POWER ALLOCATION AND OPTIMAL TX/RX STRUCTURES FOR MIMO SYSTEMS

R. Cendrillon, O. Rousseaux and M. Moonen
SCD/ESAT, Katholieke Universiteit Leuven, Belgium
{raphael.cendrillon, olivier.rousseaux, marc.moonen}@esat.kuleuven.ac.be

Etienne Van den Bogaert, Jan Verlinden
Alcatel Bell, Antwerp, Belgium
{etienne.van_den_bogaert, jan.vj.verlinden}@alcatel.be

In this paper we investigate power allocation in MIMO-OFDM systems. We describe the optimal power allocations under two different constraints: a constraint on the total power of all transmitters (TXs) which is applicable in wireless applications, and a constraint on the power of each TX which is more relevant in wireline applications.
We describe the optimal TX/RX structure which in combination with the optimal power allocation achieves the MIMO-OFDM channel capacity (under the chosen constraints), with low complexity. Simulations show the benefits of using a total power constraint in place of a per-TX power constraint are largest when the TXs see channels with significantly different attenuations.

1 INTRODUCTION

In highly frequency selective channels the loading of power across frequency has a significant impact on system performance. The well known waterfilling algorithm[1] describes the optimal power allocation for single-input single-output (SISO) channels. In this paper we investigate the problem of power allocation in multi-input multi-output (MIMO) systems. These systems are encountered in the wireless environment where multiple TX/RX antennas are used to increase data-rate and mitigate channel fades through the use of spatial multiplexing and TX/RX diversity respectively[2]. The MIMO approach also finds application in wireline environments like digital subscriber line (DSL) where it can be used to enable crosstalk cancellation[3].

2 MIMO SYSTEM MODEL

We restrict our attention to OFDM transmission. The signal sent by each transmitter has a cyclic prefix (CP) appended. We assume the CP is of sufficient length so transmission on each tone can be modeled independently. Transmission of one OFDM-block is then described

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{z}_k$$

where $\mathbf{x}_k \triangleq \left[x_k^1, \cdots, x_k^N\right]^T$ is the vector of transmitted signals on tone k. There are N co-located transmitters (TX) and co-located receivers (RX) in the system and x_k^n is the signal of TX n at tone k. $\mathbf{y}_k$ and $\mathbf{z}_k$ have similar structures. $\mathbf{y}_k$ is the vector of received signals on tone k. $\mathbf{z}_k$ is the vector of additive noise on tone k. We assume $\mathcal{E}\left\{\mathbf{z}_k\mathbf{z}_k^H\right\} = \sigma_k^2\mathbf{I}_N$. This is without loss of generality since a noise-whitening pre-filter can be applied at the RXs. $\mathbf{H}_k$ is the $N \times N$ channel transfer matrix on tone k. $h_k^{n,m} \triangleq [\mathbf{H}_k]_{n,m}$ is the channel from TX m to RX n on tone k. We define the transmit correlation matrix on tone k $\mathbf{S}_k \triangleq \mathcal{E}\left\{\mathbf{x}_k\mathbf{x}_k^H\right\}$ and its elements $s_k^{n,m} \triangleq [\mathbf{S}_k]_{n,m}$

The capacity of the system is

$$\begin{aligned} C &= \sum_k I(\mathbf{x}_k; \mathbf{y}_k) \qquad (1) \\ &= \sum_k \log\left|\mathbf{I}_N + \sigma_k^{-2}\mathbf{H}_k\mathbf{S}_k\mathbf{H}_k^H\right| \end{aligned}$$

where $I(\mathbf{a}; \mathbf{b})$ denotes the mutual information between $\mathbf{a}$ and $\mathbf{b}$.

3 POWER ALLOCATION IN MIMO SYSTEMS

We now investigate power allocations $\mathbf{S}_k$ that maximize the capacity of the MIMO channel C.

