RESEARCH   **Open Access**

Reduced feedback for selective fading MIMO broadcast channels

Lina Mroueh[1*], Stéphanie Rouquette-Léveil[2] and Jean-Claude Belfiore[3]

Abstract

In this article, we analyze the selective multiple-input multiple-output broadcast channel, where links are assumed to be selective in both time and frequency. The assumption of full channel knowledge at the transmitter side requires a large amount of feedback, and it is therefore not practical to be implemented in real systems. A more feasible solution with finite rate feedback originally proposed by Jindal in *IEEE Trans. Inf. Theory* is applied here to the selective fading case, where the minimal number of feedback bits required to achieve the full multiplexing is derived. We show that the correlation between time frequency channels can be used in order to minimize the number of feedback bits to the transmitter side while conserving the maximal multiplexing gain. Finally, the practical implementation of a time-frequency channel quantization scheme is addressed, and a low-complexity scheme that also achieves the multiplexing gain is proposed.

Keywords: time-frequency selective channel, underspread channel, broadcast channel, random vector quantization, Grassmann manifold, multiplexing gain

1. Introduction and motivation

The challenge of the next generation of wireless communication is to offer at the receiver side a high data rate with a high quality of service. The multiple-input multiple-output (MIMO) transmission and the multiuser communication have been recently introduced in almost all new standards. These two techniques of transmission have been extensively studied in the literature over the last few years aiming to boost the quality of service of wireless systems close to the one of wireline systems.

In this article, we consider the broadcast channel (BC), where a common source transmits data simultaneously to different receivers that do not cooperate. We assume that communication occurs on channels that exhibits memory in both time and frequency. Our objective here is to propose a transmission strategy for the high data rate communications when the channel is known partially at the transmitter side.

When full CSIT is assumed at the transmitter side, the capacity region of the BC has been characterized in [1]. It has been shown that the Dirty Paper coding technique achieves the maximal capacity region. Despite of its

optimality, this technique is not feasible to be implemented in practical system as it brings high complexity at the transmitter side. Many more practical downlink transmission techniques including linear precoding schemes (e.g., channel inversion [2] and block diagonalization (BD) in [3]) and nonlinear precoding schemes (e.g., vector perturbation technique [4]) have been proposed. Although the vector perturbation technique improves the error performance compared to linear precoding schemes, this comes at the expense of an increased complexity caused by the use of a sphere encoder at the transmitter side. Linear precoding schemes considered in this paper are less complex to operate than other precoding schemes and are shown to achieve the full multiplexing gain at the high SNR regime [5].

The full CSIT assumption is not generally of a practical interest as it requires a large amount of feedback. The quantization schemes of flat fading or frequency selective MIMO BC has been extensively addressed in literature [6]-[13]. A limited feedback solution with finite rate feedback for the flat fading channel has been studied by Jindal for the case of single antenna users [6] and later extended to the MIMO case in [7]. More realistic feedback schemes with noisy feedback scheme and

* Correspondence: lina.mroueh@isep.fr
[1]Institut Supérieur d'Electronique de Paris, 75006 Paris, France
Full list of author information is available at the end of the article

delay were considered in [8]. For the frequency selective fading MIMO BC, most of state-of-the-art techniques proposed in literature, e.g., [9]-[12] deal with the quantization of the frequency response with a focus on the quantization codebook design. A complete tutorial on these limited feedback strategies with their applications in standards can be found in [13]. The majority of these works pointed out that using an adequate number of feedback bits that scales as SNR, the full multiplexing gain can be also achieved using limited feedback.

While most of the above results address the case where the channels between source and destination are assumed to be flat fading or frequency selective, real communications occur on channel that exhibits memory in time and frequency [14]. The time-frequency selective channel gives an accurate model of the wireless channel, and especially for the case of applications that exhibit duration and bandwidth that exceed the coherence time and bandwidth of the channel. A complete description of the considered channel model can be found in [14,15].

A. Contributions

In this paper, we analyze the selective MIMO BC, where links are selective in both time and frequency. Based on the fact that the time-frequency selective channel model can be well approximated by the parallel correlated (time-frequency) channels as in [15], we show how the correlation between these time-frequency channels can be used in a selective MIMO BC to reduce the number of feedback bits while conserving the full multiplexing gain. The two main contributions of this paper can be summarized as following: (i) We show that the time-frequency selective channel can be characterized by a finite number of Gaussian invariants parameters, and we propose strategies with a limited number of feedback bits to quantize these invariants parameters. (ii) A practical approach to achieve the full multiplexing gain with a low computational complexity scheme is proposed to quantize the time-frequency selective channel.

B. Outline

The rest of the article is organized as following. In Section 2, we present the channel and the signal model, and we propose a general representation of the selective channel. Then, using this channel representation, we show in Section 3 how the correlation between these time-frequency channels can be used in a selective MIMO BC to reduce the number of feedback bits while conserving the full multiplexing gain. Moreover, a practical feedback scheme with a low computational complexity is also addressed in this section. The optimality of the reduced feedback channel quantization is also illustrated using numerical results. Finally, Section 4 concludes the article.

C. Notation

The notation used in this paper is as follows. Boldface lower case letters $\mathbf{v}$ denote vectors, boldface capital letters $\mathbf{M}$ denote matrices. $\mathbf{M}^\dagger$ denotes conjugate transposition. $\mathbf{I}_N$ stands for the $N \times N$ identity matrix. $\mathcal{CN}$ represents the complex Gaussian random variable. $\mathbb{E}_X$ is the mathematical expectation w.r.t. to the random variable X. The pulse distribution is denoted by $\delta_n = 1$ if $n = 0$ and 0 otherwise. $\lfloor x \rfloor$ denotes the floor of x and $\langle \mathbf{x}, \mathbf{y} \rangle$ is the scalar product between two vectors $\mathbf{x}$ and $\mathbf{y}$.

2. System and channel model

In this article, we consider a multiple-antenna BC where a source S with n_t transmit antennas wants to communicate simultaneously with K destinations D_i having n_r receive antennas each, with $n_t \geq K n_r$. We assume that all communications occur on time-frequency selective fading channels. In the following, we start by briefly recalling from [15] the approximate decomposition of time-frequency selective channels into statistical correlated parallel channels for the point-to-point case. Then, the corresponding input-output relation at each destination for the BC is provided.

A. Time-frequency selective SISO channel model

In wireless communication, the multipath approximation is often used to describe the wireless channel model. In this case, the received signal is the sum of all multipath components and the line of sight. Each path induces a variation of the signal strength due to the Doppler spread v caused by the movement of the transmitters, receivers, and scatterers and a delay shift τ at the receiver side. In general, the number of paths is very high, which makes logical to model the multipath effect by a continuous linear time-varying (LTV) system. The LTV system is described by a linear operator $\mathbb{H}$ that maps an input signal into an output signal $r(t)$, related by the following noise-free relationship,

$$r(t) = (\mathbb{H}s)(t) = \int_\tau h(t, \tau) s(t - \tau) d\tau, \tag{1}$$

where $s(t)$ is the transmitted signal and $h(t, \tau)$ is the time-varying impulse response.

The LTV system is also characterized by two other functions. The delay-Doppler spreading function $S_{\mathbb{H}}(v, \tau)$ defined as Fourier transform $(t \to v)$ of $h(t, \tau)$ and the time-varying transfer function $L_{\mathbb{H}}(t, f)$ defined as the Fourier transform $(\tau \to f)$ of $h(t, \tau)$.

B. WSSUS assumption and statistical channel description

In digital communications, the linear operator $\mathbb{H}$ is random and LTV channel models are studied under the wide-sense stationary and uncorrelated scattering (WSSUS) assumption. This property consists in assuming

that the random channel $\mathbb{H}$ is wide-sense stationary in time t and uncorrelated in scattering (delay) τ, which means that

$$\mathbb{E}\left[h(t, \tau)h(t', \tau')\right] = R_h(t - t', \tau)\delta(\tau - \tau').$$

The WSSUS property implies that the time-varying transfer function $L_{\mathbb{H}}(t, f)$ is wide-sense stationary in both time and frequency, and the spreading function $S_{\mathbb{H}}(v, \tau)$ is uncorrelated in delay τ and in Doppler v, i.e.,

$$\mathbb{E}\left[L_{\mathbb{H}}(t, f) L_{\mathbb{H}}^*(t',f')\right] = R_{\mathbb{H}}(t - t', f - f'),$$
$$\mathbb{E}\left[S_{\mathbb{H}}(v, \tau)S_{\mathbb{H}}^*(v', \tau')\right] = C_{\mathbb{H}}(v, \tau)\delta(\tau - \tau')\delta(v - v').$$

The scattering function $C_{\mathbb{H}}(v, \tau)$ is the 2D Fourier transform of the time-frequency correlation function $R_{\mathbb{H}}(\Delta t, \Delta f)$ such that

$$R_{\mathbb{H}}(\Delta t, \Delta f) = \int_{\tau}\int_{v} C_{\mathbb{H}}(v, \tau)e^{j2\pi v\Delta t}e^{-j2\pi\tau\Delta f}d\tau\, dv. \quad (2)$$

C. Underspread LTV operator

As a consequence of the limited velocity of transmitter, receiver, and scatterers in the propagation environment, the maximum Doppler shift is limited to v_0. We also assume that the maximum delay is bounded by 0 and $+\tau_0$. The scattering function $C_{\mathbb{H}}(v, \tau)$ is therefore supported on a rectangle of area $\Delta_{\mathbb{H}} = \tau_0 v_0$ known as *channel spread*, such that

$$C_{\mathbb{H}}(v, \tau) = 0 \quad \text{for} \quad (v, \tau) \notin [0, +v_0] \times [0, +\tau_0]. \quad (3)$$

The underspread assumption is of a practical interest as the time taken by channel to change significantly $(1/v_0)$ is usually much longer than the delay spread τ_0, i.e.,

$$\tau_0 v_0 \ll 1. \quad (4)$$

Channels satisfying these characteristics are called *underspread selective fading channels* and are shown by Kozek [16] to satisfy the two following properties.

- The underspread LTV operator $\mathbb{H}$ admits structured sets of orthonormal eigenfunctions $\{g_{m, l}(t)\}$ that are independent of the channel operator, well localized in time and frequency and known as *Weyl-Heisenberg* (WH) set. This set is obtained by translating in time and modulating in frequency a prototype $g(t)$. In the following, this set is denoted as

$$(g(t), T, F) = \left\{g_{m,n}(t) = g(t - mT)e^{j2\pi nFt}\right\}. \quad (5)$$

where $m, n \in \mathbb{Z}$, T and F are the grid parameter of WH set. The triple $g(t), T, F$ are chosen such that $g(t)$ has unit energy and that $g_{m, n}(t)$ form an orthonormal base, i.e.,

$$\langle g_{m,n}(t), g_{k,p}(t)\rangle = \int_{t} g_{m,n}(t)g_{k,p}(t)dt = \delta_{m,k}\delta_{n,p}.$$

Finally, the grid parameters T and F should satisfy $TF > 1$ to guarantee that $g_{m, n}(t)$ form an orthonormal basis and are well localized in time and frequency[a] (more details about the choice of grid parameters can be found in [15] and references therein). Heuristically, the optimal choice of TF that minimizes the inter-symbol interference (ISI) and inter-carrier interference (ICI) and maximizes the number of degrees of freedom is $TF \approx 1.25$.

- For grid parameters chosen such that $T \leq \frac{1}{v_0}$ and $F \leq \frac{1}{\tau_0}$ and hence

$$\Delta_{\mathbb{H}} \leq \frac{1}{TF} \leq 1, \quad (6)$$

it has been shown in [16], [15] that the eigenvalues of $\mathbb{H}$ can be well approximated by,

$$(\mathbb{H}g_{m,l})(t) \approx L_{\mathbb{H}}(mT, lF)g_{m,l}(t),$$

where $L_{\mathbb{H}}(t, f) = \mathcal{F}_{\tau \to f}\{h(t, \tau)\}$ is the time-varying transfer function.

D. Signaling scheme: equivalent parallel model

In the following, we base our results on the strategy used by Dirusi et al. [15] to derive the discrete time-frequency input output relation. This strategy consists of transmitting and receiving on the orthonormal channel eigenfunctions $g_{m, l}(t)$. The transmitted signal can be therefore written using the channel WH set as,

$$s(t) = \sum_{m=0}^{M-1}\sum_{l=0}^{N_c-1} x[m, l]g_{m,l}(t), \quad (7)$$

where $D = MT$ is the approximate time duration of $s(t)$ and $W = N_cF$ is its approximate bandwidth.

