf{A}_n) \) becomes sub-stochastic, a positive unique solution emerges.

Remark 3.2: Unlike with ZF-BF, the number of users with MMSE beamforming is unrestricted, but the maximum user rates is limited. Thus, using TDMA is good practice for reducing the MAI and increasing the rate to the required level.

IV. DELAY GUARANTEE SDMA-TDMA SCHEDULER

From Section III, in order to reduce (with MMSE) or completely suppress (with ZF) the MAI power and increase user rates, users must be partitioned into subsets. Small sets can benefit from high rate per user but results in a large number of sets; hence, high packet delays due to time multiplexing.

Thus, we seek for a partitioning algorithm with a set size tuning parameter. The SUS algorithm [13] has such a parameter, \( \alpha \), determining the semiorthogonality tolerance level; i.e., the minimum magnitude of the orthogonal components of the user gain vectors in a SUS set. The algorithm for selecting a single set in a single subcarrier \( n \) is given in [13]. The extension to multiple sets selection and multiple subcarriers is straightforward. Here, we merely provide guidelines to generate sets that can meet rate and delay requirements.

(i) Our numerical examples indicate that the user sets are almost insensitive in the frequency within our band of 520–620 MHz. Therefore, good practice is to generate a single group of SUS sets using the gains of a middle subcarrier. Having the same user sets across all subcarriers also benefits from an efficient and low cost fast Fourier transform (FFT) implementation of the OFDM channel.

(ii) Our numerical investigations, not reported herein, also show that it is almost infeasible to attain perfect orthogonality by antenna element placement. Consequently, a semi-orthogonality tolerance level of \( \alpha < 0.01 \) result into sets with 1-2 users. With a typical WRAN of 100 users, it implies time multiplexing over at least 50 user sets. Since current technology requires at least 1 ms for set switching, the delay would be of at least 50 ms in each direction.

Since a 50 ms delay per hop is unacceptable for most QoS applications, \( \alpha \) should be sufficiently large to result into at most 10 user sets; e.g., with 12 antenna elements and 72 randomly placed users, 10 user sets is attained with \( \alpha = 0.4 \).

(iii) One may argue that delay sensitive users can be allocated to the same user sets, which could then be scheduled more often than others within each TDMA cycle. This idea cannot be implemented from two main reasons. Firstly, user gain vectors are determined by their stationary physical locations, whereas delay sensitivity depends on the running application that varies in time. Thus, it is impractical to select user sets on a time scale that QoS requirements changes. Secondly, placing delay sensitive users in the same set may conflict with the semi-orthogonality criterion.

By the three guidelines above it is almost inevitable to schedule the SUS user sets using round-robin (RR).

V. PERFORMANCE EVALUATION

First, we compare the beamforming performance between ZF and MMSE on a typical CSIRO WRAN constellation. Then, we evaluate the impact of the free design parameters, namely the impact of the semi-orthogonality tolerance factor, \( \alpha \), and the number of the BS antenna elements, \( M \).

A. ZF vs. MMSE

Using the analytical formulas derived above, we computed the performance of the following system.

- A BS antenna array laid out in 3 levels with elements placed circularly in each level. Each antenna element is implemented by a \( \lambda/2 \) dipole, where \( \lambda \) is the electromagnetic wavelength, yielding an antenna gain of \( G_{G_0} = 1.6 \).
- An OFDM channel with 1680 subcarriers in the frequency band of 520 - 620 MHz. Each subcarrier bandwidth is 3906.25 Hz yielding a total bandwidth of 6.56 MHz in each direction.
- 72 stationary users randomly distributed within a tier of [5,100] km around the BS. Each user premise uses a single Yagi antenna yielding an antenna gain of \( G_{rcv} = 15 \).
- We assume that signals are attenuated according to free space line-of-sight path loss with a decaying exponent of 2.3.
- The background white noise power is set using the model \(-174 + 10 \log 10 \) (BW) dBm.
- User sets are selected using the SUS algorithm.

Using \( \alpha = 0.4 \), we received 10 user sets with size distribution of \( (0.0, 0.1, 0.0, 0.1, 0.1, 0.0, 0.0, 0.1, 0.2, 0.1, 0.2, 0.0) \) over 1-12 number of users. For comparison along the same baseline, we use the same user sets for both, ZF-BF and MMSE-BF. We compare the transmitter powers required to meet a downlink rate of 9.6 Mb/s per user.

Remark 5.1: The required rate of our broadband-to-the-bush WRAN is actually 12 Mb/s. However, since MMSE-BF cannot support it for this specific constellation, we compare the BF methods for the largest rate that MMSE can support.

Fig. 1 depicts the user rates with ZF-BF. The various rates around 9.6 Mb/s results from the numerical method determining the water-filling powers within a margin error of 5%. The total power required to meet a rate of 9.6 Mb/s
across all users is 81 Watts. The transmission and actual MAI powers are summarized in the first row of Table I. Observe that although in theory the MAI powers with ZF is zero, in practice it cannot be completely suppressed due to numerical precision in calculating the pseudo-inverse for the weight matrix. Additionally, the weight matrix is calculated from estimated gains which are error prone. Nevertheless, MAI power is 5 order of magnitude lower than the background white noise, and therefore negligible.

Fig. 2 depicts the user rates with MMSE-BF. The computed rates match exactly 9.6 Mb/s since the matrix inversion is exact and the computation has a close to zero error tolerance. The transmission and MAI powers are summarized in the second row of Table I. Observe that unlike with ZF, the MAI powers with MMSE is 4 order of magnitude higher than the background white noise. This highlights its impact on the transmission power control for MMSE.