3.1 POWER CONSTRAINT FOR ALL TRANSMITTERS

We first focus our attention on the maximization of C under a constraint P on the total power of all TXs

$$\max_{\{\mathbf{S}_k\}_{k=1\ldots K}} C \quad \text{s.t.} \quad \mathbf{S}_k \in \mathbb{R}_+^N,\ \forall k \qquad (2)$$

$$\sum_n \sum_k s_k^{n,n} \leq P \qquad (3)$$

This constraint is well suited to wireless applications and is then motivated by considering the limitations on the analog front-end (AFE) which drives the multi-element antenna. Note that $\mathbb{R}_+^N$is the set of all positive semi-definite matrices of size $N \times N$. Naturally any valid transmit PSD must be within this set.

Using the SVD $\mathbf{H}_k \overset{\text{svd}}{=} \mathbf{U}_k\Phi_k\mathbf{V}_k^H$ where $\Phi_k \triangleq \operatorname{diag}\left\{\varphi_k^1, \ldots, \varphi_k^N\right\}$. The optimal transmit correlation matrix is then

$$\mathbf{S}_k = \mathbf{V}_k\mathbf{D}_k\mathbf{V}_k^H \qquad (4)$$

where

$$\mathbf{D}_k \triangleq \left[\frac{1}{\lambda}\mathbf{I}_N - \sigma_k\Phi_k^{-2}\right]^+$$

and $[x]^+ \triangleq \max(0, x)$. $\frac{1}{\lambda}$ is the waterfilling level and is chosen such that the TX power constraint (3) is met with equality. Note that the diagonal values of $\mathbf{D}_k$ are found through a conventional waterfilling algorithm. Waterfilling is applied to the equivalent channel $\tilde{\mathbf{h}}$ which is formed by concatenation of all singular values φ_k^n

$$\tilde{\mathbf{h}} \triangleq \left[\varphi_1^1, \ldots \varphi_1^N, \ \ldots \ , \varphi_K^1, \ldots, \varphi_K^N\right]$$

Proof: See Appendix.

3.2 POWER CONSTRAINT PER TRANSMITTER

In many scenarios it is more relevant to consider a constraint P_n on the power of each TX instead of a constraint P on the total power of all TXs. An example is DSL transmission. In DSL the use of several modems as a MIMO system yields significant benefits in terms of interference cancellation[3]. In this case the limitation is on the transmit power that the AFE of each modem can support.

Maximizing capacity under a constraint on each TX leads to the optimization problem

$$\max_{\{\mathbf{S}_k\}_{k=1\ldots K}} C \quad \text{s.t.} \quad \mathbf{S}_k \in \mathbb{R}_+^N, \ \forall k \tag{5}$$

$$\sum_k s_k^{n,n} \leq P_n, \ \forall n \tag{6}$$

The object function C is concave while the constraints form a convex set of feasible solutions. As a result we can solve this problem using standard convex optimisation techniques.

Unfortunately we do not know of a closed form solution to the optimisation (5). This is the subject of ongoing work. Instead we use standard numerical techniques for solving convex problems (e.g. interior point methods).

4 OPTIMAL TX/RX STRUCTURE FOR MIMO SYSTEMS

We now describe the optimal TX and RX structure from [4] which in combination with the power allocation of the previous sections achieves MIMO channel capacity.

Using the eigenvalue decomposition $\mathbf{S}_k \overset{\text{eig}}{=} \mathbf{Q}_k \mathbf{M}_k \mathbf{Q}_k^H$. Define the equivalent channel $\overline{\mathbf{H}}_k \triangleq \mathbf{H}_k \mathbf{Q}_k \mathbf{M}_k^{1/2}$ and its SVD $\overline{\mathbf{H}}_k \triangleq \overline{\mathbf{U}}_k \overline{\Phi}_k \overline{\mathbf{V}}_k^H$.

