MIMO SYSTEMS IN FREQUENCY SELECTIVE FADING: CHANNEL MODEL, SHANNON CAPACITY, LINEAR CODING

Venceslav Kafedziski
Faculty of Electrical Engineering
University “SS Cyril and Methodius”
1000 Skopje, Republic of Macedonia
kafedzi@etf.ukim.edu.mk

Abstract—We address MIMO systems in the presence of frequency selective fading. We first describe our channel model, then we derive capacity formula, and finally we develop a linear coding procedure for the obtained channel model. Starting from a physical mobile channel model with known scatterer geometry, we develop a MIMO FIR channel model. Our channel model results in stationary, ergodic vector process for the FIR coefficients. We derive capacity of this MIMO FIR channel, assuming that both transmitter and receiver have perfect channel side information. To derive capacity, we use random ergodic operator theory. We develop an approach of linear coding of discrete time continuous amplitude sources over MIMO FIR channel using source and channel eigendecomposition on finite blocks. Thus, we reduce the problem to linear coding of vector sources on vector channels. For each channel frame we use MIMO-OFDM. We analyze the impact of antenna correlation on performance.

1. INTRODUCTION

MIMO systems have recently attracted much attention [1]. By using multiple antennas at both transmitter and receiver, channel capacity can be increased [2]. It can be shown that capacity increases linearly with the number of antennas [3], [4] under some assumptions for MIMO channel. MIMO systems are currently considered for application in broadband wireless fixed cellular networks. They can seriously improve the economics of deployment, thanks to increased coverage and data rates MIMO can offer. Potential applications in WLAN have resulted in the appearance of new 802.11n standard and first products are envisioned to appear by the end of 2004. They will have increased throughput and range compared to current WLAN standards. The next great application of MIMO is undoubtedly the next generation broadband mobile packet-based cellular networks. The initial work on MIMO systems has been directed towards flat fading MIMO channels. The demand of high transmission rates requires the use of wideband systems which are affected by a frequency selective fading. Therefore, in this work we address MIMO systems in frequency selective fading. Assuming slow fading, and using the block fading channel model, to mitigate the effects of frequency selectivity we use MIMO-OFDM. OFDM techniques have already found widespread applications, such as in ADSL, DVB-T, DAB, WLAN and broadband fixed wireless access. OFDM systems present an excellent alternative to systems with equalization [5], [6]. In OFDM systems channel is divided into a number of parallel narrowband subchannels. As a result of that, it can be assumed that each of the subchannels experiences frequency non-selective fading. In other words, intersymbol interference is eliminated.

There is much literature on spatial channel modeling, which has implications for MIMO systems. Lot of work has been done on spatial channel models for applications with smart antennas, and those models can also be used with MIMO systems. In fact, channel models developed for SISO systems that use angles of arrival of the incoming waves can be used for MIMO systems as well, by assuming a number of antennas at both transmitter and receiver. We derived such SISO channel models in [7], [8], which we extended to MIMO channels in [9].

Capacity of MIMO channel with flat fading was derived in [3]. Distributions of eigenvalues of Wishart random matrices were used. Limits when the number of antennas increases to infinity were determined. In [10] channel state information to both transmitter and receiver was assumed and standard water-filling was used to maximize the mutual information. Correlation among channels was also addressed in this work. Capacity of MIMO channel with frequency selective fading and no correlation among component channels was derived in [11], [12] under the assumption that decoupling of time and frequency variable can be assumed (slow fading). In [13], [14] we derived a capacity expression in terms of limiting eigenvalue distribution function of a sequence of channel matrices of increasing size that holds under most general conditions when decoupling of time and frequency variable cannot be assumed (fast fading).

Linear coding of analog stationary sources on SISO channels described by linear filters is developed in [15], and on MIMO channels in [16]. We use the solution for finite blocks of data, where this problem reduces to coding of vector sources on vector channels, described in [17]. We addressed linear coding of Gaussian sources and images on mobile SISO channels in [18], [7] and linear coding of Gaussian sources on MIMO channels in [19], [9]. Related work is described in [20], [21] where linear coding for MIMO is studied in the context of equalization. We consider a more general setting of coding a non-white source on MIMO FIR channels with stationary, ergodic coefficients, so that linear coding becomes a joint source channel coding procedure.