The projection of the noisy received signal $y(t) = \mathbb{H}s(t) + z(t)$ onto the channel WH set is given by,

$$y[m, l] = \langle y(t), g_{m,l}(t)\rangle$$

which imply that,

$$y[m, l] = \langle \mathbb{H}x(t), g_{m,l}(t)\rangle + \langle z(t), g_{m,l}(t)\rangle,$$
$$\approx L_{\mathbb{H}}(mT, lF)x[m, l] + z[m, l]. \quad (8)$$

Note that due to the orthonormal WH set, $z[m, l]$ are i.i.d for all $(m, l) \in \{0 \ldots M - 1\} \times \{0 \ldots N_c - 1\}$, such that $z[m, l] \in \mathcal{CN}(0, 1)$ and $\mathbb{E}[z[m, l]z[m', l']] = \delta_{m,m'}\delta_{l,l'}$. In

the rest of the paper, we let n denote the time-frequency slot (m, l), with $n = 0 \ldots N - 1$ and $N = MN_c$ is the total number of time-frequency slots. We finally denote by **h** the $N \times 1$ vector containing $L_{\mathbb{H}}(mT, lF)$ parameters with $(m, n) \in \{0 \ldots M - 1\} \times \{0 \ldots N_c - 1\}$.

It can be easily seen from (8) that using the WH set, the time-frequency parallel channel is converted into N parallel statistically dependent time-frequency channels where the correlation matrix between the scalar sub-channels is denoted by $\mathbf{R}_{\mathbb{H}} = \mathbb{E}\,[\mathbf{h}\mathbf{h}^{\dagger}]$. As shown in Section 2-B, the time-varying transfer function $L_{\mathbb{H}}(t, f)$ is wide-sense stationary in time and in frequency, and the covariance matrix $\mathbf{R}_{\mathbb{H}}$ is a two-level Hermitian Toeplitz matrix. Using the same notations as in [15], the covariance matrix $\mathbf{R}_{\mathbb{H}}$ can be written as,

$$\mathbf{R}_{\mathbb{H}} = \begin{bmatrix} \mathbf{R}_{\mathbb{H}}[0] & \mathbf{R}_{\mathbb{H}}^{\dagger}[1] & \ldots & \mathbf{R}_{\mathbb{H}}^{\dagger}[M-1] \\ \mathbf{R}_{\mathbb{H}}[1] & \mathbf{R}_{\mathbb{H}}[0] & \ldots & \mathbf{R}_{\mathbb{H}}^{\dagger}[M-2] \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{R}_{\mathbb{H}}[M-1] & \mathbf{R}_{\mathbb{H}}[M-2] & \ldots & \mathbf{R}_{\mathbb{H}}[0] \end{bmatrix},$$

where $\mathbf{R}_{\mathbb{H}}[k]$ is a Toeplitz matrix with coefficients $R_{\mathbb{H}}[k, n]$ given in (2) and defined in (9).

$$\mathbf{R}_{\mathbb{H}}[k] = \begin{bmatrix} R_{\mathbb{H}}[k, 0] & R_{\mathbb{H}}^*[k, 1] & \ldots & R_{\mathbb{H}}^*[k, N_c - 1] \\ R_{\mathbb{H}}[k, 1] & R_{\mathbb{H}}[k, 0] & \ldots & R_{\mathbb{H}}^*[k, N_c - 2] \\ \vdots & \vdots & \ddots & \vdots \\ R_{\mathbb{H}}[k, N_c - 1] & R_{\mathbb{H}}[k, N_c - 2] & \ldots & R_{\mathbb{H}}[k, 0] \end{bmatrix}. \quad (9)$$

E. Multiuser BC model

In the rest of this article, we consider a MIMO BC depicted in Figure 1 where a source S with n_t transmit antennas wants to communicate simultaneously with K destinations D_i having n_r receive antennas each, with $n_t \geq Kn_r$. This condition can be of a practical interest if we consider that the K users are selected randomly among a large number of users and have an aggregate number of receive antennas that does not exceed the number of transmit antennas.

The fading process considered in this paper are characterized by non-disjoint scattering function $C_{\mathbb{H}}(\nu, \tau)$ such that

$$C_{\mathbb{H}}(\nu, \tau) = 0 \quad \text{for} \quad (\nu, \tau) \notin [0, +\nu_0] \times [0, +\tau_0]. \quad (10)$$

By choosing the sampling period T, and sampling frequency F such that $T \leq 1/\nu_0$ and $F \leq 1/\tau_0$, channels can be diagonalized using the same WH sets. At each destination k, the received signal $\mathbf{y}_k(n)$ is given by

$$\mathbf{y}^{[k]}(n) = \mathbf{H}^{[k]}(n)\mathbf{x}(n) + \mathbf{n}^{[k]}(n), \quad (11)$$

where $\mathbf{H}^{[k]}(n) \in \mathbb{C}^{n_r \times n_t}$ is the channel matrix at the time-frequency slot n, and $n = 0 \ldots N - 1$ with N being

the total number of time-frequency slots. The vector $\mathbf{x}(n) \in \mathbb{C}^{n_t \times 1}$ is the transmitted signal, and $\mathbf{n}^{[1]}(n), \ldots, \mathbf{n}^{[K]}(n)$ are independent complex Gaussian noise terms with unit variance. The transmitter is subject to an average power constraint P, such that

$$\mathrm{Tr}\,\{\,\mathbb{E}[\mathbf{x}(n)\mathbf{x}(n)^{\dagger}]\} \leq P. \quad (12)$$

We assume that channels are spatially uncorrelated, that for a given time-frequency slot n, $\mathbf{H}^{[k]}(n)$ has i.i.d $\mathcal{CN}(0, 1)$ entries. The channels corresponding to different destinations are assumed to be statistically independent. However, channels are correlated across n for a given destination k, according to

$$[\mathbf{H}^{[k]}(0) \ldots \mathbf{H}^{[k]}(N-1)] = \mathbf{H}_w^{[k]}(\mathbf{R}_{\mathbb{H}}^{1/2} \otimes \mathbf{I}_{n_t}), \quad \forall k \quad (13)$$

where $\mathbf{R}_{\mathbb{H}}$ is the $N \times N$ correlation between the scalar sub-channels with rank equal to ρ, $\mathbf{H}_w^{[k]}$ is an $n_r \times Nn_t$ matrix with i.i.d. $\mathcal{CN}(0, 1)$ entries. For simplicity of notations, we assume that all scalar subchannels $h_{i,j}^{[k]}$ have the same correlation function.

In the following, we assume that the transmitter does not know the instantaneous value of the channel but knows the probabilistic channels' law[b] including the knowledge of $\mathbf{R}_{\mathbb{H}}$.

F. Impact of the correlation on the channel model

In this section, we propose a general representation of the time-frequency selective channel. We show that the MIMO channel between the source and each destination at each time-frequency slot can be written as given in Lemma 1.

Lemma 1 (Time-frequency MIMO channel matrix): The user k channel matrix $\mathbf{H}^{[k]}(n) \in \mathbb{C}^{n_r \times n_t}$ at a time-frequency slot n can be written as

$$\mathbf{H}^{[k]}(n) = \mathbf{H}_{\omega}^{[k]}\mathbf{\Gamma}(n) \quad (14)$$
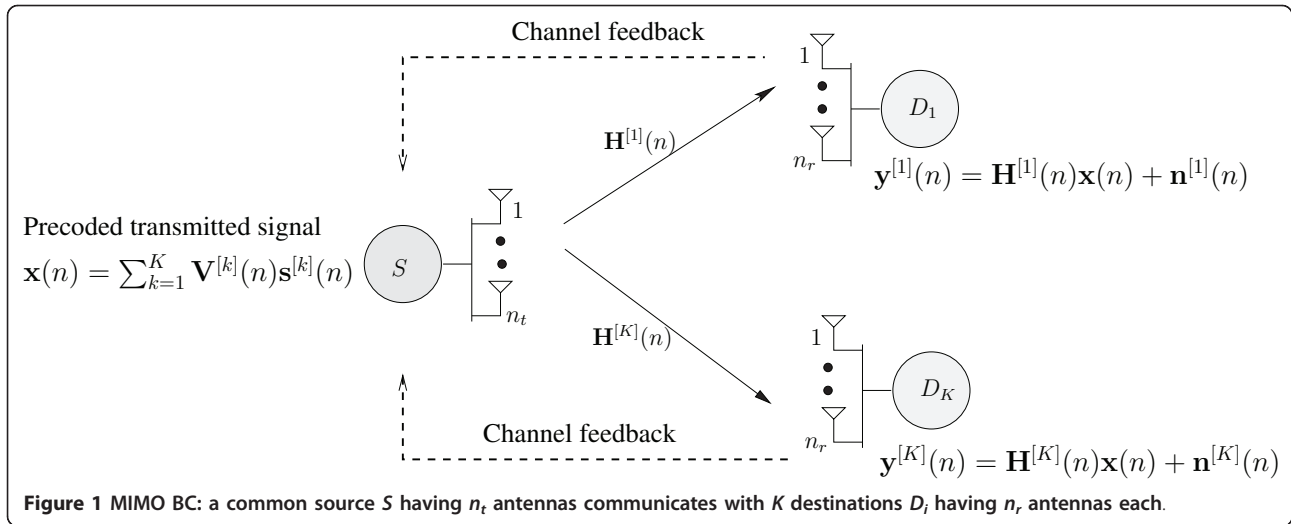
where $\mathbf{H}_{\omega}^{[k]} \in \mathbb{C}^{n_r \times \rho n_t}$ is a Gaussian matrix with i.i.d $\mathcal{CN}(0, 1)$ entries such that

$$\mathbf{H}_{\omega}^{[k]} = \left[\mathbf{H}_{\omega,0}^{[k]}\, \mathbf{H}_{\omega,1}^{[k]} \ldots \mathbf{H}_{\omega,\rho-1}^{[k]}\right],$$

and $\mathbf{\Gamma}(n)$ is a $\rho n_t \times n_t$ deterministic matrix that depends only on the channel statistics, such that,

$$\mathbf{\Gamma}(n) = \begin{bmatrix} \sigma_0 w_{n,1} \\ \vdots \\ \sigma_{\rho-1} w_{n,\rho} \end{bmatrix} \otimes \mathbf{I}_{n_t}. \quad (15)$$

where σ_{p-1}^2 are the eigenvalues of the covariance matrix $\mathbf{R}_{\mathbb{H}}$, and $w_{i,\,p}$ is the ith entry of the eigenvector $\mathbf{w}_p$ of $\mathbf{R}_{\mathbb{H}}$ corresponding to σ_{p-1}^2.

Figure 1 MIMO BC: a common source *S* having n_t antennas communicates with *K* destinations D_i having n_r antennas each.

Proof: Please refer to Appendix A for the proof of this lemma. ∎

In the following, we set

$$\sigma_t^2 = \max_{n=0...N-1} \sum_{i=0}^{\rho-1} \sigma_i^2 |w_{n,i}|^2.$$

G. Physical interpretation of Lemma 1

The channel model in Lemma 1 gives a general representation of any selective fading channel and models the cases where the channel is selective either in time, frequency or in time and frequency.

1) Time selective channel (or block fading channel

The time selective channel models the case when the channel remains constant during a block *n* and changes in a statistical independent manner across block, such that

$$\mathbf{H}^{[k]}(n) = \mathbf{H}_\omega^{[k]}(n), \quad n = 0 ... N - 1.$$

In this case, the correlation matrix $\mathbf{R}_\mathbb{H} = \mathbf{I}_N$ and consequently $\Gamma(n) = \mathbf{e}_n \otimes \mathbf{I}_{n_t}$, where the elements $e_{n,j}$ of the $N \times 1$ vector $\mathbf{e}_n$ are such that $e_{n,j} = \delta_{n-j}$.

