RESEARCH

Open Access

Decoupled signal detection for the uplink of massive MIMO in 5G heterogeneous networks

Leonel Arévalo[1*], Rodrigo C. de Lamare[1], Martin Haardt[2] and Raimundo Sampaio-Neto[1]

Parts of this paper have been published at CAMSAP 2015

Abstract

Massive multiple-input multiple-output (MIMO) systems are strong candidates for future fifth-generation (5G) heterogeneous cellular networks. For 5G, a network densification with a high number of different classes of users and data service requirements is expected. Such a large number of connected devices needs to be separated in order to allow the detection of the transmitted signals according to different data requirements. In this paper, a decoupled signal detection (DSD) technique which allows the separation of the uplink signals, for each user class, at the base station (BS) is proposed for massive MIMO systems. A mathematical signal model for massive MIMO systems with centralized and distributed antennas in heterogeneous networks is also developed. The performance of the proposed algorithm is evaluated and compared with existing detection schemes in a realistic scenario with distributed antennas. A sum-rate analysis and a computational cost study for DSD are also presented. Simulation results show an excellent performance of the proposed algorithm when combined with linear and successive interference cancellation detection techniques.

Keywords: Massive multiple-input multiple-output (MIMO) systems, Heterogeneous networks, Fifth-generation (5G) cellular networks, Decoupled signal detection (DSD)

1 Introduction

Large-scale multiple-input multiple-output (MIMO) systems, also known as massive MIMO, are a promising technology which use a large number of antennas to serve a high number of user terminals at the same time without requiring extra bandwidth resources [1–4]. This new greater scale version of traditional MIMO systems, where a restricted number of antennas is used, is designed to exploit the benefits of extra degrees of freedom obtained by the use of more antennas [5]. Massive MIMO can increase the spectral efficiency 10 times or more when compared with its predecessor [6]. In the Long Term Evolution (LTE) standard [7], which uses traditional MIMO systems, with as many as eight antenna ports at the base station (BS), and operates in the frequency-division

duplex (FDD) mode, the users estimate the channel response and feed it back to the BS. For massive MIMO with hundreds of antenna elements, this might not be feasible, due to the large number of channel coefficients that each user needs to estimate being proportional to the number of the antennas at the BS. In this paper we focus on the uplink, the reason is that the most natural transmission mode to operate in massive MIMO is the time-division duplexing (TDD) mode, where a reciprocity between the uplink and downlink channels can be obtained, if we use appropriate calibration techniques to combat the distortions induced by hardware imperfections, since the base station can offer more processing resources aimed at estimating the channel between users' terminals and the BS. On the downlink, it is possible to use different precoding schemes to mitigate the interference for the received signals at the mobile users. Such precoding schemes rely on the channels estimated on the uplink.

*Correspondence: leonel_arevalo@cetuc.puc-rio.br
[1]Center for Studies in Telecommunications (CETUC), Pontifical Catholic University of Rio de Janeiro (PUC-Rio), Rio de Janeiro, Brazil
Full list of author information is available at the end of the article

One of the main research challenges of massive MIMO is to develop computationally simple ways to process the large number of signals received at the BS. The interference between antennas and users, propagation effects such as correlation, path loss and shadowing, thermal noise and signal degradation due to the hardware imperfections need to be suppressed. Linear detection techniques such as maximum radio combining (MRC) and zero forcing (ZF) are a good option in terms of computational complexity; however, their performance is not compatible with the growing demand for high data rates. The performance of linear detectors can be improved using some nonlinear sub-optimal detector based on successive interference cancellation (SIC) [8], e.g., multibranch SIC (MB-SIC) [9, 10] and multifeedback SIC (MF-SIC) [11]. However, SIC-based detectors have a considerable computational cost for high-dimensional systems. If all signals from all active users are coupled in the detection process, the BS could spend unnecessary processing resources since some types of users might not require a very high performance.