In Section II we derive a MIMO FIR channel model from a physical mobile channel model. In Section III we prove...
that Shannon capacity of MIMO FIR channel with stationary, ergodic coefficients exists and is given in terms of limiting eigenvalue distribution of channel matrices of increasing size. In section IV we describe decomposition of slowly varying MIMO FIR channels by DFT and eigenanalysis resulting in a MIMO OFDM system. In Sections V and VI respectively, we describe optimal linear coding for vector channels, and for mobile channels with slow time variation. We present the simulation results in Section VII and conclude.

II. FROM A PHYSICAL MOBILE CHANNEL TO A MIMO FIR CHANNEL

Literature on MIMO channel modeling is abundant and includes [22], [23], [24], [10], [25], [26]. We assume that the principal mechanism of wave propagation is scattering, i.e. that LOS wave does not exist. The distribution of scatterers can be arbitrary, but usually urban models assume uniform distribution of scatterers in a circle around the mobile. This is similar to Jakes model, where the angle distribution of the waves incident on the mobile is assumed uniform. In suburban or rural models it is assumed that the scatterers are distributed in an ellipse with the mobile and the base station in its foci. Both models are representatives of WSSUS channel models [27], [8]. We model the physical mobile channel by placing a number of scatterer clusters in an ellipse with the mobile (subscriber unit SU) and the base station (BS) in its foci [7], [9]. This is described in Figure 1. In our example BS is transmitting, and the mobile (SU) is receiving, $\theta_T$, $\theta_R$ and $\theta_v$

![Fig. 1. Physical Mobile Channel](image)

determine transmit antenna position, receive antenna position, and the velocity vector position, relative to the line connecting the base station and the mobile, respectively. The transmit antenna and the receive antenna spacing is equal to $d_T = k_T \lambda$ and $d_R = k_R \lambda$ respectively, where $\lambda = c/f_0$ is the carrier wavelength ($f_0$ is the carrier frequency) and $k_R$ and $k_T$ are design parameters.

We assume that each scatterer cluster consists of a large number of elementary scatterers distributed uniformly in a region with known geometry. The channel impulse response (CIR) of the channel between $j$-th transmit and $i$-th receive antenna can be written as

$$e^{(i,j)}(t;\tau) = \sum_{m=1}^{M} A^{(i,j)}_{m}(t) \delta(\tau - \tau_m)$$  (1)

where $A^{(i,j)}_{m}(t)$ are zero mean complex Gaussian processes for $i = 1, \ldots, R$, $j = 1, \ldots, T$ and $M$ is the number of scatterer regions. $A^{(i,j)}_{m}(t)$ are obtained by summing a large number of elementary scatterer waves with random initial phases, phase delays computed from propagation paths (time delays), and elementary Doppler shifts, depending on the angles between the velocity vector and the elementary scatterer wavefronts impinging on the respective receive antenna. Gaussianity of $A^{(i,j)}_{m}(t)$ is a consequence of Central Limit Theorem. The positions of transmit antennas affect the phase delays only. The positions of receive antennas affect both the phase delays and Doppler shifts. We obtain equivalent time varying CIR’s $h^{(i,j)}(t;\tau), t = 1, \ldots, R, j = 1, \ldots, T$, by including the transmit and receive filter impulse responses. The concatenation of the transmit and receive filter represents a Nyquist filter. By sampling $h^{(i,j)}(t;\tau)$ time and delay variable at symbol rate we obtain the discrete time CIR’s $h^{(i,j)}(n), n \in \mathbb{Z}, l \in \mathbb{Z}$, where $l$ is the discrete delay variable, and $n$ is the discrete time variable. Note that this is IIR (infinite impulse response) in the delay variable. We obtain FIR (finite impulse response) in the delay variable by keeping the coefficients with significant power only and setting the rest of the coefficients to zero. In other words, we obtain $h^{(i,j)}(n), l = 0, \ldots, \lambda^{(i,j)}, n \in \mathbb{Z}$. The scatterer geometry, the distance $r_0$ between the mobile and the base station, the carrier frequency $f_0$ and the mobile velocity vector $v$ determine channel time and frequency dispersion. In addition to that, the transmission symbol rate $f_s$ affects time and frequency dispersion as well, such that higher $f_s$ means higher time dispersion, but lower frequency dispersion. We introduce two parameters that define the amount of ISI (time dispersion) and time variation (frequency dispersion). The amount of ISI depends on the parameter $\alpha = c/(f_s r_0)$, where $f_s$ is the symbol rate and $r_0$ is the distance between the transmitter and receiver. Lower values of $\alpha$ result in higher amount of ISI, i.e. in a bigger number of non-zero FIR coefficients. We measure the channel time variation in terms of the normalized maximum Doppler frequency $f_d/f_s$, where $f_d = \|v\|f_0/c$. The channel symbol frame size, on which a single channel estimation at the receiver is performed, is chosen such that the channel time variation during a single frame is insignificant. Under certain assumptions, in [7] we have shown that for the assumed physical model, the CIR of the SISO system is a stationary, ergodic process. Similarly, the vector CIR process $[h^{(1,1)}_{0}(n), \ldots, h^{(1,1)}_{0}(n), h^{(R,T)}_{0}(n), \ldots, h^{(R,T)}_{0}(n)], n \in \mathbb{Z}$ of the MIMO system is a stationary, ergodic vector process.