2) Frequency selective channel

For the frequency selective channel, it is well known that the channel can be decomposed into *N* parallel statistically dependent sub-channels when using a cyclic prefix OFDM system. In this case, *n* stands for the frequency slot, *N* the total number of subcarriers, and the correlation matrix $\mathbf{R}_\mathbb{H}$ is a circulant matrix with rank ρ equal to the number of channels taps *L*. The eigenvectors of $\mathbf{R}_\mathbb{H}$ are the columns of an $N \times N$ FFT matrix. Consequently, $\Gamma(n)$ in Lemma 1 is such that

$$\Gamma(n) = \frac{1}{\sqrt{N}} \begin{bmatrix} \sigma_0 \, \omega^0 \\ \vdots \\ \sigma_{L-1} \, \omega^{n(L-1)} \end{bmatrix} \otimes \mathbf{I}_{n_t},$$

where $\omega = e^{j\frac{2\pi}{N}}$ and σ^2 is the variance of the channel taps.

3) Time frequency selective channel

As stated before, due to the delay time-varying channel, the considered LTV channel induces ISI and ICI at each receiver side. As mentioned in Section 2-D, by projecting the transmitted signal and the received signal on the channel eigenfunctions, the ISI and ICI interferences are canceled. The LTV channel is then decomposed into parallel time-frequency channel. The time-frequency channels change at each time-frequency slot. However, for an LTV channel that it is characterized by a scattering function that is compactly supported in a rectangle as in (3), it is well known from [17] that this variation depends only on a finite number of parameters that are *invariant* during all the duration of the transmission.

It can be deduced from Lemma 1 that when statistical channel knowledge is available at the receiver side, the knowledge of the time-varying channel $L_{i,j,\mathbb{H}}^{[k]}(mT, lF)$ at any time frequency slot $n = (m, l)$ depends only on the knowledge of the Gaussian vector $\mathbf{h}_{\omega, i, j} = [h_{\omega, i, j}[0] ... h_{\omega, i, j}[\rho - 1]]$, where

$$h_{i,j}^{[k]}(n) = L_{i,j,\mathbb{H}}^{[k]}(mT, lF) = \sum_{s=0}^{\rho-1} \sigma_s w_{n,s} h_{\omega,i,j}[s]. \quad (16)$$

Another different way to identify the LTV channel has also been established by Kailath [17] where a channel reconstruction strategy based on the 2D Shannon sampling theorem was proposed. A complete description on

the LTV channel identification can be found in Chapters 1 and 2 of [14]. It should be emphasized here that the vector containing the finite number of Kailath parameters' required to identify the channel is not Gaussian with i.i.d. entries and cannot be easily quantized using classical quantization technique. The statistical distribution of these parameters is not easy to be derived. The Kailath channel identification strategy states also that the LTV channel can be identified by a finite number of parameters, which is equal to $(\lfloor \nu_0 D \rfloor + 1)(\lfloor \tau_0 W \rfloor + 1)$. This number of free parameters can give a rough estimate of the rank ρ of the covariance matrix $\mathbf{R}_{\mathbb{H}}$, which is nothing but the number of independent Gaussian parameters required to identify the channel. This implies that ρ is approximatively equal to

$$\rho \approx (\lfloor \nu_0 D \rfloor + 1)(\lfloor \tau_0 W \rfloor + 1). \tag{17}$$

Remark 1: From the choice of the grid parameters ($\nu_0 T \leq 1$, $\tau_0 F \leq 1$ and $TF = 1.25$), it can be easily shown that in this case, that the ratio $\frac{\rho}{N}$ is bounded by

$$\frac{\rho}{N} \leq TF\Delta_{\mathbb{H}} + \frac{\nu_0 T}{N_c} + \frac{\tau_0 F}{M} + \frac{1}{MN_c}.$$

For wireless channel, the channel spread is generally $\Delta_{\mathbb{H}} \in [10^{-7} \, 10^{-3}]$. Moreover, the total number of time-frequency slots is generally very large. This implies that $\frac{\rho}{N}$ scales as

$$\frac{\rho}{N} \sim \max(\Delta_{\mathbb{H}}, \frac{1}{M}, \frac{1}{N_c})$$

and, therefore, the rank of the covariance matrix ρ is much lower than the total number of time-frequency slots.

Remark 2: As stated in Section 2-D, the covariance matrix $\mathbf{R}_{\mathbb{H}}$ is a two-level Hermitian Toeplitz matrix. The eigenvectors of such matrix are not generally well structured as it is the case for circulant matrix unless its dimensions are sufficiently large, i.e., $M \to \infty$ and $N \to \infty$ as shown in [18]. The eigenvectors can be well approximated in this case by he eigenvectors of a two-level circulant matrix. For the two-level circulant matrix, the eigenvectors correspond to the kronecker product between all the columns of the fast Fourier transform matrix with dimensions $M \times M$ and $N_c \times N_c$. In the following, no restriction on the values of M and N is considered. However, we assume that the correlation matrix $\mathbf{R}_{\mathbb{H}}$ is known at the transmitter side[c] and the channel matrix $\mathbf{\Gamma}(n)$ can be deduced straightforwardly from the knowledge of $\mathbf{R}_{\mathbb{H}}$ as shown in Lemma 1.

3. Reduced feedback for the selective MIMO BC
The main objective of this section is to show how to achieve the total multiplexing gain in a selective fading MIMO BC as illustrated in Figure 1 when a limited feedback bits are used to quantize the channel. For this, we start first by giving some basic preliminaries on the linear precoding over the MIMO BC in Sections 3-A and 3-B. Then, we give in Section 3-C a global overview on the general concept of the proposed quantization schemes of LTV selective fading channel. In Section 3-D, we propose quantization schemes that take advantage of the correlation between time-frequency to reduce the number of feedback bits when a zero forcing or a BD scheme are used. A practical approach to achieve the full multiplexing while keeping a low computational complexity is proposed in Section 3-E. Numerical illustrations are provided in Section 3-F.

A. Basic preliminaries
When linear precoding is used, the transmitted signal vector $\mathbf{x}(n)$ is a linear function of the destinations' data symbols $\mathbf{s}^{[k]}(n) \in \mathbb{C}^{n_r \times 1}$. Let $\mathbf{V}^{[k]}(n)$ denote the precoding matrix of user k at a time-frequency slot, such as

$$\mathbf{x}(n) = \sum_{k=1}^{K} \mathbf{V}^{[k]}(n)\mathbf{s}^{[k]}(n). \tag{18}$$

The received signal for user k is given by,

$$\mathbf{y}^{[k]}(n) = \mathbf{H}^{[k]}(n)\mathbf{V}^{[k]}(n)\mathbf{s}^{[k]}(n) \\ + \sum_{j \neq k} \mathbf{H}^{[k]}(n)\mathbf{V}^{[j]}(n)\mathbf{s}^{[j]}(n) + \mathbf{n}^{[k]}(n),$$

where the second term represents the multi-user interference from every other user's signal. In the rest of this section, let $\mathbf{h}_i^{[k]}(n)$ denote the channel seen by the receive antenna i of user k at a time-frequency slot n.

1) Zero forcing (ZF)
The zero forcing (ZF) precoder eliminates the multiuser and the inter-antenna interference. In this case, the lth columns $\mathbf{v}_l^{[j]}(n)$ of the precoding matrices $\mathbf{V}^{[j]}(n)$ are chosen such that,

$$\mathbf{h}_i^{[k]}(n)\mathbf{v}_l^{[j]}(n) = 0, \quad \forall j \neq k \in [1, K], \forall i, l \in [1, n_r],$$
$$\mathbf{h}_i^{[k]}(n)\mathbf{v}_l^{[k]}(n) = 0, \quad \forall l \neq i \in [1, n_r],$$

The received signal is given by

$$y_i^{[k]}(n) = \mathbf{h}_i^{[k]}(n)\mathbf{v}_i^{[k]}(n) \, s_i^{[k]}(n) \; + n_i^{[k]}(n), \\ i = 1 \ldots n_r.$$

2) Block diagonalization
When the BD precoding schemes is used, the precoding matrices are chosen in order to eliminate the multiuser interference only, such that

$$\mathbf{H}^{[k]}(n)\mathbf{V}^{[j]}(n) = 0 \quad \forall j \neq k \in [1, K].$$

The received signal at each user side is given by

$$\mathbf{y}^{[k]}(n) = \mathbf{H}^{[k]}(n)\mathbf{V}^{[k]}(n)\mathbf{s}^{[k]}(n) + \mathbf{n}^{[k]}(n).$$

3) Achieving the full multiplexing gain

When full CSIT is assumed, it is known from [5] that both schemes achieve the full multiplexing gain of min $(n_t, Kn_r) = Kn_r$, which means that the sum capacity C_{sum} (P) of the BC scales as,

$$C_{\text{sum}}(P) \sim Kn_r \log_2 P.$$

More recently, it has been shown in [6,7] that the full CSIT assumption can be reduced to a partial CSIT when performing digital quantization of the flat MIMO channel using a number of feedback bits $N_{f,\,\text{flat}}$ that scales as

$$N_{f,\text{flat}} = \begin{cases} n_r(n_t - 1)\log_2 P & \text{with a ZF precoder,} \\ n_r(n_t - n_r)\log_2 P & \text{with a BD precoder.} \end{cases}$$

B. A first approach: straightforward approach (SA)

The first approach to quantize the time-frequency selective channel can be straightforwardly deduced from the quantization of the flat fading channel. Each time-frequency channel $\mathbf{H}^{[k]}(n)$ can be simply seen a flat fading channel and is quantized by its own using

$$N_{f,n} = N_{f,\text{flat}} \quad n = 0 \dots N - 1,$$

feedback bits in order to achieve the full multiplexing gain. Over the N time-frequency slots, the total number of feedback bits is such that

$$N_f = N_{f,0} + \dots + N_{f,N-1} = N N_{f,\text{flat}}.$$

Although, the straightforward strategy achieves the full multiplexing gain, it is not optimal in the sense that the number of feedback bits sent to the source is very large. Moreover, this feedback contains redundant information about the channels. In the next two subsections, we will show how the correlation between the time-frequency channels can be used in order to reduce the number of feedback bits.

C. Quantizing the selective fading channel: general concepts

In this section, we give a global overview on the general concept of the proposed quantization of LTV selective fading channel.

The training phase in Figure 2 aims to estimate and quantize these channel *invariant* parameters and includes: (1) The estimation of the LTV invariants parameters at each receiver side. (2) The estimated invariant

parameters of the channel are quantized and feedback to the transmitter side.

1) Perfect channel estimation at the receiver side

For the MIMO BC, each receiver needs to estimate the ρn_t invariants parameters[d] at each receive antennas. For this, the transmitter sends a common pilot common n_t $\times$ 1 vector signal of length $n_t\rho$ to all receivers such that,

$$\mathbf{x}(n) = \mathbf{e}_p \quad n = 0 \dots n_t\rho - 1 \tag{19}$$

where $p = (n \bmod n_t) + 1$ and $\mathbf{e}_j$ is the j-th vector of the $n_t \times n_t$ identity matrix with entries $e_{j,\,n} = \delta_{n\,-\,j}$.