In the next generation of wireless communication systems [12], it is expected that a large number of users with different configurations and requirements are connected to the network. Therefore, it is necessary to design heterogeneous networks capable of interconnecting the different user types with each other [13]. The received signals from this large number of connected devices such as metering equipment, sensors, environmental monitoring devices, health care gadgets, security management products, smart grid components, smart phones, and tablets need to be separated in order to detect the transmitted information according to their different data requirements. In this context, distributed antenna systems (DAS) with massive MIMO are a promising alternative for the 5G cellular architecture [14], where the BS will be equipped with a large number of antennas and some remote antenna arrays or radio heads will be distributed around the cell and connected to the BS via optical fiber. The signals associated with different remote antenna arrays are processed at the BS. DAS have low path loss effects and improve the coverage and the spectral efficiency [15, 16]. The energy consumption of users is reduced and the transmission quality is improved due to the shorter distances between users and some remote antenna arrays. For this vision of 5G wireless networks, which includes a combination of massive MIMO, heterogeneous networks, and distributed antenna systems, efficient signal processing techniques at the BS are necessary.

In this paper, we propose an algorithm for the uplink of massive MIMO systems to separate the combined received signal of all users at the BS into independent signals for each user class. The proposed decoupled signal detection (DSD) applies a decomposition into multiple independent single user class signals, where all users in a class have the same data requirements and a common complex modulation. Assuming that the channel state information (CSI) was previously estimated, DSD employs a common channel inversion and QR decomposition to decouple the received signal. Applying the proposed algorithm, the computational cost of the signal processing is reduced and it is possible to have flexibility on the detection procedures at the BS. A signal model for heterogeneous networks with different classes of users and an arbitrary configuration of centralized antenna systems (CAS) and DAS is also introduced in this paper. A sum rate analysis and a computational complexity study for the proposed DSD are presented. The performance of the proposed scheme is evaluated in a realistic scenario with propagation effects and compared with existing approaches.

The remainder of this paper is organized as follows. In Section 2 the proposed massive MIMO signal model is presented. The proposed DSD scheme is presented in Section 3. The sum rate analysis for the DSD scheme is described in Section 4. Section 5 presents simulation results. Section 6 gives the conclusions.

2　Massive MIMO signal model

5G cellular networks will be able to support a large number and diverse classes of devices, i.e., they will be heterogeneous, where a single macro cell may need to support 10,000 or more low-rate devices, along with its traditional high-rate mobile users [17]. This will require joint user classification according to their data rate requirements. Thus, a user in a high definition video call may have higher data rate requirements compared to a mobile user transmitting voice. Then we can classify them in different classes of users. In this section, a signal model for heterogeneous networks with different classes of users, as depicted in Fig. 1, and an arbitrary configuration of CAS and DAS is presented. We consider the uplink channel scenario of a massive MIMO system with N different classes of active users transmitting simultaneously signals to one base station (BS) equipped with D remote antenna arrays distributed around the cell and N_B receive antennas at the BS. Each remote array of antennas has Q antennas linked to the BS via wired links. Therefore, the total number of receive antennas is $N_r = N_B + DQ$. The choice of N_B, D, and Q is made based on the features of the network and type of application scenario. For example, suppose that we have a city with a high density of users or devices in the center of a cell and sparsely distributed users or devices in the remaining part of the cell. In this case, we could use a number of centralized antennas to deal with the high density of users and distributed antennas to serve the remaining devices. The cardinality of the nth user class $|C_n|$ represents the number of users of
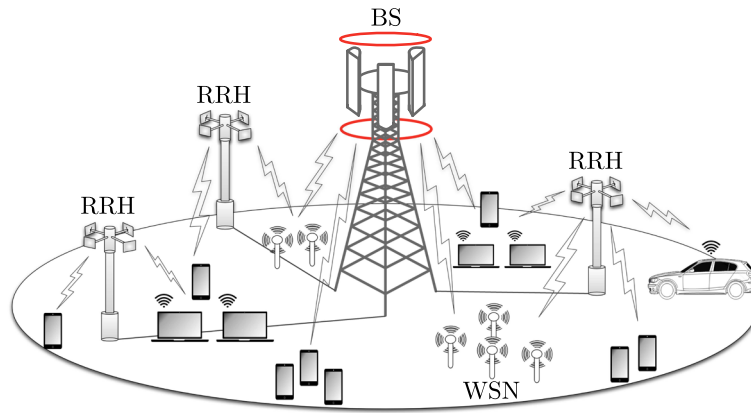
Fig. 1 Heterogeneous wireless network: base station (BS), remote radio head (RRH), and wireless sensor networks (WSN)

the class n. The total number of active users is given by $K = \sum_{n=1}^{N} | C_n |$. The kth user in the nth user class transmits data divided into $N_{t_{k,n}}$ sub-streams through $N_{t_{k,n}}$ antennas, where $N_r \geqslant N_t = \sum_{n=1}^{N} \sum_{k=1}^{|C_n|} N_{t_{k,n}}$ and N_t is the total number of transmit antennas. The received signal vector at the BS from all active users in all user classes is given by