Correlation between component CIR processes $h^{(i,s)}_{l}$ and $h^{(k,s)}_{l}$ at lag $l$ is computed as follows:

$$\frac{\mathbb{E}[|h^{(i,s)}_{l}|^2]|h^{(k,s)}_{l}|}{\sqrt{\mathbb{E}[|h^{(i,s)}_{l}|^2]|\mathbb{E}[|h^{(k,s)}_{l}|^2]|}}$$  (2)

We determined that, in our channel model, increasing the size of the scatterer clusters, together with increasing $d_T$ and $d_R$, results in decorrelation of component CIR processes. It is well known that correlation decreases capacity when transmitter has no channel state information. When transmitter has perfect channel state information, the capacity of MIMO correlated channel can even increase in the low SNR region, compared to uncorrelated MIMO channel, because of the water filling on the transmitted powers, as observed in [10], [28]. In the high
SNR region, capacity of MIMO correlated channel always suffers performance loss.

III. CAPACITY OF MIMO FIR CHANNELS WITH STATIONARY, ERGODIC COEFFICIENTS

Consider a MIMO-FIR channel with $T$ transmit and $R$ receive antennas. Each of $T \times R$ channels is described by its impulse response $[h_{i,j}(1), h_{i,j}(2), \ldots, h_{i,j}(R,T)]$ for $i = 1, \ldots, R, j = 1, \ldots, T$. Set $\nu = \max\{\nu(i,j)\}$. According to Section II it can be assumed that

$$\mathbf{h}(n) = [h_{i,j}^{(1)}(n), \ldots, h_{i,j}^{(R,T)}(n), \ldots, h_{i,j}^{(1,1)}(n), \ldots, h_{i,j}^{(R,R,T)}(n)]$$

(3)

for $n \in \mathbb{Z}$, is a zero mean stationary ergodic Gaussian proper complex vector process with process measure $m_{\nu}$ and $\sum_{i=1}^{T} \sum_{j=1}^{R} \sum_{\nu=0}^{\nu} \mathbb{E}||h_{i,j}(\nu)||^2 < \infty$. The signal received at the $i$-th receive antenna at time instant $n$ is:

$$Y_n^{(i)} = \sum_{k=1}^{T} \sum_{j=1}^{R} h_{k,j}^{(i,j)}(n) X_{n-k}^{(j)} + W_n^{(i)} \quad \text{for} \quad i = 1, \ldots, R \quad (4)$$

where $W_n^{(i)}$, $n \in \mathbb{Z}$, is an additive zero mean proper complex Gaussian noise process with variance $\sigma_w$, for each $i = 1, \ldots, R$.

Consider the infinite matrix operator:

$$H = \begin{bmatrix}
\cdots & \cdots & \cdots & \cdots & 0 & H_0(-i) & H_0(0) & 0 & \cdots \\
& \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
& & \cdots & 0 & H_0(i) & H_0(0) & 0 & \cdots \\
& & & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
& & & & \cdots & \cdots & \cdots & \cdots & \cdots \\
& & & & & \cdots & \cdots & \cdots & \cdots \\
& & & & & & \cdots & \cdots & \cdots \\
& & & & & & & \cdots & \cdots \\
& & & & & & & & \cdots \\
\end{bmatrix}$$

(5)

where

$$H_k(n) = \begin{bmatrix}
h_{k,0}^{(1)}(n) & \cdots & \cdots & h_{k,0}^{(1,R,T)}(n) \\
\vdots & \cdots & \cdots & \vdots \\
h_{k,0}^{(R,1)}(n) & \cdots & \cdots & h_{k,0}^{(R,R,T)}(n)
\end{bmatrix}$$

for $k = 0, \ldots, \nu$ are $R \times T$ matrices. The MIMO channel input output relationship for time $-N, \ldots, N$ is

$$Y^{(2N+1)R} = H_{2N+1} X^{(2N+1)T} + W^{(2N+1)R}$$

(6)

where $X^{(2N+1)T}$ is a vector of size $(2N+1)T$, $Y^{(2N+1)R}$ and $W^{(2N+1)R}$ are vectors of size $(2N+1)R$, and $H_{2N+1}$ of size $(2N+1)R \times (2N+1)T$ is obtained by block truncating $H$ around the block $H_0(0)$.