Assuming a perfect estimation at the receiver side[e] and using the above pilot sequence, the value of the time-frequency channel between transmit antenna i and receiver antenna j is available at each antenna j for a time-frequency slot equal to $n_t q + (i - 1)$ with $q = 0 \dots \rho$ $- 1$. The ρ *invariant* parameters $h_{\omega,\,i,\,j}[s]$ $(s = 0 \dots \rho - 1)$ of the channel between transmit antenna i and receive antenna j can be therefore evaluated as a solution of a linear system with ρ independent equations and ρ unknown variables, given by

$$h_{i,j}^{[k]}(n_t q + (i - 1)) = \sum_{s=0}^{\rho-1} \sigma_s w_{n_t q + (i-1), s} h_{\omega, i, j}[s],$$

with $q = 0 \dots \rho - 1$. The duration of this estimation phase is given by,

$$T_e = \left(\left\lfloor \frac{n_t\rho}{N_c} \right\rfloor + 1 \right) T. \tag{20}$$

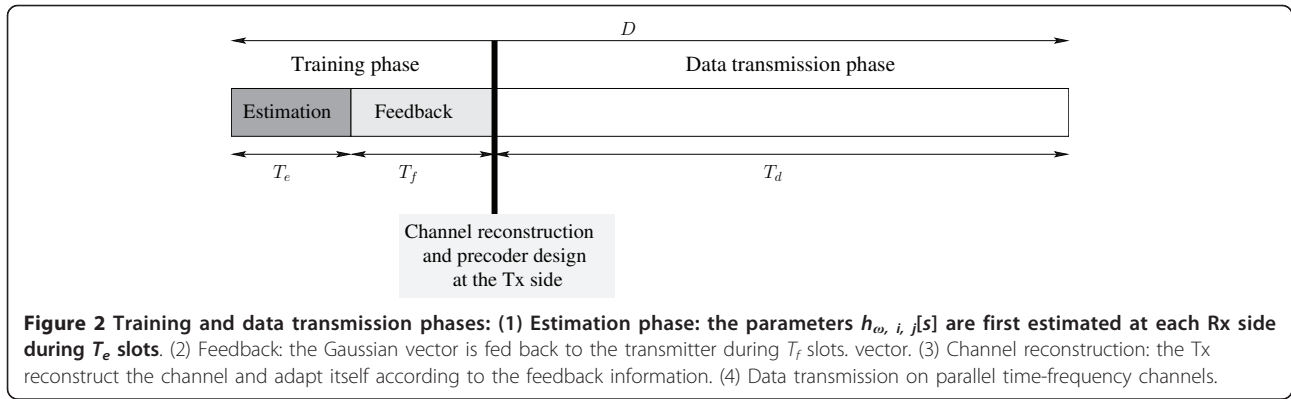### 2) Quantization and feedback of the Gaussian vector

The estimated invariant parameters of the channel at each receiver are then quantized using N_f bits. This quantization is based on grouped or partitioned feedback that will be detailed in Sections 3-D and 3-E. The duration of this phase depends on the uplink reliable rate R_u on which the feedback bits are conveyed to the transmitter side and is such that

$$T_f = K\frac{N_f}{R_u}. \tag{21}$$

3) Channel reconstruction and precoder design

After this training phase, the transmitter should be able to reconstruct the channel using the quantized channel invariant parameters and the statistical channel knowledge as shown in Lemma 1. At each time-frequency slot, the linear precoder is adapted to the quantized time-frequency channel, and data are transmitted to the different users as shown in Figure 2.

Generally, the exact derivation of the effective maximal rate should take into account the training phase, meaning that the effective duration of transmission is,

Figure 2 Training and data transmission phases: (1) Estimation phase: the parameters $h_{\omega,\,i,\,j}[s]$ are first estimated at each Rx side during T_e slots. (2) Feedback: the Gaussian vector is fed back to the transmitter during T_f slots. vector. (3) Channel reconstruction: the Tx reconstruct the channel and adapt itself according to the feedback information. (4) Data transmission on parallel time-frequency channels.

$$T_d = D - (T_e + T_f). \tag{22}$$

However, to make a fair comparison with the perfect CSIT (where the training phase is often omitted), the rate gap between the quantized rate and the perfect CSIT that will be considered in the following is derived considering only the effective time-frequency slot where data information is transmitted using $N_d = \frac{T_d}{D} N_c$ time-frequency slots.

D. Grouped reduced feedback (GRF) for selective fading BC

In this section, the estimated elements are grouped into one $1 \times \rho n_t$ vector (respectively into one $\rho n_t \times n_r$ matrix) and quantized using a random vector quantization (RVQ) when zero forcing precoder is used (respectively using a Grassmannian quantization with BD).

1) Zero forcing with grouped reduced feedback (GRF-ZF)

At each time frequency slot n, the channel seen by each receive antenna j for a user k can be written using Lemma 1 as

$$\mathbf{h}_j^{[k]}(n) = \mathbf{h}_{\omega,j}^{[k]} \Gamma(n), \quad j = 1...n_r, \quad k = 1...K. \tag{23}$$

where $\mathbf{h}_{\omega,j}^{[k]} \in \mathbb{C}^{1 \times \rho n_t}$ is a Gaussian vector with i.i.d entries. As it can be noticed from (23), it is sufficient to know $\mathbf{h}_{\omega,j}^{[k]} \in \mathbb{C}^{1 \times \rho n_t}$ to determine the channel at each time frequency slot and at each antenna $j = 1 ... n_r$. For the selective fading BC when a zero forcing precoder is used, we prove in Theorem 1 that it is sufficient to quantize n_r unit norm vectors, $\tilde{\mathbf{h}}_{\omega,j}^{[k]} = \frac{1}{||\mathbf{h}_{\omega,j}^{[k]}||} \mathbf{h}_{\omega,j}^{[k]}$ at each user k to achieve the full multiplexing gain using a RVQ technique. We assume that each destination uses n_r different codebooks to quantize each vector $\tilde{\mathbf{h}}_{\omega,j}^{[k]}$ in order to prevent quantizing two different vectors by the same vector. The quantization codebook $\mathcal{C} = \{\omega_1, ..., \omega_{2^B}\}$ containing 2^B unit norm Gaussian $1 \times \rho n_t$ vectors is assumed to be known at the transmitter and receiver

side. At each user k, each antenna j feeds the index $F^{[k,\,j]}$ of the ω vectors that is closest (in term of its angle) to its channel vector $\mathbf{h}_{\omega,j}^{[k]}$. The minimal number of feedback bits required to achieve the full multiplexing gain is summarized in Theorem 1 as following.

Theorem 1: For the K-selective MIMO BC with n_t transmit antennas at the source and n_r receive antennas at the destinations ($n_t \geq K n_r$) when a zero forcing scheme is used, the total spatial multiplexing gain of $K n_r$ can be achieved using the above RVQ scheme if the number of feedback bits N_f sent by each user scales as

$$N_f = n_r(\rho n_t - 1) \log_2 P, \tag{24}$$

where ρ is the rank of the selective fading channel covariance matrix. At high SNR, the rate loss incurred by the above quantization scheme is upper bounded by,

$$\Delta R_k \leq n_r \log_2 \left(1 + \sigma_t^2 \frac{\rho(K-1)\zeta}{\rho n_t - 1} \right), \tag{25}$$

Proof: The proof of this theorem is mainly based on the previous quantization result of the flat fading channel in [6] and is detailed in Appendix B.

2) BD with grouped reduced digital feedback (GRF-BD)

In this section, we propose a quantization scheme for the BD when a time-frequency selective channel is considered. Based on the observation that time-frequency selective channel slots are correlated, we compute the minimal number of feedback bits required to achieve the full multiplexing gain.

As the time-frequency channel matrices $\mathbf{H}^{[k]}(n)$ are correlated, it is not necessary that the receiver feeds back the channel at each time-frequency slot. It can be easily deduced from Lemma 1 that the knowledge of $\mathbf{H}_\omega^{[k]} \in n_r \times \rho n_t$ is sufficient to know the channel at each time-frequency slot. Usually, quantized matrices are chosen in a Grassmannian manifold $\mathcal{G}(T, M)$, where $T > M$. That's why, the quantization problem of selective fading channel consists in finding a quantization for

(with[f] $\rho n_t \geq n_r$). The quantization codebook is supposed to be known at the transmitter side as well at the receivers side. Each receiver uses a different codebook $\mathbf{C}_k$ of 2^B unitary matrices in $\mathbb{C}^{\rho n_t \times n_r}$, such that

$$\mathbf{C}_k = \{\mathbf{W}_1, ..., \mathbf{W}_{2^B}\}.$$

Each user (to say user k) feeds back the index of the $\mathbf{W}$ matrix that is closest in term of its chordal distance to the channel matrix $\mathbf{H}_\omega^{[k]\dagger} \in \mathbb{C}^{\rho n_t \times n_r}$, i.e.,

$$\hat{\mathbf{H}}_\omega^{[k]} = \arg\min_{\mathbf{W} \in \mathcal{C}_k} d^2(\mathbf{H}_\omega^{[k]}, \mathbf{W}), \tag{26}$$

where d is the chordal distance between two matrices. Each of the 2^B unitary matrices are chosen independently and are uniformly distributed over a Grassmannian $\mathcal{G}(\rho n_t, n_r)$. As shown in [7] (and references therein), the distortion associated with a given codebook C_k for the quantization of $\mathbf{H}_\omega^{[k]\dagger} \in \mathbb{C}^{\rho n_t \times n_r}$ is such that

$$D_s = \mathbb{E}\,[d^2(\mathbf{H}_\omega^{[k]}, \hat{\mathbf{H}}_\omega^{[k]})] \leq \bar{D}, \tag{27}$$

where $\bar{D}$ is equivalent when the number of bits B goes to infinity to

$$\bar{D} \xrightarrow{B \to \infty} C\, 2^{-\frac{B}{n_r(\rho n_t - n_r)}}. \tag{28}$$

and C is a constant independent of B given by

$$C = \frac{\Gamma(\frac{1}{g})}{(g+1)} \left(\frac{1}{g!} \prod_{i=1}^{n_r} \frac{(\rho n_t - i)!}{(n_r - i)!} \right)^{-\frac{1}{g}}. \tag{29}$$

with $g = n_r(\rho n_t - n_r)$ is the dimensionality of the Grassmannian manifold. The minimal number of feedback bits required to achieve the full multiplexing gain is summarized in Theorem 2 as following.

Theorem 2: For the K-selective MIMO BC with n_t transmit antennas at the source and n_r receive antennas at the destinations($n_t \geq Kn_r$) when a BD scheme is used, the total spatial multiplexing gain of Kn_r can be achieved using the above grouped Grassmannian quantization scheme if the number of feedback bits N_f broadcast by each user scales as,

$$N_f = n_r(\rho n_t - n_r) \log_2 P, \tag{30}$$

where ρ is the rank of the selective fading channel covariance matrix. At high SNR, the rate loss incurred by the above quantization scheme is upper bounded by

$$\Delta R_k \leq n_r \log_2 \left(1 + \sigma_t^2 \frac{\rho(K-1)}{\rho n_t - n_r} C \right), \tag{31}$$

where C is a constant defined as

$$C = \frac{\Gamma(\frac{1}{g})}{(g+1)} \left(\frac{1}{g!} \prod_{i=1}^{n_r} \frac{(\rho n_t - i)!}{(n_r - i)!} \right)^{-\frac{1}{g}},$$

with $g = n_r(\rho n_t - n_r)$ is the dimensionality of the Grassmannian manifold.

Proof: Please refer to Appendix C for the proof of this theorem. ■

E. Partitioned-reduced feedback (PRF): a practical approach to achieve the full multiplexing gain

It can be noticed from Section 3-D that when grouping at each receiver all the invariants parameters of the channel into one $1 \times \rho n_t$ vector (respectively one $\rho n_t \times n_r$ matrix), the size of the codebook required to achieve full multiplexing gain is very large, and consequently the search complexity of the optimal vector (respectively matrix) in the codebook becomes very high.

For instance, if we consider a MIMO BC with a ZF precoder with $n_t = 6$ antennas, a covariance matrix rank $\rho = 12$ and $n_r = 2$, the number of feedback bits required to quantize a $1 \times \rho n_t$ vector at an SNR = 30 dB is 707 bits. In order to quantize this vector, one needs to search the optimal vector in a codebook of size 2^{707}, which is not always feasible to be implemented in a practical system.

Motivated by this issue, we propose in this subsection a practical approach to quantize the time-frequency selective channel, which will be called in the following PRF. The proposed strategy guarantees to exploit all the available degrees of freedom in the MIMO BC when a *zero forcing* precoder is used while keeping a low computational complexity.

1) PRF strategy

The main objective of the proposed PRF scheme with ZF precoding is to quantize the channel vector containing the invariants Gaussian parameters $\mathbf{h}_{\omega,j}^{[k]} \in \mathbb{C}^{1 \times \rho n_t}$ seen at the receive antenna j of user k in a partitioned way.

This requires to write $\mathbf{h}_{\omega,j}^{[k]}$ as the concatenation of ρ vectors $\mathbf{h}_{\omega,j}^{[k]}[i] \in \mathbb{C}^{1 \times n_t}$ $(i = 0...\rho - 1)$ such that

$$\mathbf{h}_{\omega,j}^{[k]} = [\mathbf{h}_{\omega,j}^{[k]}[0] \,...\, \mathbf{h}_{\omega,j}^{[k]}[\rho - 1]].$$

The proposed PRF scheme consists to find for each partitioned channel vector $\mathbf{h}_{\omega,j}[i]$:

- Its quantized norm $\| \mathbf{h}_{\omega,j}^{[k]}[i] \|_a$ using a noisy analog feedback scheme. We assume that these coefficients are sent β times on an unfaded uplink AWGN channel with the same power as the downlink scheme.