Begin with a set of normalized frequency domain symbols $\overline{\mathbf{x}}_k$ which are generated by the encoder at tone k. These are normalized such that $\mathcal{E}\left\{\overline{\mathbf{x}}_k \overline{\mathbf{x}}_k^H\right\} = \mathbf{I}_N$. Before transmission we apply the pre-filter $\mathbf{P}_k \triangleq \mathbf{Q}_k \mathbf{M}_k^{1/2} \overline{\mathbf{V}}_k$ to the normalized symbols. Hence $\mathbf{x}_k = \mathbf{P}_k \overline{\mathbf{x}}_k$ and the transmitted signal $\mathbf{x}_k$ has the optimal PSD, ie. $\mathcal{E}\left\{\mathbf{x}_k \mathbf{x}_k^H\right\} = \mathbf{S}_k$.

At the RXs we apply the filter $\mathbf{W}_k \triangleq \overline{\Phi}_k^{-1} \overline{\mathbf{U}}_k^H$. Our estimate of the transmitted symbols is thus formed

$$\begin{aligned} \widehat{\mathbf{x}}_k &= \mathbf{W}_k \mathbf{y}_k \\ &= \mathbf{W}_k \left(\mathbf{H}_k \mathbf{P}_k \overline{\mathbf{x}}_k + \mathbf{z}_k\right) \\ &= \overline{\mathbf{x}}_k + \overline{\Phi}_k^{-1} \overline{\mathbf{z}}_k \end{aligned}$$

where $\overline{\mathbf{z}}_k \triangleq \overline{\mathbf{U}}_k^H \mathbf{z}_k$. Note that $\mathcal{E}\left\{\overline{\mathbf{z}}_k \overline{\mathbf{z}}_k^H\right\} = \mathcal{E}\left\{\mathbf{z}_k \mathbf{z}_k^H\right\} = \sigma_k^2 \mathbf{I}_N$.

Note that under the total power constraint (3): $\mathbf{Q}_k = \mathbf{V}_k$, $\mathbf{M}_k = \mathbf{D}_k$, $\overline{\mathbf{H}}_k \triangleq \mathbf{U}_k \Phi_k \mathbf{D}_k^{1/2}$, $\overline{\mathbf{U}}_k = \mathbf{U}_k$, $\overline{\Phi}_k = \Phi_k \mathbf{D}_k^{1/2}$ and $\overline{\mathbf{V}}_k = \mathbf{I}_N$. Hence we have the following simplifcations: $\mathbf{P}_k = \mathbf{V}_k \mathbf{D}_k^{1/2}$ and $\mathbf{W}_k = \mathbf{D}_k^{-1/2} \Phi_k^{-1} \mathbf{U}_k^H$.

Applying an ideal single-input single-output (SISO) code to the scalar stream $\bar{x}_k^n$ allows us to achieve the rate

$$c_{\mathrm{SISO}}^n = \sum_k \log_2 \left(1 + \sigma_k^{-2} \left(\overline{\varphi}_k^n\right)^2\right) \text{ bps/Hz}$$

with vanishing probability of error. This leads to a total rate

$$\begin{aligned}
C_{\mathrm{SISO}} &= \sum_n c_{\mathrm{SISO}}^n \\
&= \sum_k \log_2 \left|\mathbf{I}_N + \sigma_k^{-2} \overline{\Lambda}_k^2\right| \\
&= \sum_k \log_2 \left|\mathbf{I}_N + \sigma_k^{-2} \mathbf{H}_k \mathbf{S}_k \mathbf{H}_k\right| \\
&= C
\end{aligned}$$

where C is the capacity of the MIMO channel as defined in (1).

So using independent SISO encoders/decoders for the N scalar streams $x_k^n, y_k^n, \ \forall n$ plus simple linear pre and post-filtering allows us to achieve the full capacity of the MIMO channel. Note this has a much lower complexity than using a maximum likelihood (ML) multi-input multi-output (MIMO) encoder/decoder for the N dimensional data-stream $\mathbf{x}_k, \mathbf{y}_k$.

5 PERFORMANCE

Operating under a total power constraint (3) rather than a power constraint on each TX (6) gives an extra degree of freedom in the power allocation problem. We now investigate the performance of the optimal power allocations under both constraints.