Set $Q = \min\{T, R\}$. Denote the capacity of the block channel (6) with $C_N = \sup I(X^{(2N+1)T}, Y^{(2N+1)R})$ where the $\sup$ is over the input probability distributions that satisfy

$$\frac{1}{(2N+1)Q} \sum_{i=-N}^{N} \sum_{j=1}^{T} E[|X_{i,j}|^2] \leq P$$

(7)

If the limit

$$C = \lim_{N \to \infty} \frac{1}{(2N+1)Q} C_N$$

(8)

exists $m_{\nu}$ a.e. on the channel sample space, we define it as the capacity of the MIMO channel given by (4).

To compute $C_N$ we use channel eigendecomposition to decouple the multiple antenna channels both in space and time. We perform water filling on the powers of the component channels of the vector Gaussian channel. If $R \leq T$ consider $A = HH^*$ ($*$ is adjoint), and if $R > T$ consider $A = H^*H$. $A$ is an essentially self-adjoint Jacobi operator $m_{\nu}$ a.e. Consider a sequence of block matrices $A_{2N+1}$ of sizes $(2N+1)Q \times (2N+1)Q$, obtained by block truncating $A$ around the block with indices $(0, 0)$. We define the empirical distribution function of the eigenvalues $\lambda_{N,i}$ of $A_{2N+1}$ as

$$F_N(\lambda) = \frac{1}{(2N+1)Q} \left(\text{number of } \lambda_{N,i} \leq \lambda\right)$$

(9)

**Theorem** Consider a MIMO channel (4) with impulse response described by a finite power stationary ergodic process $\mathbf{h}(n)$, $n \in \mathbb{Z}$, given by (3). The capacity (8) is given by

$$C = \int_{0}^{\infty} \log(\max(\frac{\theta}{\sigma_w^2}, 1)) dF_c(\lambda)$$

(10)

$$P = \int_{0}^{\infty} \max(\theta - \frac{\sigma_w^2}{\lambda}, 0) dF_c(\lambda)$$

(11)

on a set in the channel sample space of $m_{\nu}$ measure equal to one, on which $F_N$ converge weakly to a fixed function $F_c$.

Proof sketch: In order to show (10)-(11) it suffices to show that a unique LEDF when $N \to \infty$ exists [29]. The vector process $g(n)$ on the diagonal and subdiagonals of $A$, with process measure $m_{\nu}$ is $Q$-stationary ([30]), i.e. stationary with respect to $Q$-shifts. We use all the results from [31] with the process stationary mean ([30]) $\tilde{m}_{\nu}$ substituted for the process measure $m_{\nu}$. Thus, to use Theorem 4.8 in [31] for a $Q$-stationary process we need to show that $\sum_{l=-\infty}^{\infty} \mathbb{E}[\tilde{m}_{\nu}][a_{l,0}] < \infty$. We have

$$\sum_{l=-\infty}^{\infty} \mathbb{E}[\tilde{m}_{\nu}][a_{l,0}] = \frac{1}{Q} \sum_{j=0}^{Q-1} \sum_{l=-[\nu+j+1]Q}^{[\nu+j+1]Q-1} \mathbb{E}[\tilde{m}_{\nu}][a_{l+j,0}] < \infty$$

IV. MIMO-OFDM

Instead of considering time indices from $-\infty$ to $\infty$, now we consider time indices from $0$ to $N - 1$. We use the following $NR \times (N + \nu)T$ block matrix:

$$H_N = \begin{bmatrix}
H_{\nu}(0) & \cdots & H_0(0) & \cdots & 0 \\
\vdots & \ddots & \vdots & \ddots & \vdots \\
0 & \cdots & H_{\nu}(N-1) & \cdots & H_0(N-1)
\end{bmatrix}$$

(12)

MIMO channel input output relationship for time $0, \ldots, N-1$ can be written as

$$Y^{NR} = H_N X^{(N+\nu)T} + W^{NR}$$

where superscript denotes the vector size.