The received signal at the base station is,

$$\gamma = \sqrt{\beta P}||\mathbf{h}_{\omega,j}^{[k]}[i]|| + e_i,$$

This implies that the unbiased MMSE estimated norm is,

$$|| \mathbf{h}_{\omega,j}^{[k]}[i]||_a = ||\mathbf{h}_{\omega,j}^{[k]}[i]|| + \frac{1}{\sqrt{1 + \beta P}}e_i, \qquad (32)$$

where e_i is the feedback Gaussian noise such that $e_i \sim \mathcal{CN}(0, 1)$.

- Its direction by finding the optimal orthonormal vector $\hat{\mathbf{h}}_{\omega,j}[i]$ in a random codebook of size 2^B that is closest in term of its angle to $\mathbf{h}_{\omega, j}[i]$ using a RVQ technique, such that,

$$\hat{\mathbf{h}}_{\omega,j}^{[k]}[i] = \arg \max_{\mathbf{v} \in C_i} < \mathbf{v}, \mathbf{h}_{\omega,j}^{[k]}[i] > .$$

Finally, the normalized corresponding $\rho n_t \times 1$ quantized vector $\tilde{\mathbf{h}}_{\omega,j}^{[k]}$ is constructed as following

$$\hat{\mathbf{h}}_{\omega,j}^{[k]} = \frac{1}{\left(\sum_{i=0}^{\rho-1} ||\mathbf{h}_{\omega,j}^{[k]}[i]||_a^2\right)^{1/2}} \begin{bmatrix} ||\mathbf{h}_{\omega,j}^{[k]}[0]||_a \hat{\mathbf{h}}_{\omega,j}^{[k]}[0] \\ ||\mathbf{h}_{\omega,j}^{[k]}[1]||_a \hat{\mathbf{h}}_{\omega,j}^{[k]}[1] \\ \vdots \\ ||\mathbf{h}_{\omega,j}^{[k]}[\rho-1]||_a \hat{\mathbf{h}}_{\omega,j}^{[k]}[\rho-1] \end{bmatrix}. \quad (33)$$

In the following, let

$$||\mathbf{h}_{\omega,j}^{[k]}||_a = \left(\sum_{i=0}^{\rho-1} ||\mathbf{h}_{\omega,j}^{[k]}[i]||_a^2\right)^{1/2},$$

and

$$||\mathbf{h}_{\omega,j}^{[k]}|| = \left(\sum_{i=0}^{\rho-1} ||\mathbf{h}_{\omega,j}^{[k]}[i]||^2\right)^{1/2}.$$

In order to derive the gap rate with the full CSIT, we characterize in Lemma 2 the angle between the normalized vector $\tilde{\mathbf{h}}_{\omega}^{[k]} = \frac{1}{||\mathbf{h}_{\omega}^{[k]}||}\mathbf{h}_{\omega}^{[k]}$ vector and its quantized vector $\hat{\mathbf{h}}_{\omega}^{[k]}$.

Lemma 2: The quantization error of the PRF scheme operating at high SNR is upper bounded by,

$$\frac{1}{2}\sin^2(\tilde{\mathbf{h}}_{\omega}^{[k]}, \hat{\mathbf{h}}_{\omega}^{[k]}) \leq \sum_{i=0}^{\rho-1} \frac{||\mathbf{h}_{\omega}^{[k]}[i]||^2}{||\mathbf{h}_{\omega}^{[k]}||^2} \sin^2(\tilde{\mathbf{h}}_{\omega}^{[k]}[i], \bar{\mathbf{h}}_{\omega}^{[k]}[i]).$$

The expected quantization error is upper bounded by,

$$\frac{1}{2}\mathbb{E}[\sin^2(\tilde{\mathbf{h}}_{\omega}^{[k]}, \hat{\mathbf{h}}_{\omega}^{[k]})] \leq 2^{-\frac{B}{n_t-1}}.$$

Proof: Please refer to Appendix D for the proof of this lemma.

2) Achieving the full multiplexing gain

For a selective fading MIMO BC with $n_t \geq Kn_r$, we show in Theorem 3 that when using the above PRF scheme with a sufficient number of feedback bits and a zero forcing precoder, the total multiplexing gain can be also achieved. The following result is summarized in the following theorem.

Theorem 3: For a selective fading MIMO BC with $n_t \geq Kn_r$, the total spatial multiplexing gain of Kn_r can be achieved using the above partitioned vector quantization scheme if

- The direction of each vector $\mathbf{h}_{\omega,j}^{[k]}[i](i = 0...\rho - 1)$ is quantized using

$$N_{f,d} = n_r\rho(n_t - 1)\log_2 P \text{ bits.}$$

- The norm of each vector $\mathbf{h}_{\omega,j}^{[k]}[i]$ is quantized using

$$N_{f,a} = \beta\rho\log_2 P \text{ bits.}$$

The total number of feedback bits N_f sent by each user is,

$$N_f = n_r\rho(n_t - 1)\log_2 P + \beta\rho\log_2 P. \qquad (34)$$

At high SNR, the upper bound on the rate loss incurred by the PRF quantization scheme is

$$\Delta R_k \leq n_r \log_2\left(1 + 2\sigma_t^2\frac{\rho(K - 1)\zeta}{\rho n_t - 1}\right). \qquad (35)$$

Proof: Please refer to Appendix E for the proof of this theorem. ∎

Remark 3: It should be emphasized here that the quantization of the directions of the sub-vectors is not sufficient alone to achieve the full multiplexing gain if it is not coupled with the feedback of these vectors' norms. This is also illustrated in the numerical results in Section 3-F and can be analytically proved following the same reasoning as above.

3) Reducing the computational search complexity

Although the number of feedback bits required by the PRF scheme is slightly increased compared to the GRF scheme, the PRF strategy reduces significantly the complexity of the system. The quantization problem of the whole vector is decomposed into ρ sub-quantization problems where each sub-quantization deals with a smaller size of vector. Instead of performing one search in one codebook of size 2^{B_g} as shown in Subsection 3-D, where $B_g = (\rho n_t - 1)\log_2 P$, the PRF schemes PRF

perform ρ searches in ρ different codebooks of size B_p each with $B_p = \rho(n_t - 1) \log_2 P$. Thus,

$$\frac{2^{B_g}}{\rho 2^{B_p}} = \frac{1}{\rho} P^{n_r(\rho-1)}.$$

It can deduced therefore that the total search complexity in the PRF scheme is reduced by a factor of $\frac{1}{\rho} P^{n_r(\rho-1)}$ compared to the GRF scheme. This factor becomes very significant for high SNR ranges.

We finally note that partitioning the vectors is not restricted to ρ vectors with n_t elements. Increasing the vectors partitions comes at the expense of an increased feedback, but a significant reduced computational complexity of the system. A tradeoff between the complexity of the system and the number of feedback bits should be considered.

F. Numerical results

In order to illustrate the proposed reduced quantized schemes, we consider the MIMO BC with $K = 3$ destinations having $n_r = 2$ receive antennas each and a source with $n_t = 6$ transmit antennas. We assume that the communication occurs over a radio channel characterized by the parameters in Table 1 (Table 2.1 in [19]). These parameters correspond to the context of the standard IEEE 802.16 (or WIMAX).

In this case, it can be easily verified that the signal duration and bandwidth are much larger than the coherence bandwidth and the coherence time of the channel. The channel is therefore selective in time and frequency. This channel can be approximately decomposed into parallel time-frequency channel using WH sequences as explained in Section 2. We assume in the following that the grids parameters are chosen such that $TF = 1.25$, where $T = 0.1$ ms and $F = 12.5$ KHz. The number of time-frequency slots is therefore,

$$N = \frac{DW}{TF} = 40000,$$

Table 1 Channel and signal parameters

| Key channel and signal parameters	Values
Carrier frequency f_c	2.5 GHz
Communication bandwidth W	1 MHz
Delay requirement D	50 ms
Doppler spread v_0	100 Hz
Coherence time T_c	2.5 ms
Delay spread τ_0	1 μs
Coherence bandwidth W_c	500 KHz
Channel spread $\Delta_{\mathbb{H}}$	10^{-4}
Reliable uplink rate	1 Mbps

and the sufficient number of parameters required to identify the channel is $\rho \approx (\lfloor v_0 D \rfloor + 1)(\lfloor \tau_0 W \rfloor + 1) = 12 \ll N$. The required estimation period can be deduced from (20) and is equal to,

$$T_e = \left(\left\lfloor \frac{n_t \rho}{N_c} \right\rfloor + 1 \right) T = T = 0.1 \text{ ms}.$$

For this channel and signal model, we compare[g] the proposed strategies: the SA with ZF (SA - ZF), the SA with ZF (SA - BD), the GRF scheme[h] with zero forcing (GRF - ZF) and the PRF. The classical comparisons of ZF and BD are extensively addressed in [6,7] and the same behavior as for the flat fading channel can be observed for the TF selective channel. In the rest of this subsection, we focus mainly on the performance of the practical proposed PRF scheme. We note here that the PRF scheme is not compatible with the BD construction as it is based on a RVQ and not on a quantization over a Grassmann manifold [7].

Figure 3 compares the number of feedback bits required by each strategy. We can notice that the number of feedback bits using reduced feedback strategies is significantly reduced compared to straightforward strategies and that the number of feedback bits required by the PRF strategy is slightly increased compared to the GRF - ZF. The total duration of the feedback period in (21) depends on this number of feedback bits and the quantization strategy. For instance, the feedback time required by the PRF scheme operating at an SNR = 30 dB is $T_f = K \frac{N_f}{R_u} = 2.15$ ms. However, for the SA-BD, the feedback time is equal to 12 ms considered as a large delay as it constitutes nearly 1/4 of the total duration of the communication.

At the end of the training phase, the source reconstructs the channel using one of the above strategies and adapts itself to the quantized channel. Figure 4 compares the sum capacity of the MIMO BC when the above quantization strategies are used. When the number of feedback bits of each strategy scales with SNR as shown in Theorems 1 and 3, we can observe that the full multiplexing gain of $\min(n_t, K n_r) = 6$ can be achieved. Moreover, the performances of the straightforward strategies are the same as the reduced ones. As stated before, the PRF scheme reduces significantly the complexity of the system as it deals with the quantization of vector of 1×6 vector instead of 1×72 as it is the case for the GRF scheme. We can also see in Figure 4 that the full multiplexing is achieved and there is a negligible rate loss between the PRF and the GRF-ZF of nearly ~0.1 dB. It can be deduced that the same performance of the GRF-ZF strategy can be achieved but with a much lower computational complexity.

Figure 3 Reduced feedback scheme versus SA.
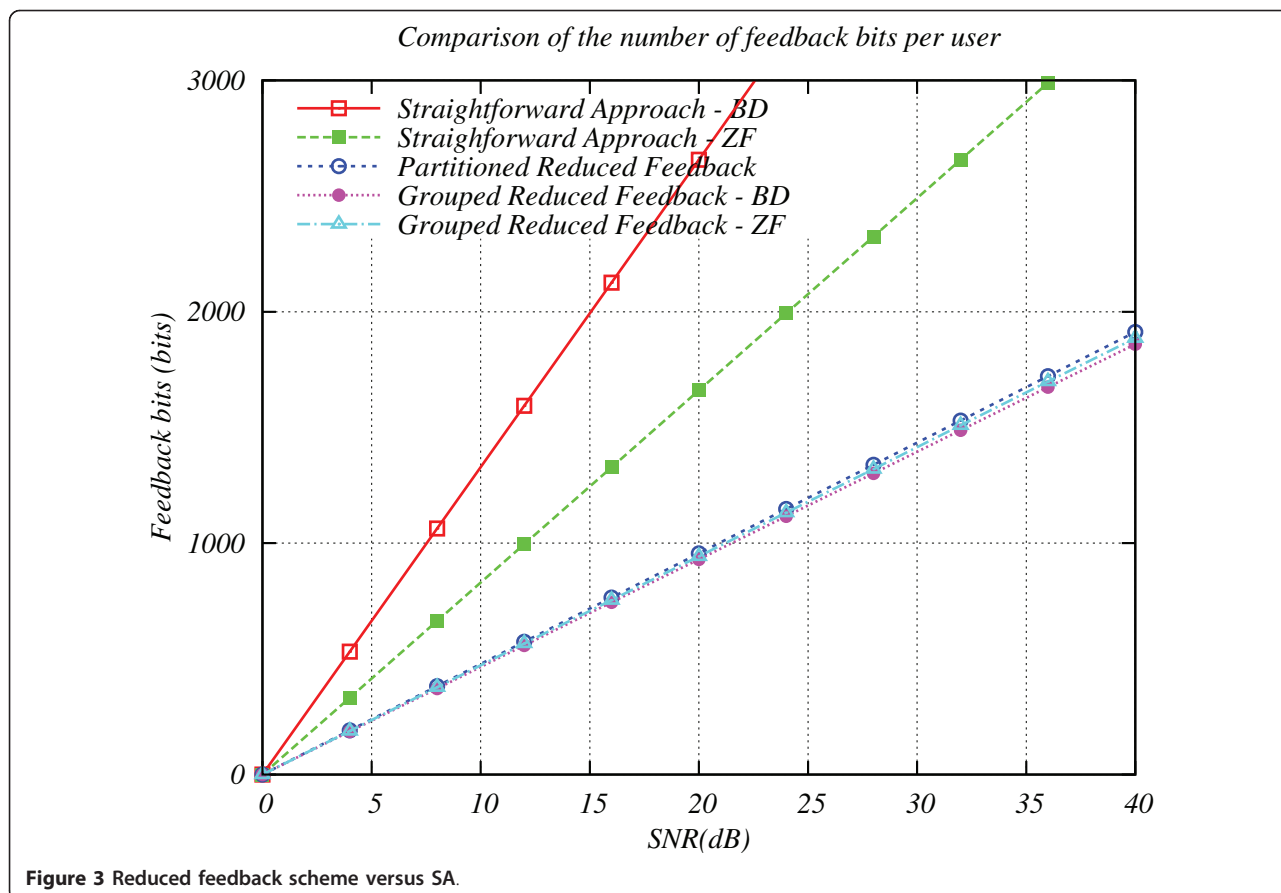
When no feedback information on the sub-vectors norms is provided to the transmitter side, we can observe in Figure 4 that the full multiplexing gain cannot be achieved. The directional knowledge should be coupled with the sub-vectors' norms quantification to achieve the full multiplexing gain.