Our simulation uses a Rayleigh channel model with $K = 16$ tones. The elements of the channel matrix $\mathbf{H}_k$ at each tone have independent, Gaussian distributions. The benefit of using a total power constraint over a power constraint on each TX is largest when the TXs see channels with significantly different gains. To introduce this into our simulation we define the parameter α which determines the spread in attenuation of the channels seen by each of the TXs.

$$\mathbf{H}_k = \left[\mathbf{h}_k^1 \ \cdots \ \mathbf{h}_k^N\right]$$

$$\mathcal{E}\left\{\mathbf{h}_k^n \left(\mathbf{h}_k^n\right)^H\right\} = \left(\frac{1}{\alpha}\right)^{n-1} \mathbf{I}_N$$

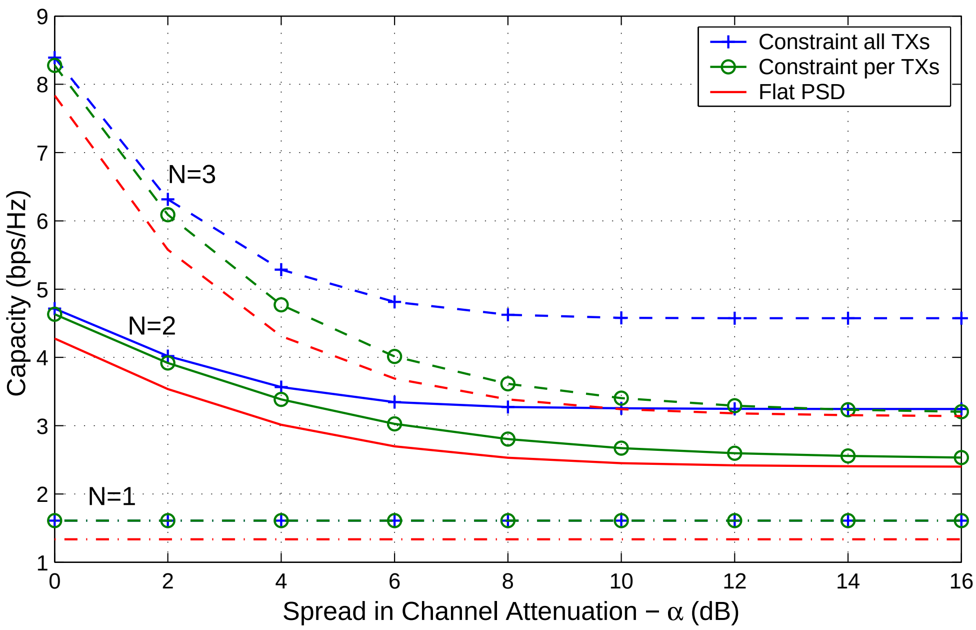

Figure 1: Capacity vs. Spread in Channel Attenuation - α

So we can expect the channels from the $n+1$th TX to be attenuated by a factor α more than the channels from the nth TX. Shown in Fig. 1 is capacity versus the spread in channel attenuation α. This is plotted for different numbers of TXs N. The capacity is shown with flat transmit PSDs where

$$\mathbf{S}_k = \frac{P}{KN}\mathbf{I}_N$$

Also plotted is the capacity with optimal transmit PSDs under a total power constraint and a per-TX power constraint as described in Sec. 3.1 and 3.2 respectively.

As can be expected, the freedom to shift power from one TX to another, as provided by the total power constraint, gives the largest gains when the TXs see channels with significantly different attenuations, that is for high values of α.

6 CONCLUSIONS

In this paper we investigated the problem of power allocation in MIMO-OFDM channels. We discussed the optimal power allocation problem under a constraint on the total power of all TXs (3), and under a constraint on the power of each TX (6). We also described the optimal TX/RX structure which in combination with the optimal power allocations achieves MIMO-OFDM channel capacity with low complexity.