We first assume that on a block of $L$ complex symbols at each transmit antenna the channel is fixed, and channel varies from block to block. We can describe this fixed channel by a matrix $H_L$ with matrix blocks $h_{i,j}(n)$ that do not depend on the time index $n$. Using the multivariate DMT described in [32], $L$ complex symbols are transmitted from each antenna during each channel usage. By using cyclic prefixes of size $\nu$,
instead of $H_L$ it is possible to consider block diagonal matrix 
$$
\tilde{H}_L = \text{diag}( \mathcal{H}(1), \ldots, \mathcal{H}(L) ) 
$$
where

$$
\mathcal{H}(k) = \begin{bmatrix}
H^{(1,1)}(k) & \cdots & H^{(1,T)}(k) \\
\vdots & \ddots & \vdots \\
H^{(R,1)}(k) & \cdots & H^{(R,T)}(k)
\end{bmatrix}
$$
is the $R \times T$ space frequency channel evaluated at DFT index $k$, for $k = 1, \ldots, L$. We further perform SVD of the matrix blocks $\mathcal{H}(k)$ of size $R \times T$ and obtain $Q = \min(R, T)$ singular values $\sigma_k^{(1)}, \ldots, \sigma_k^{(Q)}$ for each $k = 1, \ldots, L$, resulting in a total of $LQ$ singular values. Thus, we obtain a decomposition of the matrix $\tilde{H}_L$ to a diagonal matrix. In other words, we decompose fixed MIMO FIR channel into a vector channel with $LQ$ components, with gains $\sigma_k^{(1)}, \ldots, \sigma_k^{(Q)}, k = 1, \ldots, L$ on each component channel. This is in effect MIMO-OFDM transmission. The details of MIMO-OFDM implementation are given in [32], where terminology used is multivariate DMT.

To analyze a slowly varying channel, we assume that we have a block diagonal matrix $\tilde{H}_{NL}$ of $N$ blocks of the form $\tilde{H}_L$ on the diagonal. This models a block fading channel. Because of channel stationarity and ergodicity, the limit $F_c$ of the empirical distribution functions of the eigenvalues of the matrix $\tilde{H}_{NL}\tilde{H}_{NL}^\dagger$ if $T \geq R$ (or $\tilde{H}_{NL}^\dagger\tilde{H}_{NL}$ if $T \leq R$) of size $NLQ \times NLQ$ when $N \to \infty$ is equal to the distribution function of the eigenvalues of the random matrix $\tilde{H}_L\tilde{H}_L^\dagger (\tilde{H}_L^\dagger\tilde{H}_L)$.

Under the assumption of independent MIMO component channels, i.e. iid entries with equal variances $\sigma^2(k)$ in the matrix $\mathcal{H}(k)$ it is possible to obtain a closed form expression for the p.d.f. of the eigenvalues of $\tilde{H}_L\tilde{H}_L^\dagger (\tilde{H}_L^\dagger\tilde{H}_L)$:

$$
\frac{1}{QL} \sum_{k=1}^L \sigma_k^{-1} \sum_{j=0}^{Q-1} \left( \frac{j}{j+S-Q} \right)^{L} \left( \sigma_j \right)^{2} \left( \sigma_k \right)^{S-Q} e^{-\lambda \sigma_k^{-1}}
$$

where $L_j^S$ is the generalized Laguerre polynomial of order $j$, and $S \equiv \max(R, T)$ (see [3], [11], [12]). The limit $F_c(\lambda)$ is the integral of this p.d.f. In [19] we reported that MIMO channel capacity obtained using this $F_c(\lambda)$ in the capacity formula (10)-(11) is not only the average capacity found in [12], but also the ergodic (Shannon) capacity, under the assumption of block fading MIMO channel. Correlation between the same entries in $\mathcal{H}(i)$ and $\mathcal{H}(k)$ for $i \neq k$ could exist and $\sigma^2(k)$ of the elements of $\mathcal{H}(k)$ could depend on $k$.

Our MIMO channel model results in correlation between the component channels, and above formula cannot be applied. To obtain channel capacity we use simulation and matrix eigenanalysis.