Finally, we can observe in Figure 5 that when the number of feedback bits $N_f = 200$ does not scale with SNR that the GRF scheme gives a better reconstruction of the channel at the transmitter side and outperforms all the other strategies. This result is not surprising as the 200 bits will be divided for the SA-ZF between all the time-frequency slots, and each vector is quantized using 4 bits only. For the PRF scheme, a part of the 200 feedback bits is used for the analog feedback of sub-vectors norm and the remaining feedback bits are used to quantize $n_r\rho$ sub-vectors. Each sub-vector is therefore quantized using only a few number of feedback bits that are not sufficient for channel reconstruction.
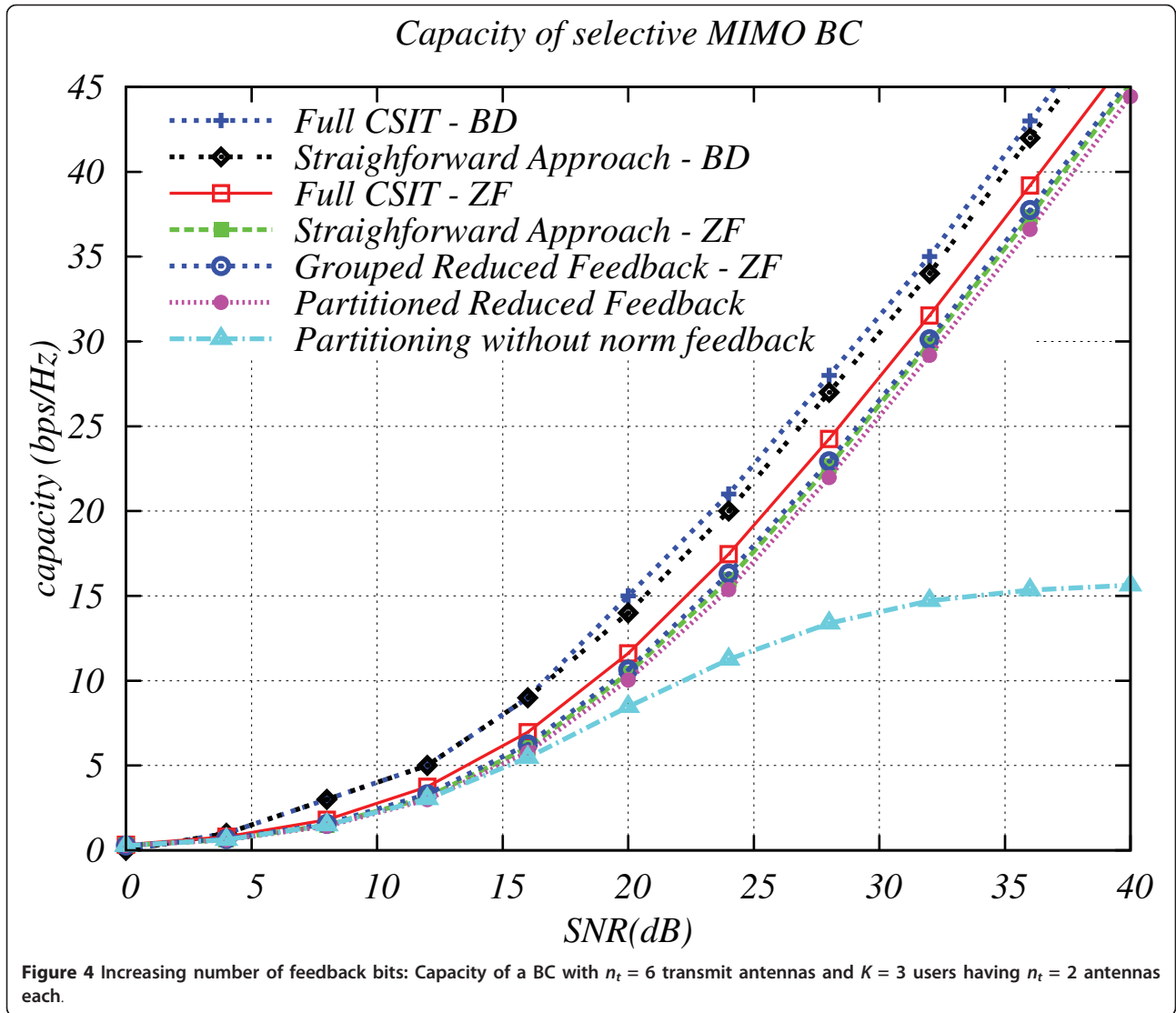
4. Conclusions and perspectives
In this article, we studied the selective MIMO BC with limited feedback. We showed that as time-frequency

channels are correlated it is not necessary to quantize each time-frequency channel. However, it is sufficient to reconstruct the channel based on a finite number of parameters by making use of the correlation in time and frequency while conversing the full spatial multiplexing gain. The optimal number of feedback bits required to achieve the full multiplexing gain is computed. Moreover, the practical implementation of a TF channel quantization scheme is addressed and a low-complexity scheme that also achieves the multiplexing gain is proposed. The design of a PRF matricial scheme compatible with the BD scheme will be addressed in our forthcoming works.

Appendix A
Proof of Lemma 1
Let $\mathbf{h}_{i,j}^{[k]} = [h_{i,j}^{[k]}(0)...h_{i,j}^{[k]}(N-1)]$ be the $N \times 1$ stacked channel vector that contains the N time-frequency channel's components, and $\mathbf{R}_{\mathbb{H}}$ its $N \times N$ Hermitian covariance matrix such that $\mathbf{R}_{\mathbb{H}} = \mathbb{E}[\mathbf{h}_{i,j}^{[k]}\mathbf{h}_{i,j}^{[k]\dagger}]$. The covariance channel matrix coefficients can be deduced from (2) and is supposed to be known at both the transmitter and the receiver side. In the following, we set $\rho = \text{rank}\{\mathbf{R}_{\mathbb{H}}\}$

Figure 4 Increasing number of feedback bits: Capacity of a BC with $n_t = 6$ transmit antennas and $K = 3$ users having $n_t = 2$ antennas each.

the rank of $\mathbf{R}_\mathbb{H}$ and $\mathbf{R}_\mathbb{H} = \mathbf{W}\Lambda\mathbf{W}^\dagger$ its eigenvalue decomposition where $\Lambda = \mathrm{diag}\{\sigma_0^2, ..., \sigma_{\rho-1}^2, 0, ..., 0\}$.

The vector $\mathbf{h}_{i,j}^{[k]}$ can be written in function of its covariance matrix $\mathbf{R}_\mathbb{H}$ such that

$$\mathbf{h}_{i,j}^{[k]} = \mathbf{R}_\mathbb{H}^{1/2}\bar{\mathbf{h}}_{i,j}^{[k]},$$

where $\bar{\mathbf{h}}_{i,j}^{[k]}$ is an i.i.d. $\mathcal{CN}(0, 1)$ vector with the same dimension as $\mathbf{h}_{i,j}^{[k]}$. Using the eigenvalue decomposition of $\mathbf{R}_\mathbb{H}$,
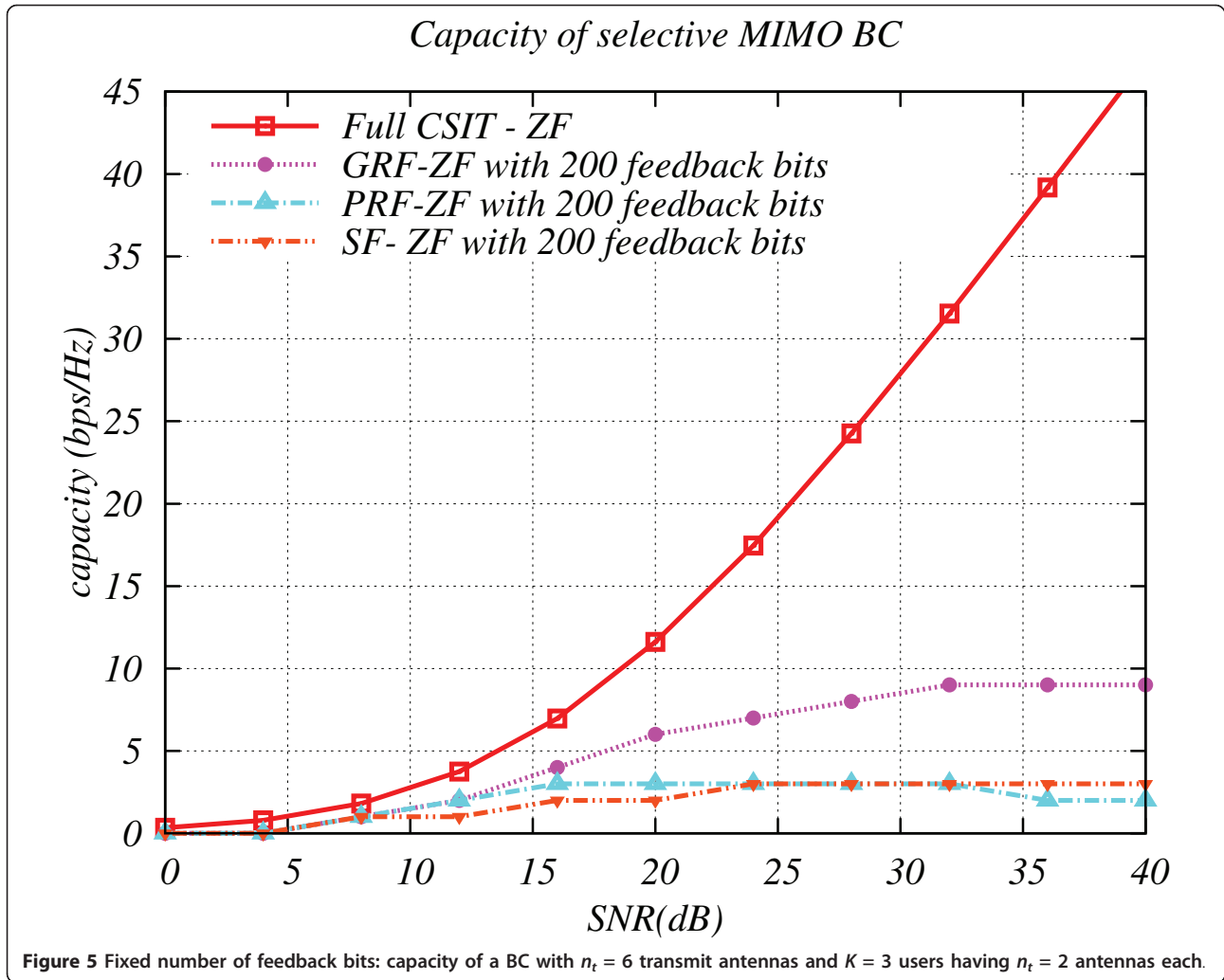
$$\begin{aligned}\mathbf{h}_{i,j}^{[k]} &= \mathbf{W}\Lambda^{1/2}\mathbf{W}^\dagger\bar{\mathbf{h}}_{i,j}^{[k]}, \\ &= \mathbf{W}\Lambda^{1/2}\mathbf{h}_{\omega,i,j}^{[k]},\end{aligned} \quad (36)$$

where $\mathbf{h}_{\omega,i,j}^{[k]}$ in (36) is also a random Gaussian vector $\mathcal{CN}(0, 1)$, since $\mathbf{W}^\dagger$ is a unitary matrix. It follows from (36) that

$$h_{i,j}^{[k]}(n) = \sum_{l=0}^{\rho-1} w_{n,l}\sigma_l h_{\omega,i,j}^{[k]}[l], \quad n = 0 \ldots N - 1. \quad (37)$$