$$\mathbf{y} = \sum_{n=1}^{N} \sum_{k=1}^{|C_n|} \mathbf{\Upsilon}_{k,n} \bar{\mathbf{H}}_{k,n} \mathbf{s}_{k,n} + \mathbf{n}, \qquad (1)$$

where $\mathbf{s}_{k,n}$ is the $N_{t_{k,n}} \times 1$ transmitted signal vector, by the kth user of the nth user class, at one time slot taken from a complex constellation, denoted by $\mathcal{A} = \{a_1, a_2, \ldots, a_O\}$. Each symbol has M bits and $O = 2^M$. The vector $\mathbf{n}$ is an $N_r \times 1$ zero mean complex circular symmetric Gaussian noise vector with covariance matrix $\mathbf{K_n} = \mathbb{E}\left[\mathbf{nn}^{\mathcal{H}}\right] = \sigma_n^2 \mathbf{I}$. Moreover, $\bar{\mathbf{H}}_{k,n}$ is the $N_r \times N_{t_{k,n}}$ channel matrix of the kth user in the class n with elements $\bar{h}_{i,j}^{(k,n)}$ corresponding to the complex channel gain from the jth transmit antenna of the kth user to the ith receive antenna. For the antenna elements located at the BS and at each remote radio head, the $D + 1$ sub-matrices of $\bar{\mathbf{H}}_{k,n} = \left[\left(\bar{\mathbf{H}}_{k,n}^{(1)}\right)^T, \left(\bar{\mathbf{H}}_{k,n}^{(2)}\right)^T, \ldots, \left(\bar{\mathbf{H}}_{k,n}^{(D+1)}\right)^T\right]^T$ can be modeled using the Kronecker channel model [18], expressed by

$$\bar{\mathbf{H}}_{k,n}^{(j)} = \left(\mathbf{R}_r^{(j)}\right)^{1/2} \mathbf{G}_{k,n}^{(j)} \left(\mathbf{R}_{t_{k,n}}\right)^{1/2} \qquad (2)$$

where $\mathbf{G}_{k,n}^{(j)}$ has complex channel gains between the kth user and the jth radio head, obtained from an independent and identically distributed random fading model whose coefficients are complex Gaussian random variables with zero mean and unit variance. $\mathbf{R}_r^{(j)}$ and $\mathbf{R}_{t_{k,n}}$ denote the receive correlation matrix of the jth radio head and the transmit correlation matrix, respectively. The components of the correlation matrices $\mathbf{R}_r^{(j)}$ and $\mathbf{R}_{t_{k,n}}$

are modeled as a variation of the model described in equations (3)-(5) of reference [18]:

$$\mathbf{R} = \begin{bmatrix} 1 & \rho & \rho^4 & \cdots & \rho^{(N_a-1)^2} \\ \rho & 1 & \rho & \cdots & \vdots \\ \rho^4 & \rho & 1 & \vdots & \rho^4 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho^{(N_a-1)^2} & \cdots & \rho^4 & \rho & 1 \end{bmatrix}, \qquad (3)$$

where N_a is the number of antennas and ρ is the correlation coefficient of neighboring antennas ($\rho = \rho_{t_x}$ for the transmit antennas and $\rho = \rho_{r_x}$ for the receive antennas), i.e., a decaying of the correlation index with antenna separation faster than exponential was adopted. Note that $\rho = 0$ represents an uncorrelated scenario and $\rho = 1$ implies a fully correlated scenario. The $N_r \times N_r$ diagonal matrix $\mathbf{\Upsilon}_{k,n}$ represents the large-scale propagation effects for the user k of the user class n, such as path loss and shadowing, given by

$$\mathbf{\Upsilon}_{k,n} = \text{diag}\Big(\underbrace{\gamma_{k,n}^1 \cdots \gamma_{k,n}^1}_{N_B} \underbrace{\gamma_{k,n}^2 \cdots \gamma_{k,n}^2}_{Q} \cdots \underbrace{\gamma_{k,n}^{D+1} \cdots \gamma_{k,n}^{D+1}}_{Q} \Big), \qquad (4)$$

where the parameters $\gamma_{k,n}^j$ denote the large-scale propagation effects from the kth user to the jth radio head described by

$$\gamma_{k,n}^j = \alpha_{k,n}^j \beta_{k,n}^j, \qquad j = 1, \ldots, D+1. \qquad (5)$$