V. OPTIMAL LINEAR CODING FOR VECTOR CHANNELS

Due to Shannon separation theorem source and channel coding can be performed separately in an optimal way [33]. In practice, due to complexity constraints (i.e. finite blocks of data used for coding), it might be beneficial to perform source and channel coding jointly. There are many techniques for joint source channel coding (see [7] and references therein). There are channel optimized source coding techniques as well as source channel coding techniques. The design is done in an iterative way to obtain optimal performance. Optimization criterion is the end to end mean square error under transmitted power constraint. Note that in joint source channel approaches we do not care about probability of error, since what really matters is the mean square error (distortion) in the received signal. Although the optimal mappings from the source space to the channel space (encoder) and from channel space to source space (decoder) that minimize distortion are nonlinear, in this work we consider linear maps, due to two reasons: simplicity, and the existence of a closed form solution. Moreover, in [14] we have proved that if the source and channel are matched in a certain way, the optimal linear coding achieves OPTA (optimal performance theoretically attainable) bound. The OPTA bound is obtained by computing the distortion from the source distortion rate curve, at a rate equal to channel capacity. By using finite blocks of data, we reduce our problem of joint source channel coding of a discrete time continuous amplitude source on a MIMO frequency selective channel to the problem of coding of a vector source on a vector channel. Linear coding uses an analog modulation procedure - multicarrier block quadrature amplitude modulation (BQAM).

We now briefly review the optimal linear coding procedure for coding of vector sources on vector channels described in [17]. Consider, as in [17], a vector channel $y = u + n$ where $u$, $y$, and $n$ are the input, output and noise $s$-dimensional vector respectively (Figure 2). Superscripts $r$ on $u$ and $y$ denote real dimensions. The $k$-dimensional source vector $s$ is obtained by taking $k$ adjacent samples of a discrete time continuous amplitude stationary source $s_n$, $n \in \mathbb{Z}$. The source encoder decorrelates $s$. The correlation matrix of the decorrelated source is the diagonal matrix $\mathcal{M} = \mathbb{E}[xx^\dagger] = \text{diag}[\mu_1, \ldots, \mu_k]$. Assume $1/k \sum_{i=1}^k \mu_i = \sigma^2$. Due to the source stationarity, $\sigma^2$ is equal to the source variance. The mapping from the source space to the channel space (the channel encoder) is the matrix $G$, and the mapping from the space of the received signals to the space of the estimated source signal (the channel decoder) is the matrix $\Gamma$. The vector $n$ denotes additive multichannel Gaussian zero mean noise with correlation matrix $\text{diag}[N_1, N_2, \ldots, N_s]$. Let $\mu_1 > \mu_2 > \cdots > \mu_k$ and $N_1 < N_2 < \cdots < N_s$. The reconstructed signal is

$$
\hat{s} = \Gamma(Gx + n)
$$

The power of the transmitted signal $u$ per source sample is

$$
P = \frac{1}{k} \text{tr}(G\mathcal{M}G^\dagger)
$$

The goal is to minimize the end to end MSE per source sample

$$
D = \frac{1}{k} \text{tr}\mathbb{E}[(x - \hat{x})(x - \hat{x})^\dagger]
$$

under the average power per source sample constraint $P \leq P_c$. 

![Fig. 2. Linear coding for a vector channel](image-url)
In [17] it was shown that the only nonzero elements of the encoder matrix $G$ and the decoder matrix $\Gamma$ are those with equal indices

$$g_i = \left(\frac{\mu_i N_i}{\theta} - N_i\right)^{1/2}/\sqrt{\mu_i}, \quad \text{for } 1 \leq i \leq l$$

(16)

$$\gamma_i = \frac{\mu_i g_i}{\mu_i \theta^2 + N_i} \quad \text{for } 1 \leq i \leq l$$

(17)

where $l \leq \min(k, s)$ is the maximum integer which satisfies $\frac{N_k}{N_{\theta}} > \theta$. The average transmitted signal power and distortion (MSE) per source are ([17])

$$P = \frac{1}{k} \sum_{i=1}^{l} \left(\frac{\mu_i N_i}{\sqrt{\theta}} - N_i\right)$$

(18)

$$D = \frac{1}{k} \left(\sum_{i=1}^{l} \mu_i + \sqrt{\theta} \sum_{i=1}^{l} \sqrt{\mu_i N_i}\right)$$

(19)

In our vector representation of MIMO channel, noise variances are obtained by repeating twice the gain on each complex channel dimensions, where in the resulting $LQ$ complex channel dimensions. In fact there are twice as many real channel dimensions, where in the real vector channel representation, gains on real subchannels are obtained by repeating twice the gain on each complex subchannel. Thus, $s = 2QL$.