Note here that the number of independent parameters $h_{\omega,i,j}^{[k]}[l]$ is nothing but the rank of the covariance matrix $\mathbf{R}_\mathbb{H}$. As channels between transmit antennas and receive antennas are not correlated, this implies that $\mathbf{H}^{[k]}(n)$ can be written as shown in (38).

$$H^{[k]}(n) = \begin{bmatrix} \sum_{l=0}^{\rho-1} w_{n,l}\sigma_l h_{\omega,1,1}^{[k]}[l] & \cdots & \sum_{l=0}^{\rho-1} w_{n,l}\sigma_l h_{\omega,1,n_t}^{[k]}[l] \\ & \vdots & \\ \sum_{l=0}^{\rho-1} w_{n,l}\sigma_l h_{\omega,n_r,1}^{[k]}[l] & \cdots & \sum_{l=0}^{\rho-1} w_{n,l}\sigma_l h_{\omega,n_r,n_t}^{[k]}[l] \end{bmatrix} \quad (38)$$

Figure 5 Fixed number of feedback bits: capacity of a BC with $n_t = 6$ transmit antennas and $K = 3$ users having $n_t = 2$ antennas each.

Equivalently,

$$\mathbf{H}^{[k]}(n) = \sum_{l=0}^{\rho-1} w_{n,l} \sigma_l \begin{bmatrix} h_{\omega,1,1}^{[k]}[l] & \cdots & h_{\omega,1,n_t}^{[k]}[l] \\ & \vdots & \\ h_{\omega,n_r,1}^{[k]}[l] & \cdots & h_{\omega,n_r,n_t}^{[k]}[l] \end{bmatrix}$$

$$= \sum_{l=0}^{\rho-1} w_{n,l} \sigma_l \mathbf{H}_{\omega,l}^{[k]}$$

By using some simple matricial manipulations, we get,

$$\mathbf{H}^{[k]}(n) = [\mathbf{H}_{\omega,0}^{[k]} \mathbf{H}_{\omega,1}^{[k]} \cdots \mathbf{H}_{\omega,\rho-1}^{[k]}] \Gamma(n)$$

where

$$\Gamma(n) = \begin{bmatrix} \sigma_0 w_{n,1} \\ \vdots \\ \sigma_{\rho-1} w_{n,\rho} \end{bmatrix} \otimes \mathbf{I}_{n_t}. \qquad (39)$$

which completes the proof.

Appendix B
Proof of Theorem 1
A. Relationship between the matrix and its quantification

Let a be the quantization error between the normalized vector $\mathbf{h}_{\omega,j}^{[k]}$ and its quantified vector $\hat{\mathbf{h}}_{\omega,j}^{[k]}$ such that,

$$a = \sin^2(\tilde{\mathbf{h}}_{\omega,j}^{[k]}, \hat{\mathbf{h}}_{\omega,j}^{[k]}).$$

As shown in Lemma 2 in [6], the vector $\mathbf{h}_{\omega,j}^{[k]}$ can be written as the sum of two vectors, one, $\hat{\mathbf{h}}_{\omega,j}^{[k]}$ in the direction of the quantized vector and the second, $\mathbf{s}_j$ is isotropically distributed in the null-space of $\hat{\mathbf{h}}_{\omega,j}^{[k]}$, independent of a as shown in [6], such that

$$\tilde{\mathbf{h}}_{\omega,j}^{[k]} = \frac{1}{||\mathbf{h}_{\omega,j}^{[k]}||} \mathbf{h}_{\omega,j}^{[k]} = \sqrt{1-a}\, \hat{\mathbf{h}}_{\omega,j}^{[k]} + \sqrt{a}\, \mathbf{s}_j.$$

This implies that,

$$\mathbf{h}_j^{[k]}(n) = \mathbf{h}_{\omega,j}^{[k]}\boldsymbol{\Gamma}(n),$$
$$= ||\mathbf{h}_{\omega,j}^{[k]}|| \left(\sqrt{1-a}\,\hat{\mathbf{h}}_{\omega,j}^{[k]}\boldsymbol{\Gamma}(n) + \sqrt{a}\,\mathbf{s}_j\boldsymbol{\Gamma}(n) \right), \quad (40)$$

with $\mathbb{E}\left[||\mathbf{h}_{\omega,j}^{[k]}||\right] = \rho n_t$.

B. Throughput analysis

Let $\Delta R_k = R_{\text{Quant}} - R_{\text{Full CSIT}}$ be the rate loss incurred by the quantization. Due to the isotropic nature of the channel matrices, the rate loss can be written such that,

$$\Delta R_k \leq \mathbb{E}\left[\frac{1}{N_d} \sum_{n=0}^{N_d-1} \sum_{j=1}^{n_r} \log_2\left(1 + \frac{P\zeta}{n_t} \sum_{i\neq k} ||\mathbf{h}_j^{[k]}(n)\hat{\mathbf{v}}^{[i]}(n)||^2 \right) \right],$$

$$\leq \frac{n_r}{N_d} \sum_{n=0}^{N_d-1} \log_2\left(1 + \frac{P\zeta}{n_t} \sum_{i\neq k} \mathbb{E}[||\mathbf{h}_j^{[k]}(n)\hat{\mathbf{v}}^{[i]}(n)||^2] \right),$$

$$\leq \frac{n_r}{N_d} \sum_{n=0}^{N_d-1} \log_2\left(1 + \frac{P\zeta}{n_t}(K-1)\mathbb{E}[||\mathbf{h}_j^{[k]}(n)\hat{\mathbf{v}}^{[i]}(n)||^2] \right)$$

Using the relation between the channel vector and its quantized channel vector in (40), and from the ZF constraint,

$$\hat{\mathbf{h}}_j^{[k]}(n)\hat{\mathbf{v}}^{[i]}(n) = \hat{\mathbf{h}}_{\omega,j}^{[k]}\boldsymbol{\Gamma}(n)\hat{\mathbf{v}}^{[i]}(n) = 0, \quad i \neq k, \quad (41)$$

it follows that

$$\mathbf{h}_j^{[k]}(n)\hat{\mathbf{v}}^{[i]}(n) = \sqrt{a}||\mathbf{h}_{\omega,j}^{[k]}||\mathbf{s}_j\boldsymbol{\Gamma}(n)\hat{\mathbf{v}}^{[i]}(n).$$

It should be emphasized here that for the selective fading channel and as a consequence of the zero forcing constraint in (41), the vector $\boldsymbol{\Gamma}(n)\hat{\mathbf{v}}^{[i]}(n)$ is isotropically distributed in the null-space of $\hat{\mathbf{h}}_{\omega,j}^{[k]}$. Moreover, it can be easily shown by using the value of matrix $\boldsymbol{\Gamma}(n)$ in (15) that the norm of this vector is such that,

$$||\boldsymbol{\Gamma}(n)\hat{\mathbf{v}}^{[i]}(n)||^2 = \sum_{i=0}^{\rho-1} \sigma_i^2 |w_{n,i+1}|^2 = \sigma_{t,n}^2.$$

It can be deduced therefore that the two vectors $\mathbf{s}_j$ and $\frac{1}{\sigma_{t,n}}\boldsymbol{\Gamma}(n)\hat{\mathbf{v}}^{[i]}(n)$ are two unit vector isotropically distributed in the null-space of $\hat{\mathbf{h}}_{\omega,j}^{[k]}$. Then by following the same reasoning as in [6], these two vectors are distributed in the ρn_t - 1 null-space of $\hat{\mathbf{h}}_{\omega,j}^{[k]}$, and the angle between these two vectors $\in \mathbb{C}^{\rho n_t-1}$ is beta distributed with parameters $\beta(1, \rho n_t - 2)$. Hence,

$$\mathbb{E}[||\mathbf{h}_j^{[k]}(n)\hat{\mathbf{v}}^{[i]}(n)||^2] = \sigma_{t,n}^2 \mathbb{E}[\beta(1, \rho n_t - 2)]\mathbb{E}[||\mathbf{h}_{\omega,j}^{[k]}||]\mathbb{E}[a],$$
$$= \frac{1}{\rho n_t - 1}\rho n_t \mathbb{E}[a].$$

It is well known from [6] that the quantization error a corresponds the minimum angle between the channel vector and the 2^B codebooks vector, and therefore it is distributed as the minimum between 2^B beta variables, and

$$\mathbb{E}[a] \leq 2^{-\frac{B}{\rho n_t - 1}}.$$

Then, the rate loss is therefore upper bounded by,

$$\Delta R_k \leq \frac{n_r}{N_d} \sum_{n=0}^{N_d-1} \log_2\left(1 + \frac{\rho(K-1)\zeta}{\rho n_t - 1}\sigma_{t,n}^2 P 2^{-\frac{B}{\rho n_t - 1}} \right),$$

The maximal multiplexing gain can be achieved, if the gap capacity between the full CSIT and the quantized capacity are independent of P. This occurs if the number of bits scale as $(\rho n_t - 1)\log_2 P$, then ΔR_k is constant and independent of P, such that

$$\Delta R_k \leq \frac{n_r}{N_d} \sum_{n=0}^{N_d-1} \log_2\left(1 + \sigma_{t,n}^2 \frac{\rho(K-1)\zeta}{\rho n_t - 1} \right),$$
$$\leq n_r \log_2\left(1 + \sigma_t^2 \frac{\rho(K-1)\zeta}{\rho n_t - 1} \right),$$

where $\sigma_t^2 = \max_{n=0\ldots N_d-1}\sigma_{t,n}^2$. Consequently,

$$\hat{R} = R - \Delta R,$$
$$\geq R - c,$$
$$\geq Kn_r \log_2 P - Kn_r \log_2\left(1 + \sigma_t^2 \frac{\rho(K-1)\zeta}{\rho n_t - 1} \right),$$

and therefore the maximal multiplexing gain can be achieved, but with a constant capacity gap.

Appendix C
Proof of Theorem 2

The relationship between the quantized channel matrix can be deduced using Lemma 1 in [7]. By doing similar manipulations as in Theorem 1, it can be shown that in this case the rate loss incurred by the quantization can be bounded such that,

$$\Delta R_k \leq n_r \log_2\left(1 + P(K-1)\sigma_t^2 \frac{\rho}{\rho n_t - n_r}D_s \right), \quad (42)$$

where D_s is the distortion error rate and is such that,

$$D_s \leq \bar{D} \quad (43)$$

where $\bar{D}$ is equivalent for large B to

$$\bar{D} \xrightarrow{B\rightarrow\infty} C\, 2^{-\frac{B}{n_r(\rho n_t - n_r)}}. \quad (44)$$

to and C is a constant independent of B defined as in (29). It follows that,

$$\Delta R_k \leq n_r \log_2 \left(1 + P\sigma_t^2 \frac{\rho(K-1)}{\rho n_t - n_r} C \, 2^{-\frac{B}{n_r(\rho n_t - n_r)}} \right).$$

It can be easily deduced that, ΔR is independent of P, if and only if B scales as $n_r(\rho n_t - n_r) \log_2 P$, then

$$\Delta R \leq n_r \log_2 \left(1 + \sigma_t^2 \frac{\rho(K-1)}{\rho n_t - n_r} C \right) = c,$$

where c is a constant independent of P. Consequently,

$$\begin{aligned}
\hat{R} &= R - \Delta R, \\
&\geq R - c, \\
&\geq K n_r \log_2 P - K n_r \log_2 \left(1 + \sigma_t^2 \frac{\rho(K-1)}{\rho n_t - n_r} C \right).
\end{aligned}$$

The maximal multiplexing gain can be therefore achieved, but with a constant capacity gap.