Here $\alpha_{k,n}^j$ is the distance based path-loss between each user and the radio heads which is calculated by

$$\alpha_{k,n}^j = \sqrt{\frac{L_{k,n}^j}{\left(d_{k,n}^j\right)^\tau}}, \qquad (6)$$

where $L_{k,n}^{j}$ is the power path loss of the link associated with the user and the jth radio head, $d_{k,n}^{j}$ is the relative distance between this user and the jth radio head, τ is the path loss exponent chosen between 2 and 4 depending on the environment. The log normal random variable $\beta_{k,n}^{j}$ which represents the shadowing between user k and the receiver is given by

$$\beta_{k,n}^{j} = 10^{\frac{\mu_{k,s}\vartheta_{k,s}^{j}}{10}}, \tag{7}$$

where $\mu_{k,s}$ is the shadowing spread in decibels and $\vartheta_{k,s}^{j}$ corresponds to a real-valued Gaussian random variable with zero mean and unit variance. Since the $N_r \times N_{t_{k,n}}$ composite channel matrix includes large-scale and small-scale fading effects, it can be denoted as $\mathbf{H}_{k,n} = \boldsymbol{\Upsilon}_{k,n}\bar{\mathbf{H}}_{k,n}$, and the expression in (1) can be written as

$$\mathbf{y} = \sum_{n=1}^{N} \mathbf{H}_n \mathbf{s}_n + \mathbf{n}, \tag{8}$$

where $\mathbf{H}_n = [\mathbf{H}_{1,n} \quad \mathbf{H}_{2,n} \dots \mathbf{H}_{|C_n|,n}]$ and $\mathbf{s}_n = \left[\mathbf{s}_{1,n}^{T} \; \mathbf{s}_{2,n}^{T} \dots \mathbf{s}_{|C_n|,n}^{T}\right]^{T}$ represent the channel matrix and the transmitted symbol vector of all users in the class n, respectively. The received signal vector can be expressed more conveniently as

$$\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{n}, \tag{9}$$

where $\mathbf{H} = [\mathbf{H}_1 \; \mathbf{H}_2 \dots \mathbf{H}_N]$ and $\mathbf{s} = \left[\mathbf{s}_1^{T} \; \mathbf{s}_2^{T} \dots \mathbf{s}_N^{T}\right]^{T}$. The symbol vector $\mathbf{s}$ of all N user classes has zero mean and a covariance matrix $\mathbf{K_s} = \mathbb{E}\left[\mathbf{s}\mathbf{s}^{\mathcal{H}}\right] = \mathrm{diag}(\mathbf{p})$, where the elements of the vector $\mathbf{p}$ are the signal power of each transmit antenna. To maintain a notational simplicity in the subsequent analysis, we assume that all antenna elements at the users transmit with the same average transmitted power σ_s^2, i.e., $\mathbf{K_s} = \sigma_s^2\mathbf{I}$. We assume that the channel matrix $\mathbf{H}$ was previously estimated at the BS. From (9) we can see that the signals arrive coupled at the BS. If we want to use different detection procedures for each user class according to its data requirements, we have to separate the received signal vector $\mathbf{y}$ into independent received signals for each user class. For the system model presented in this work, when the number of remote radio heads is set to zero, i.e., $D = 0$, the DAS architecture is reduced to the CAS scheme with $N_r = N_B$.

3 Decoupled signal detection

As presented in Section 2, in heterogeneous networks different classes of users send parallel data streams, through the massive MIMO channel operating with distributed antennas, which arrive superposed at the BS. In this 5G context, we need to separate the data streams for each category of users efficiently. In this section, we describe the proposed decoupled signal detection (DSD) which

allows us to separate the received signal of the nth user class from the others. To this end, we consider that the process of authentication, identification, and channel estimation was already made, i.e, the BS is able to identify the users by classes according to their data requirements. Similar approaches to the proposed algorithm have been proposed for the downlink, such as block diagonalization (BD)-based techniques [19–22]. However, unlike prior work in which downlink BD is used, for the proposed DSD scheme, it is not necessary to use any precoding at the transmit side. The receiver only needs to know the channels between users and receive antennas. Moreover, the concept of separating the users with respect to the classes in heterogeneous networks according to its requirements is a new approach. The first steps to construct the concept proposed in this paper were presented in [23, 24].