VI. OPTIMAL LINEAR CODING FOR MOBILE CHANNELS

Due to channel stationarity and ergodicity it is possible to use a finite state channel model for coding of discrete time continuous amplitude source ([18], [7]). $K$ representative channel states can be obtained by classifying a long realization of the sampled vector CIR $[h_0^{(1,1)}(nL), \ldots, h_0^{(R,T)}(nL), \ldots, h_{\nu}^{(1,1)}(nL), \ldots, h_{\nu}^{(R,T)}(nL)], n = 0, \ldots, N - 1$, where $L$ is the channel frame size, and $N$ is a large number.

Assuming perfect CSI, the number of states is $K = N$. For each channel state $i = 1, \ldots, N$, we compute RT DFT’s of size $L$, to obtain $H_i^{(1,1)}(k), \ldots, H_i^{(R,T)}(k)$, for $k = 1, \ldots, L$ and $i = 1, \ldots, N$. For each channel state and each $k$ we perform SVD of the blocks $H_i(k)$, as described in Section IV, to obtain singular values and a vector channel of dimension $LQ$. The design of the optimal linear coder for the obtained vector channel proceeds as described in Section V. In [18], [7], [19] we described two approaches for linear coding on stationary, ergodic FIR channels. In Approach 1, to make use of the channel ergodicity, we encode on $N$ channel blocks of size $L$ corresponding to $N$ channel states. We concatenate the singular values evaluated for all $N$ blocks into one vector of size $NLQ$ $(2NLQ$ real dimensions) and order them in decreasing order. We proceed to design a single encoder/decoder pair as described in Section V. The transmitter has CSI in order to transmit the portion of source vector, corresponding to the current channel state. We evaluated Approach 1 performance limit in terms of source and channel LEDF (limiting eigenvalue distribution functions) in [14], [19]

$$P = \int_{\lambda F_{c}^{-1}F_{c}(\lambda) \geq \theta} \frac{F_{c}^{-1}F_{c}(\lambda)}{\lambda} dF_{c}(\lambda)$$

(20)

$$D = 1 - \int_{\lambda F_{c}^{-1}F_{c}(\lambda) \geq \theta} \left(\frac{F_{c}^{-1}F_{c}(\lambda)\theta}{\lambda}\right) dF_{c}(\lambda)$$

(21)

In Approach 2 we design a separate encoder/decoder pair for each channel state, on channel blocks of size $L$ (with $2LQ$ real dimensions of the decomposed real channel), although the design is joint. Joint design is achieved by taking the same value of the design parameter $\theta$ for the design in each state. Performance limit is evaluated in [19] in terms of the distribution functions of the order statistics of the squared singular values on blocks of size $L$.

In a practical approach, described in [18], [7], which is a hybrid between the two approaches, we order the channel state powers (sums of moduli of DFT coefficients squared) in a decreasing order. We then select certain number of states with total probability $p$, with lowest power (“bad” states), and choose not to transmit when the channel is in one of these states. In order not to lose information, we transmit twice as many source samples during the states with probability $p$ and highest powers (matching “good” states). For that purpose, only half of the KLT coefficients of each source block are transmitted during “good” states, which enables transmission of two source blocks on a single channel block. In other words, when the channel is in a “good” state, we use a source compression of ratio two. During the remaining “fair” states, we transmit a single source block on a single channel block.

VII. SIMULATION RESULTS

As an example of MIMO system we use two transmit and two receive antennas, $R = T = 2$. We assume a carrier frequency $f_0 = 1$ GHz ($\lambda = 0.3m$). We chose $r_0 = 5$ km and symbol rate (on each transmit antenna) of $f_s = 400$ ksymbols/s resulting in the value of parameter $a = 0.15$ and time dispersion of five FIR coefficients. A slowly varying channel, which is suitable for channel estimation and multicarrier transmission, is obtained for the assumed symbol rate $f_s = 400$ ksymbols/s and channel blocks (channel frame) on each antenna of $L = 16$, even at vehicle speeds of $|v| = 100$ km/h (normalized Doppler is $f_d/f_s = 0.00023$).