It was seen that applying a power constraint to the total power of all TXs provides an extra degree of freedom in power allocation. This allows power to be redistributed from TXs which have poor channels to TXs whose channels have low attenuation. As can be expected, using the total power constraint in place of the per-TX power constraint provides significant gains. These gains are largest when the TXs see channels with significantly different attenuations. This was confirmed by simulation.

In this paper we used numerical methods to solve the problem of power allocation with a per-TX power constraint. A closed form solution for this problem is an important extension to this work.

APPENDIX

Using the eigenvalue decomposition

$$\mathbf{S}_k \overset{\text{eig}}{=} \mathbf{Q}_k \mathbf{D}_k \mathbf{Q}_k^H$$

where $\mathbf{D}_k = \mathrm{diag}\left\{d_k^1, \ldots, d_k^N\right\}$ is a diagonal matrix whose diagonal elements contain the eigenvalues of $\mathbf{S}_k$ and $\mathbf{Q}_k$ contains the corresponding eigenvectors. The constraint

$$\mathbf{S}_k \in \mathbb{R}_+^N \quad \leftrightarrow \quad d_k^n \geq 0,\ \forall n \tag{7}$$

Now, since $\mathbf{D}_k = \mathbf{Q}_k^H \mathbf{S}_k \mathbf{Q}_k$, (7) can be rewritten

$$\left(\mathbf{q}_k^n\right)^H \mathbf{S}_k \mathbf{q}_k^n \geq 0,\ \forall n$$

Define the Lagrangian of the optimisation (2)

$$J = C + L_1 + L_2$$

where

$$\begin{aligned} L_1 &\triangleq \sum_k \sum_n \mu_k^n \left(\mathbf{q}_k^n\right)^H \mathbf{S}_k \mathbf{q}_k^n \\ L_2 &\triangleq \lambda \left(P - \sum_n \sum_k s_k^{n,n}\right) \end{aligned}$$

Note that the objective C is concave whilst the constraints form a convex set of feasible solutions. Thus the K.K.T. conditions

$$\nabla_{\mathbf{S}_k} J = 0 \tag{8}$$

$$\mu_k^n d_k^n = 0,\ \forall k, n \tag{9}$$

$$\lambda \left(P - \sum_n \sum_k s_k^{n,n}\right) = 0 \tag{10}$$

are sufficient for optimality

KKT CONDITION 1

Now

$$\begin{aligned}\nabla_{\mathbf{S}_k} C &= \left[\mathbf{S}_k + \sigma_k^2 \left(\mathbf{H}_k^H \mathbf{H}_k\right)^{-1}\right]^{-1} \\ \nabla_{\mathbf{S}_k} L_1 &= -\lambda \mathbf{I}_N \\ \nabla_{\mathbf{S}_k} L_2 &= \mathbf{Q}_k \mathrm{diag}\left\{\mu_k^1, \ldots, \mu_k^N\right\} \mathbf{Q}_k^H\end{aligned}$$

Hence (8) implies

$$\begin{aligned}\left[\mathbf{S}_k + \sigma_k^2 \left(\mathbf{H}_k^H \mathbf{H}_k\right)^{-1}\right]^{-1} &= \lambda \mathbf{I}_N - \mathbf{Q}_k \mathrm{diag}\left\{\mu_k^1, \ldots, \mu_k^N\right\} \mathbf{Q}_k^H \\ &= \mathbf{Q}_k \left[\lambda \mathbf{I}_N - \mathrm{diag}\left\{\mu_k^1, \ldots, \mu_k^N\right\}\right] \mathbf{Q}_k^H\end{aligned}$$

and

$$\begin{aligned}\mathbf{S}_k &= \mathbf{Q}_k \mathrm{diag}\left\{\lambda - \mu_k^1, \ldots, \lambda - \mu_k^N\right\}^{-1} \mathbf{Q}_k^H - \sigma_k^2 \left(\mathbf{H}_k^H \mathbf{H}_k\right)^{-1} \\ &= \mathbf{Q}_k \mathrm{diag}\left\{\lambda - \mu_k^1, \ldots, \lambda - \mu_k^N\right\}^{-1} \mathbf{Q}_k^H - \sigma_k^2 \mathbf{V}_k \Phi_k^{-2} \mathbf{V}_k^H\end{aligned}$$