The received signal vector (8) can be written as:

$$\mathbf{y} = \mathbf{H}_n \mathbf{s}_n + \sum_{m=1,m\neq n}^{N} \mathbf{H}_m \mathbf{s}_m + \mathbf{n}, \tag{10}$$

where $\mathbf{H}_n$ and $\mathbf{s}_n$ are the channel matrix and the transmitted symbol vector for the nth user class, respectively. From (10), we can see that the nth user class has inter-user class interference.

3.1 Proposed decoupling strategy

To remove the presence of the other classes of users in the detection procedure for the nth user class, we can employ a linear operation to project the received signal vector $\mathbf{y}$ onto the subspace orthogonal or almost orthogonal to the subspace generated by the signals of the interfering classes. In DSD, a matrix $\mathbf{A}_n$ is calculated employing a channel inversion method and a QR decomposition [25, 26], in order to decouple the nth user's class received signal from other user's class signals. To compute $\mathbf{A}_n$, we construct the matrix $\tilde{\mathbf{H}}_n$ excluding the channel matrix of the nth user class in the following form:

$$\tilde{\mathbf{H}}_n = [\mathbf{H}_1 \dots \mathbf{H}_{n-1} \; \mathbf{H}_{n+1} \dots \mathbf{H}_N], \tag{11}$$

where $\tilde{\mathbf{H}}_n \in \mathbb{C}^{N_r \times (N_t - N_{t_n})}$ and $N_{t_n} = \sum_{k=1}^{|C_n|} N_{t_{k,n}}$ is the number of transmit antennas in the nth user class. After that, the objective is to obtain a matrix $\mathbf{A}_n$ that satisfies the following condition:

$$\mathbf{A}_n \tilde{\mathbf{H}}_n = \mathbf{0}, \quad \forall n \in (1 \dots N). \tag{12}$$

To compute $\mathbf{A}_n$, DSD first computes the MMSE channel inversion of the combined channel matrix $\mathbf{H}$ given by

$$\begin{aligned} \mathbf{H}^{\dagger} &= \mathbf{H}^{\mathcal{H}} \left(\mathbf{H}\mathbf{H}^{\mathcal{H}} + \sigma^2\mathbf{I}\right)^{-1} \\ &= \left[(\ddot{\mathbf{H}}_1)^{T} \; (\ddot{\mathbf{H}}_2)^{T} \; \dots \; (\ddot{\mathbf{H}}_N)^{T}\right]^{T} \end{aligned} \tag{13}$$

where $\sigma^2 = \sigma_n^2/\sigma_s^2$, $\mathbf{H}^\dagger \in \mathbb{C}^{N_t \times N_r}$ and $\ddot{\mathbf{H}}_n \in \mathbb{C}^{N_{t_n} \times N_r}$. Then, the matrix $\ddot{\mathbf{H}}_n$ is approximately in the null space of $\tilde{\mathbf{H}}_n$, that is

$$\ddot{\mathbf{H}}_n \tilde{\mathbf{H}}_n \approx \mathbf{0}, \quad \forall n \in (1 \ldots N). \tag{14}$$

To decouple the nth user group from the other user groups, we employ a QR decomposition as described by

$$\ddot{\mathbf{H}}_n = \mathbf{R}_n \mathbf{Q}_n, \quad \forall n \in (1 \ldots N), \tag{15}$$

where $\mathbf{R}_n \in \mathbb{C}^{N_{t_n} \times N_{t_n}}$ is an upper triangular matrix and $\mathbf{Q}_n \in \mathbb{C}^{N_{t_n} \times N_r}$ is a matrix with orthogonal rows and composed by approximately orthonormal basis vectors of the left null space of $\tilde{\mathbf{H}}_n$. Then we have

$$\mathbf{Q}_n \tilde{\mathbf{H}}_n \approx \mathbf{0}. \tag{16}$$

From (16), we can see that $\mathbf{Q}_n$ is a good approximation for $\mathbf{A}_n$ in (12). Using $\mathbf{A}_n = \mathbf{Q}_n$ as a linear combination with the received signal vector in (10), we have

$$\begin{aligned} \mathbf{y}_n &= \mathbf{A}_n \mathbf{y} \\ &= \mathbf{Q}_n \mathbf{H}_n \mathbf{s}_n + \mathbf{Q}_n \sum_{\substack{m=1 \\ m \neq n}}^{N} \mathbf{H}_m \mathbf{s}_m + \mathbf{Q}_n \mathbf{n}, \end{aligned} \tag{17}$$