In our simulations we set $d_T$ and $d_R$ equal to $10\lambda$ or $0.5\lambda$ and $\theta_T$ and $\theta_R$ equal to 0 or $\pi/2$. In all the simulations we set a vehicle speed of $|v| = 100$ km/h and $\theta = 0$. Simulation results were obtained by statistically averaging over large number of scatterer configurations. The centers of scatterer regions were placed randomly in an ellipse, with uniform probability distribution. We choose scatterer regions of size $50m \times 50m$ or $250m \times 250m$ with uniform distribution of elementary scatterers within those regions.
In Figure 3 we show capacity results for the four MIMO configurations and SISO. Power and capacity expressions are normalized by the number of antennas, and therefore, it looks that SISO results in higher capacity than some cases of MIMO. Antenna correlation results in small performance gains in low SNR region, and performance losses in high SNR region. This is especially remarkable in the \( \theta_R = \theta_T = 0, d_R = d_T = 0.5\lambda \) case. Thus, our conclusions for the impact of correlation on the capacity of MIMO frequency selective channel agree with the observations in [10], [28] for the impact of correlation on the capacity of MIMO frequency nonselective channel.

We consider a transmission of a zero mean autoregressive Gaussian source \( s_n = \rho s_{n-1} + z_n \), with variance \( \sigma^2 = 1 \), where \( z_n \) is Gaussian noise with variance \( \sigma_z^2 = \sigma^2(1 - \rho^2) \).

Channel frame of \( L = 16 \) results in a vector channel of \( s = 2QL = 64 \) real dimensions \( (Q = 2 \) for \( 2 \times 2 \) MIMO). Setting \( k = s \) implies that the size of the source blocks is 64.

The simulation results of MIMO-OFDM source transmission \( (\rho = 0.99) \) for scatterer regions of \( 50m \times 50m \), are shown in Figure 4 and Figure 5. Results are represented as signal to distortion ratio per source dimension \( SDR = \sigma^2/D \), in terms of channel SNR per source dimension \( SNR = P/\sigma_w^2 \). We show plots of Approach 1 and Approach 2 coding bounds as well as the OPTA bound (all obtained when \( N \to \infty \)). The Approach 2 bound is lower than the Approach 1 bound, which is the ultimate performance limit of linear coding. Due to large antenna correlation, we compute OPTA and Approach 1 coding bounds by eigendecomposition of large source and channel matrices. In fact, due to slow channel variation, we use eigendecomposition in each channel state and concatenate all the computed eigenvalues in a single vector of length \( 2NQL \).

Approach 1 coding bound is lower than the OPTA bound since linear coding is suboptimal for arbitrary variances of the components of the decorrelated vector source and arbitrary subchannel gains on the equivalent vector channel.

In [9] we presented results for \( \rho = 0.9 \) as well. This resulted in a lower performance compared to \( \rho = 0.99 \). All three curves were closer than for \( \rho = 0.99 \), which showed that source eigenvalues were better matched to channel gains (on the large channel block of size \( NL \) for Approach 1, or on channel blocks of size \( L \) for Approach 2) in this case.

In Figure 6 we show Approach 1 and OPTA bounds for clusters of size \( 50m \times 50m \) and \( \rho = 0.99 \) for the four considered MIMO configurations.

In-line antenna configuration \( (\theta_T = \theta_R = 0) \), results in higher correlation than broad-side antenna configuration \( (\theta_T = \theta_R = \pi/2) \), as also observed in [24], [10] for the case of frequency nonselective fading. Highest correlation is obtained when in-line antennas with spacing between the antenna elements of \( 0.5\lambda \) are chosen at both the transmitter and receiver. From Figure 6 (a) we conclude that very high MIMO correlation somewhat increased linear coding performance in the low SNR region, and considerably decreased performance in the high SNR region. The same conclusion holds for OPTA bounds shown in Figure 6 (b). In Figure 6 we also show Approach 1 and OPTA bound for SISO system. For very high antenna correlation \( (\theta_T = \theta_R = 0 \) and \( d_T = d_R = 0.5\lambda \)) performance of SISO seems superior to MIMO. But, we should not forget that we transmit twice as much information with \( 2 \times 2 \) MIMO compared to SISO.

Increasing the size of the scatterer clusters to \( 250m \times 250m \) resulted in decreased antenna correlation (see [9]).
VIII. CONCLUSION

We presented several issues related to MIMO systems in frequency selective fading channels: channel model, Shannon capacity, and linear coding. Starting from a physical mobile channel model, we developed a MIMO FIR channel model with stationary, ergodic coefficients. Then, we derived capacity of the obtained MIMO FIR channel. Assuming a slowly varying channel and MIMO-OFDM transmission, we proposed a linear coding algorithm for discrete time continuous amplitude sources.

REFERENCES


[13] V. Kafedziski, “Capacity of time varying single and multiple antenna channels with ISI using random ergodic operators,” in Proceedings of