A quite unexpected result is the total transmission power required for attaining 9.6 Mb/s per user. **MMSE requires 1155 Watts compared to 81 Watts required by ZF.** It places ZF-BF as a substantially “greener” solution.

This typical WRAN example demonstrates that energy and bandwidth wise, ZF is far more superior to linear MMSE-ZF. Primarily, it requires as low as 81 Watts (compared to 200-300 Watts specified for WiMAX, and 1155 Watts required with MMSE-BF). Additionally, its MAI suppressing enables to increase each user rate by merely increasing its transmission power. Since MAI power is 5 order of magnitude lower than the background white noise, the power can be increased by 4 order of magnitude until it will affect the rate of other users. With MMSE-BF on the other hand, the maximum attained rate is MAI limited and cannot support 12 Mb/s in our example.

Another question may be raised. The comparison above uses the information theoretic rates; how the two BF methods compare with respect to the rates excluding bit error rate (BER)? MMSE is designed to minimize the mean error rate, so it might still be better than ZF. Note however that the error rates depend on the modulation/demodulation and coding techniques and in general, BER is monotonic with SINR. Since both methods must attain the same SINR in order to achieve their the same information theoretic capacity, their expected BER and error-free rates should also be the same.

**B. The Impact of the Semi-orthogonality Tolerance Factor**

Above, we established the superiority of ZF-BF over MMSE-BF; next we examine the impact of the semi-orthogonality tolerance factor, $\alpha$. By definition, one would expect that when $\alpha$ increases, the number of SUS user sets would decrease. Also that the magnitude of the combined channel gains, $\{\gamma_{k,n}(S_n)\}$ (see (10)), would decrease. The latter should consequently result in higher transmission power.

Indeed, considering the system of Subsection V-A with 96 users and bandwidth requirements of 12 Mb/s per user, the impact of $\alpha$ is given in Table II. In those cases where $\alpha$ is annotated by ‘x’, one of the users has $\gamma = 0$ preventing it from any successful transmission.

It can be observed that the trend of the number of sets is as expected but there are some irregularities with the expected required powers. One irregular point occurs when the power is reduced by increasing $\alpha$ from 0.1 to 0.2. This is explained by the fact that one of the users with $\alpha = 0.2$ has $\gamma = 0$ and therefore is not allocated any power. The same phenomena occurs when increasing $\alpha$ from 0.5 to 0.52. The case of $\alpha = 0.4$ is particularly inefficient where a large decrease in the number of sets result in a sharp increase in the transmission powers with one user having $\gamma = 0$.

One observation from this example is that the final selection of $\alpha$ should be considered carefully. Although the general trend is approximately expected, slight variation can make a difference in the number of supported users. Another observation is that to meet delay requirements and multiplex only among 11 user sets, rather than among 33 sets, the transmission power needs to be roughly doubled.

**C. The Impact of the Number of BS Antenna elements**

Here, we examine the impact of the number of antenna elements, $M$, on the transmission power. Since the antenna
elements require minimum distance between each other, the larger is $M$ the larger are the plates the elements are placed on. This could present technical construction issues.

On the other hand, a larger $M$ will increase the size of each semi-orthogonality user set (bounded by $M$), which will subsequently decrease the number of user sets. Thus, by Subsection V-B, one can expect a lower transmission power.

For the system of Subsection V-A with 72 users and bandwidth requirements of 12 Mb/s per user, the impact of $M$ is given in Table III. For the same baseline comparison, $\alpha$ is tuned to attain 9-11 user sets for each $M$, except for $M = 6$. The power penalty for $M = 6$ amounts to $\approx 10^4$ Watts.

It can be observed that the number of antenna elements has a notable impact when increasing $M$ from 18 to 21 saving about 180 Watts. When $M$ equals 9,12,15 and 18, the powers are very similar ranging between 216 to 237 Watts not exhibiting monotonicity. For $M = 6$, we can achieve a power of 235 Watts but only if we use 15 user sets. When reducing the number of sets to 13 by using $\alpha = 0.7$, the power jumps to an order of $10^5$ Watts.

This example demonstrates that the number of BS antenna elements must be designed in a very early stage of the deployment since it may involve element construction that could be difficult to change in later stages.

VI. CONCLUSIONS

Applications requiring QoS are expanding their reach to remote rural areas where optical fibers are not economical. For those areas, wireless access via OFDM MIMO with multiuser beamforming appears as the best economical solution.

An economical broadband wireless solution needs to minimize the number of BS and support high bandwidth and low delay per end user. These two metrics conflict each other. Indeed, with a smaller number of BS, each one must cover more users. More users per BS require more multiplexing which result in lower bandwidth and higher delay per user.

SDMA multiplexing favors delay-sensitive users since they share the channel concurrently. However, the generated MAI could reduce the bandwidth per user. When MAI exists, higher transmission power could not always increase the per-user bandwidth, e.g., with DS-CDMA. We showed that it is also a problem with OFDM MIMO when MMSE-BF is used. Conversely, ZF-BF practically suppresses all MAI, hence bandwidth per user can be increased by increasing its transmission power. This is demonstrated in this paper.

Although ZF-BF can meet the bandwidth requirement, it may not meet the delay requirement since the number of SDMA multiplexed users is limited to the BS antenna elements. As a side remark, MMSE-BF also suffers from a size limitation due to MAI. We addressed the size limitation by TDMA multiplexing and limiting the number of user sets.

Overall, we showed that energy-wise, ZF-BF is superior to MMSE-ZF and by using parameters $\alpha$ and $M$, we can maintain the required bandwidth and delay with only a total power budget of several tens of Watts.

REFERENCES