where $\mathbf{y}_n \in \mathbb{C}^{N_{t_n} \times 1}$ is the equivalent received signal vector for the user class n and the term $\mathbf{Q}_n \sum_{\substack{m=1 \\ m \neq n}}^{N} \mathbf{H}_m \mathbf{s}_m \approx \mathbf{0}$ represents the residual inter-user class interference. Then, we can transform the received signal vector into parallel single-user class signals as described by

$$\mathbf{y}_n = \check{\mathbf{H}}_n \mathbf{s}_n + \mathbf{n}_n, \quad \forall n \in (1 \ldots N), \tag{18}$$

where $\check{\mathbf{H}}_n = \mathbf{Q}_n \mathbf{H}_n \in \mathbb{C}^{N_{t_n} \times N_{t_n}}$ is the equivalent channel matrix of the nth user class after DSD and $\mathbf{n}_n = \mathbf{Q}_n \sum_{\substack{m=1 \\ m \neq n}}^{N} \mathbf{H}_m \mathbf{s}_m + \mathbf{Q}_n \mathbf{n} \in \mathbb{C}^{N_{t_n} \times 1}$ is the equivalent noise vector. Note that $\mathbf{H}^\dagger$ in (13) could have been provided by the pseudo-inverse of $\mathbf{H}$. This option satisfies the zero interference constraint in (12); however, it does result in a noise enhancement effect and has a restriction in terms of the dimension, i.e., $N_r \geq N_t$. The proposed strategy can be used even when $N_r < N_t$ and provides a balance between the inter-user class interference and the noise effects since the noise is taken into account in the computation of (13).

Another option to compute a basis for the left null space of $\tilde{\mathbf{H}}_n$ is performing the SVD transformation $\tilde{\mathbf{H}}_n = \tilde{\mathbf{U}}_n \tilde{\Sigma}_n \tilde{\mathbf{V}}_n^{\mathcal{H}}$, where $\tilde{\Sigma}_n \in \mathbb{C}^{N_r \times (N_t - N_{t_n})}$ is a rectangular diagonal matrix with the singular values of $\tilde{\mathbf{H}}_n$ on the diagonal, $\tilde{\mathbf{U}}_n \in \mathbb{C}^{N_r \times N_r}$ and $\tilde{\mathbf{V}}_n^{\mathcal{H}} \in \mathbb{C}^{(N_t - N_{t_n}) \times (N_t - N_{t_n})}$ are unitary matrices. If r_n is the rank of $\tilde{\mathbf{H}}_n$ that corresponds to the number of non-zero singular values, i.e., $r_n = \text{rank}(\tilde{\mathbf{H}}_n) \leq N_t - N_{t_n}$, the SVD can be expressed equivalently as:

$$\tilde{\mathbf{H}}_n = \begin{bmatrix} \tilde{\mathbf{U}}_{1,n} & \tilde{\mathbf{U}}_{0,n} \end{bmatrix} \tilde{\Sigma}_n \begin{bmatrix} \tilde{\mathbf{V}}_{1,n} & \tilde{\mathbf{V}}_{0,n} \end{bmatrix}^{\mathcal{H}}, \tag{19}$$

where $\tilde{\mathbf{U}}_{1,n} \in \mathbb{C}^{N_r \times r_n}$ consists of the first r_n left singular vectors, $\tilde{\mathbf{U}}_{0,n} \in \mathbb{C}^{N_r \times (N_r - r_n)}$ holds the last $N_r - r_n$ left singular vectors, $\tilde{\mathbf{V}}_{1,n} \in \mathbb{C}^{(N_t - N_{t_n}) \times r_n}$ consists of the first r_n singular vectors and $\tilde{\mathbf{V}}_{0,n} \in \mathbb{C}^{(N_t - N_{t_n}) \times (N_t - N_{t_n} - r_n)}$ hold the last $N_t - N_{t_n} - r_n$ singular vectors. Thus $\tilde{\mathbf{U}}_{0,n}$ and $\tilde{\mathbf{V}}_{0,n}^{\mathcal{H}}$ form an orthogonal basis for the left null space and the null space of $\tilde{\mathbf{H}}_n$, respectively. Then, the alternative solution for (12) could be:

$$\mathbf{A}_n = \tilde{\mathbf{U}}_{0,n}^{\mathcal{H}}. \tag{20}$$

Although the matrix $\tilde{\mathbf{U}}_{0,n}^{\mathcal{H}}$ eliminates the inter-user class interference effectively, i.e., $\tilde{\mathbf{U}}_{0,n}^{\mathcal{H}} \sum_{\substack{m=1 \\ m \neq n}}^{N} \mathbf{H}_m \mathbf{s}_m = \mathbf{0}$, when we use $\mathbf{y}_n = \mathbf{Q}_n \mathbf{y}$ in the first approach the noise effects in the detection procedure are mitigated due to the fact that the computation of $\mathbf{Q}_n$ takes the noise into account. Thus, noise effects are reduced which improves the performance, even in the presence of residual interference. In addition, the equivalent channel matrix $\check{\mathbf{H}}_n^1 = \tilde{\mathbf{U}}_{0,n}^{\mathcal{H}} \mathbf{H}_n$ has dimensions $(N_r - r) \times N_t$ as opposed to the matrix $\check{\mathbf{H}}_n^2 = \mathbf{Q}_n \mathbf{H}_n$ which has dimensions $N_{t_n} \times N_{t_n}$. For this reason the computational complexity of the detector is lower if we use the matrix $\mathbf{Q}_n$ to decouple the received signal vector.

The fact that we obtain a square equivalent channel matrix also allows the possibility of using lattice reduction (LR)-based detectors which have a better performance for square channel matrices [27]. Further, the computational complexity to compute the channel inversion (13) and N QR decompositions (15) of matrices with dimensions $N_{t_n} \times N_r$ is much lower than the computational cost of computing N SVD transformations (19) of matrices with dimensions $N_r \times (N_t - N_{t_n})$. For these reasons, in this paper we will focus on the first alternative.

As it will be presented in the next section, the equivalent received signal vector in (18) shows that the process in (11)-(17) is an effective algorithm to separate the user classes at the BS and we can consider the data stream of the nth user class as independent of the received signals of the other user classes. In practice, this allows the possibility of using different transmission and reception schemes for each user class. We can now implement the traditional detectors for each class of users separately which also allows the possibility of using more complex detection schemes due to the reduction of the dimensions of the matrices that need to be processed. The description of the proposed algorithm is presented in Algorithm 1.

3.2 Detection algorithms
In this subsection, we examine signal detection algorithms for massive MIMO in heterogeneous networks. To detect the data stream for each class of users independently, we assume that the DSD algorithm described in Algorithm 1 was previously employed.

Algorithm 1 : The DSD algorithm

Channel and noise variance estimated with training sequence

(1)   Initialization: $\mathbf{H}, \sigma$
             Get the received signal
(2) $\mathbf{y} = \sum_{n=1}^{N} \mathbf{H}_n \mathbf{s}_n + \mathbf{n}$
             $= \mathbf{Hs} + \mathbf{n}$
             Compute the combined channel inversion
(3) $\mathbf{H}^{\dagger} = \mathbf{H}^{\mathcal{H}} \left(\mathbf{HH}^{\mathcal{H}} + \sigma^2 \mathbf{I} \right)^{-1}$
             $= \begin{bmatrix} \ddot{\mathbf{H}}_1 & \ddot{\mathbf{H}}_2 & \dots & \ddot{\mathbf{H}}_N \end{bmatrix}$
(4)   **Do for** $n = 1$ **to** N
             Applying the QR decomposition
(5) $[\mathbf{R}_n \ \mathbf{Q}_n] = \mathrm{RQ}(\ddot{\mathbf{H}}_n)$
(6) $\mathbf{A}_n = \mathbf{Q}_n$
             Compute the equivalent received signal vector
(7) $\mathbf{y}_n = \mathbf{A}_n \mathbf{y}$
             $= \check{\mathbf{H}}_n \mathbf{s}_n + \mathbf{n}_n$
             Applying the detection procedures
(8) $\hat{\mathbf{s}}_n = \mathrm{Detector} \left(\mathbf{y}_n, \check{\mathbf{H}}_n = \mathbf{A}_n \mathbf{H}_n \right)$
(9)   **End**
             Compute the overall estimated signal vector
(10) $\hat{\mathbf{s}} = \begin{bmatrix} \hat{\mathbf{s}}_1^T & \hat{\mathbf{s}}_2^T & \dots & \hat{\mathbf{s}}_N^T \end{bmatrix}^T$