Appendix D
Proof of Lemma 2
Let $a_p = \sin^2(\tilde{\mathbf{h}}_\omega^{[k]}, \hat{\mathbf{h}}_\omega^{[k]})$ denote the quantization error between the channel vector and its quantized version. Then,

$$\begin{aligned}
a_p &= 1 - \langle \tilde{\mathbf{h}}_\omega^{[k]}, \hat{\mathbf{h}}_\omega^{[k]} \rangle^2 \\
&= \left(1 - \langle \tilde{\mathbf{h}}_\omega^{[k]}, \hat{\mathbf{h}}_\omega^{[k]} \rangle \right) \left(1 + \langle \tilde{\mathbf{h}}_\omega^{[k]}, \hat{\mathbf{h}}_\omega^{[k]} \rangle \right)
\end{aligned}$$

In order to find an upper bound on a_p, we propose to find separately upper bounds on $1 - \langle \tilde{\mathbf{h}}_\omega^{[k]}, \hat{\mathbf{h}}_\omega^{[k]} \rangle$ and $1 + \langle \tilde{\mathbf{h}}_\omega^{[k]}, \hat{\mathbf{h}}_\omega^{[k]} \rangle$. Using (33), the scalar product between this two vectors can be related to the partitioned vectors as

$$\langle \tilde{\mathbf{h}}_\omega^{[k]}, \hat{\mathbf{h}}_\omega^{[k]} \rangle = \sum_{i=0}^{\rho-1} \frac{\|\mathbf{h}_\omega^{[k]}[i]\| \, \|\mathbf{h}_{\omega,a}^{[k]}[i]\|}{\|\mathbf{h}_\omega^{[k]}\| \, \|\mathbf{h}_\omega^{[k]}\|} \sqrt{1 - a_i},$$

where

$$a_i = \sin(\mathbf{h}_\omega^{[k]}[i], \hat{\mathbf{h}}_\omega^{[k]}[i]).$$

As $\sqrt{1 - a_i} \leq 1$, then

$$1 + \langle \tilde{\mathbf{h}}_\omega^{[k]}, \hat{\mathbf{h}}_\omega^{[k]} \rangle \leq 1 + \sum_{i=0}^{\rho-1} \frac{\|\mathbf{h}_\omega^{[k]}[i]\| \quad \|\mathbf{h}_{\omega,a}^{[k]}[i]\|}{\|\mathbf{h}_\omega^{[k]}\| \quad \|\mathbf{h}_{\omega,a}^{[k]}\|}.$$

Let $\mathbf{x} = [\mathbf{h}_\omega^{[k]}[0]...\mathbf{h}_\omega^{[k]}[\rho-1]]$ and $\mathbf{y} = [\mathbf{h}_{\omega,a}^{[k]}[0]...\mathbf{h}_\omega^{[k]}[\rho-1]]$ then,

$$1 + \langle \tilde{\mathbf{h}}_\omega^{[k]}, \hat{\mathbf{h}}_\omega^{[k]} \rangle \leq 1 + \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\| \, \|\mathbf{y}\|}.$$

The Cauchy-Schwartz inequality,

$$\langle \mathbf{x}, \mathbf{y} \rangle \leq \|\mathbf{x}\| \quad \|\mathbf{y}\|$$

implies that,

$$1 + \langle \tilde{\mathbf{h}}_\omega^{[k]}, \hat{\mathbf{h}}_\omega^{[k]} \rangle \leq 2.$$

Next, by noticing that $\sqrt{1 - a_i} \geq 1 - a_i$, we get the upper bound on

$$1 - \langle \tilde{\mathbf{h}}_\omega^{[k]}, \hat{\mathbf{h}}_\omega^{[k]} \rangle \leq \varepsilon_a + \varepsilon_d$$

where,

$$\varepsilon_a = 1 - \frac{1}{\|\mathbf{h}_\omega^{[k]}\| \, \|\mathbf{h}_{\omega,a}^{[k]}\|} \sum_{i=0}^{\rho-1} \|\mathbf{h}_\omega^{[k]}[i]\| \, \|\mathbf{h}_{\omega,a}^{[k]}[i]\|,$$

and

$$\varepsilon_d = \frac{1}{\|\mathbf{h}_\omega^{[k]}\| \, \|\mathbf{h}_{\omega,a}^{[k]}\|} \sum_{i=0}^{\rho-1} \|\mathbf{h}_\omega^{[k]}[i]\| \, \|\mathbf{h}_{\omega,a}^{[k]}[i]\| \, a_i.$$

At high SNR, the analog estimation $\|\mathbf{h}_{\omega,a}^{[k]}\|$ in (32) is approximately equal to $\|\mathbf{h}_\omega^{[k]}\|$. This implies that,

$$\varepsilon_a \sim 0$$

$$\varepsilon_d \sim \frac{1}{\|\mathbf{h}_\omega^{[k]}\|^2} \sum_{i=0}^{\rho-1} \|\mathbf{h}_\omega^{[k]}[i]\|^2 a_i.$$

The expected quantization error is therefore,

$$\mathbb{E}[\varepsilon_d] = \sum_{i=0}^{\rho-1} \mathbb{E}\left[\frac{\|\mathbf{h}_\omega^{[k]}[i]\|^2}{\|\mathbf{h}_\omega^{[k]}\|^2} \right] \mathbb{E}[a_i]$$

The random variable $\frac{\|\mathbf{h}_\omega^{[k]}[i]\|^2}{\|\mathbf{h}_\omega^{[k]}\|^2}$ is beta distributed with parameters $a = n_t$ and $b = (\rho - 1)n_t$. The expectation of this variable is $a/(a + b) = 1/\rho$. The expectation of quantization error incurred by the RVQ is upper bounded by $2^{-\frac{B}{n_t-1}}$. Then, at high SNR, the quantization error is upper bounded by,

$$\mathbb{E}[\varepsilon_d] \leq 2^{-\frac{B}{n_t - 1}}.$$

This completes the proof.

Appendix E
Proof of Theorem 3
For each receive vector, the number of feedback bits required to estimated the partitioned vectors norms is at

most equal to the capacity of the uplink AWGN channel $N_{f,a} = \beta\rho \log_2 P$.

The upper bound can be derived following the same steps as Appendix B. It turns out that at high SNR, the gap rate is upper bounded by,

$$\Delta R_k \leq n_r \log_2\left(1 + 2\sigma_t^2 \frac{\rho(K-1)\zeta}{\rho n_t - 1} \sin^2(\tilde{\mathbf{h}}_\omega^{[k]}, \hat{\mathbf{h}}_\omega^{[k]})P\right).$$

Using Lemma 2, the expected gap rate is upper bounded by,

$$\Delta R_k \leq n_r \log_2\left(1 + 2\sigma_t^2 \frac{\rho(K-1)\zeta}{\rho n_t - 1} 2^{-\frac{B}{n_t-1}}P\right). \quad (45)$$

The maximal multiplexing gain can be achieved, if the gap capacity between the full CSIT and the quantized capacity are independent of P. This occurs if the number of bits required to quantize each part of the vector scales as $(n_t - 1)\log_2 P$. In total there is ρ parts in each vector, and the total number of feedback bits is therefore $\rho(n_t - 1)\log_2 P + \beta\rho \log_2 P$.

Endnotes

[a]Please note that the considered WH set with parameters T and F is a Riez sequence and is constructed as a dual of a WH frame characterized by grid parameters $\tilde{T} = \frac{1}{T}$, and $\tilde{F} = \frac{1}{F}$ such that $\tilde{T}\tilde{F} \leq 1$. [b]This assumption is commonly used for when considering non-coherent setting as defined in [15] and references therein. [c]In practical system, this assumption can be feasible as it requires only the feedback of N different values of $R_{\mathbb{H}}[k, n]$. The two-level Toeplitz matrix can be constructed according to (9). [d]This corresponds to parameters $h_{\omega, i, j}[s]$ in (16). [e]For the perfect estimation, we assume that the I/O relationship in (11) is noisy-free. [f]This inequality can be easily verified by noticing that $\rho \geq 1$, $K \geq 1$ and $n_t \geq Kn_r$. This implies that $\rho n_t \geq n_t \geq Kn_r \geq n_r$. [g]For simplicity, we only consider the data transmission over the first 100 TF slots that follow the training phase to plot the numerical results. [h]We note here that the implementation of the GRF with BD is not possible with Matlab when dealing with a 72×2 channel using the classical numerical generation as in [7].

Author details
[1]Institut Supérieur d'Electronique de Paris, 75006 Paris, France [2]Thales Communication, Paris, France [3]Télécom ParisTech, 75013 Paris, France

References
1. H Weingarten, Y Steinberg, S Shamai, The capacity region of the gaussian multiple-input multiple-output broadcast channel. IEEE Trans Inf Theory, 52, 3936–3964 (2006)
2. CB Peel, BM Hochwald, AL Swindlehurst, A vector perturbation technique for near capacity multiantenna multiuser communication Part I: channel inversion and regularization. IEEE Trans Inf Theory. 53(195-202), 4787–4792 (2005)
3. L-U Choi, RD Murch, A transmit preprocessing technique for multiuser MIMO systems using a decomposition approach. IEEE Trans Wirel Commun. 3, 20–24 (2004). doi:10.1109/TWC.2003.821148
4. CB Peel, BM Hochwald, AL Swindlehurst, A vector perturbation technique for near capacity multiantenna multiuser communication Part II: perturbation. IEEE Trans Inf Theory. 53(3), 537–544 (2005)
5. J Lee, N Jindal, High SNR analysis for MIMO broadcast channels: dirty paper coding vs. linear precoding. IEEE Trans Inf Theory. 53(12), 4787–4792 (2007)
6. N Jindal, MIMO broadcast channels with finite rate feedback. IEEE Trans Inf Theory. 52(11), 5045–5058 (2006)
7. N Ravindran, N Jindal, Limited feedback-based block diagonalisation for the MIMO broadcast channels. IEEE J Sel Areas Commun. 26(8), 1473–1482 (2008)
8. G Caire, N Jindal, M Kobayashi, N Ravindran, Multiuser MIMO achievable rates with downlink training and channel state feedback. IEEE Trans Inf Theory. 56(6), 2845–2866 (2010)
9. H Shirani-Mehr, G Caire, Channel state feedback schemes for multiuser MIMO-OFDM downlink. IEEE Trans Commun. 57(9), 2713–2723 (2009)
10. J Thukral, Spatial multiplexing in multiuser networks with limited feedback. Ph.D. thesis, ETH Zurich. (2009)
11. M Trivellato, S Tomasin, N Benvenuto, On channel quantization and feedback strategies for multiuser MIMO-OFDM downlink systems. IEEE Trans Commun. 57(9), 2645–2654 (2009)
12. D Love, R Heath, Limited feedback diversity techniques for correlated channels. IEEE Trans Veh Technol. 55(2), 718–722 (2006). doi:10.1109/TVT.2005.863346
13. D Love, R Heath, V Lau, D Gesbert, B Rao, M Andrews, An overview of limited feedback in wireless communication systems. IEEE J Sel Areas Commun. 26(8), 1341–1365 (2008)
14. F Hlawatsch, G Matz, Wireless Communications Over Rapidly Time-Varying Channels (Academic Press, 2010)
15. G Durisi, UG Schuster, H Bölcskei, SS Shitz, Noncoherent capacity of underspread fading channels. IEEE Trans Inf Theory. 56(1), 367–395 (2010)
16. W Kozek, Matched weyl heisemberg expansions of nonstationary environments. Ph.D. dissertation, Vienna University of Technology, March 1997
17. T Kailath, Measurements on time-variant communications channels. IRE Trans Inf Theory. 8(5), 229–236 (1962). doi:10.1109/TIT.1962.1057748
18. PA Voois, A theorem on the asymptotical distribution of toeplitz-block-toeplitz matrices. IEEE Trans Signal Process. 44(7), 1837–1841 (1996). doi:10.1109/78.510633
19. D Tse, P Viswanath, Fundamentals of wireless communications (Cambridge University Press, 2005)