In the proposed solution (4), $\mathbf{Q}_k = \mathbf{V}_k$ so (8) is satisfied if

$$\mathbf{S}_k = \mathbf{V}_k \left[\mathrm{diag}\left\{\lambda - \mu_k^1, \ldots, \lambda - \mu_k^N\right\}^{-1} - \sigma_k^2 \Phi_k^{-2}\right] \mathbf{V}_k^H$$

KKT CONDITION 2

From the previous condition we have

$$\mathbf{D}_k = \mathrm{diag}\left\{\lambda - \mu_k^1, \ldots, \lambda - \mu_k^N\right\}^{-1} - \sigma_k^2 \Phi_k^{-2}$$

Examining (9) we find two cases.

Case 1 $d_k^n > 0$

$d_k^n > 0$ implies $\mu_k^n = 0$ hence

$$d_k^n = \frac{1}{\lambda} - \sigma_k^2 \left(\varphi_k^n\right)^{-2} > 0$$

and

$$d_k^n = \left[\frac{1}{\lambda} - \sigma_k^2 \left(\varphi_k^n\right)^{-2}\right]^+, \ \forall n \text{ s.t. } d_k^n > 0$$

Case 2 $d_k^n = 0$

$d_k^n = 0$ implies $\mu_k^n = \lambda - \left(\sigma^2 \left(\varphi_k^n\right)^{-2}\right)^{-1}$. Hence

$$\begin{aligned} \lambda &= \mu_k^n + \left(\sigma^2 \left(\varphi_k^n\right)^{-2}\right)^{-1} \\ &\geq \left(\sigma^2 \left(\varphi_k^n\right)^{-2}\right)^{-1} \end{aligned}$$

since $\mu_k^n \geq 0$. Hence

$$\frac{1}{\lambda} - \sigma^2 \left(\varphi_k^n\right)^{-2} \leq 0$$

and

$$d_k^n = \left[\frac{1}{\lambda} - \sigma^2 \left(\varphi_k^n\right)^{-2}\right]^+ = 0, \ \forall n \text{ s.t. } d_k^n = 0$$

Combining both cases yields

$$\mathbf{D}_k = \left[\frac{1}{\lambda}\mathbf{I}_N - \sigma_k^2 \Phi_k^{-2}\right]^+$$

KKT CONDITION 3

Note that $\lambda = 0$ implies $s_k^{n,n} = \infty$. Clearly this violates the power constraint (3). Hence $\lambda > 0$. Using (10) this implies

$$P = \sum_n \sum_k s_k^{n,n}$$

So any optimal solution must meet the total power constraint with equality.

At this point notice that the solution

$$\mathbf{S}_k = \mathbf{V}_k \left[\frac{1}{\lambda}\mathbf{I}_N - \sigma_k \Phi_k^{-2}\right]^+ \mathbf{V}_k^H$$

satisfies all 3 KKT conditions and thus is optimal.

REFERENCES

[1] T. Cover and J. Thomas, *Elements of Information Theory*. Wiley, 1991.

[2] A. Naguib, N. Seshadri, and A. Calderbank, “Increasing Data Rate over Wireless Channels,” *IEEE Signal Processing Mag.*, pp. 76–92, May 2000.

[3] G. Ginis and J. Cioffi, “Vectored Transmission for Digital Subscriber Line Systems,” *IEEE J. Select. Areas Commun.*, vol. 20, pp. 1085–1104, June 2002.

[4] G. Taubock and W. Henkel, “MIMO Systems in the Subscriber-Line Network,” in *Proc. of the 5th Int. OFDM-Workshop*, pp. 18.1–18.3, 2000.