3.2.1 Linear detectors

In linear detectors, the equivalent received signal vector for the nth user class $\mathbf{y}_n \in \mathbb{C}^{N_{t_n} \times 1}$ is processed by a linear filter to eliminate the channel effects [28]. The two linear detectors considered here are given by

$$
\mathbf{W}_n^{\chi} = \begin{cases} \left(\check{\mathbf{H}}_n^{\mathcal{H}} \check{\mathbf{H}}_n \right)^{-1} \check{\mathbf{H}}_n^{\mathcal{H}}, & \chi = \text{Zero Forcing} \\ \left(\check{\mathbf{H}}_n^{\mathcal{H}} \check{\mathbf{H}}_n + \sigma^2 \mathbf{I} \right)^{-1} \check{\mathbf{H}}_n^{\mathcal{H}}, & \chi = \text{MMSE} \end{cases}
$$

$$(21)$$

where $\sigma^2 = \sigma_n^2 / \sigma_s^2$ and $\check{\mathbf{H}}_n \in \mathbb{C}^{N_{t_n} \times N_{t_n}}$ is the equivalent channel matrix of the nth user class. Note that for the MMSE detector, we consider the autocorrelation matrix of the equivalent noise vector as $\mathbf{K}_{\mathbf{n}_n} \approx \sigma_n^2 \mathbf{I}$. As the residual interference is very small, an excellent performance can be obtained with this approximation. The linear hard decision of $\mathbf{s}_n$ is carried out as follows:

$$
\hat{\mathbf{s}}_n = \mathbb{C} \left(\mathbf{W}_n^{\chi} \mathbf{y}_n \right),
$$

$$(22)$$

where the function $\mathbb{C}(x)$ returns the point of the complex signal constellation closest to x. The linear detectors have a lower computational complexity when compared with the non-linear detectors. However, due to the impact

of interference and noise, linear detectors offer a limited performance.

3.2.2 Successive interference cancellation

The successive interference cancellation (SIC) detector for the nth user class in (18) consists of a bank of linear detectors, each detects a selected component $s_{n,i}$ of $\mathbf{s}_n$. The component obtained by the first detector is used to reconstruct the corresponding signal vector which is then subtracted from the equivalent received signal to further reduce the interference in the input to the next linear receive filter. The successively cancelled received data vector that follows a chosen ordering in the ith stage is given by

$$
\mathbf{y}_{n,i} = \mathbf{y}_n - \sum_{j=1}^{i-1} \check{\mathbf{h}}_{n,j} \hat{s}_{n,j},
$$

$$(23)$$

where $\check{\mathbf{h}}_{n,j}$ correspond to the columns of the channel matrix $\check{\mathbf{H}}_n$ and $\hat{s}_{n,j}$ is the estimated symbol obtained at the output of the jth linear detector.

3.2.3 Multiple-branch SIC detection

In the multi-branch scheme [10] for the nth user class, different orderings are explored for SIC, each ordering is referred to as a branch, so that a detector with L branches produces a set of L estimated vectors. Each branch uses a column permutation matrix $\mathbf{P}_n$. The estimate of the signal vector of branch l, $\hat{\mathbf{x}}_n^{(l)}$, is obtained using a SIC receiver based on a new channel matrix $\check{\mathbf{H}}_n^{(l)} = \check{\mathbf{H}}_n \mathbf{P}_n^{(l)}$. The order of the estimated symbols is rearranged to the original order by

$$
\hat{\mathbf{s}}_n^{(l)} = \mathbf{P}_n^{(l)} \hat{\mathbf{x}}_n^{(l)}, \qquad l = 1, \dots, L.
$$

$$(24)$$

A higher detection diversity can be obtained by selecting the most likely symbol vector based on the ML selection rule, that is

$$
\hat{\mathbf{s}}_n = \arg \min \left\| \mathbf{y}_n - \check{\mathbf{H}}_n \hat{\mathbf{s}}_n^l \right\|^2, l = 1, \dots, L.
$$

$$(25)$$

Other detectors could be used with the proposed DSD technique and this is up to the designer to choose the detector.

4 Sum-rate analysis

In this section, a performance analysis for the proposed scheme is presented in terms of the sum rate. We consider that the channel matrix $\mathbf{H}$ was previously estimated at the BS, assume Gaussian signalling and that the received signal vector was decoupled for each user class. Considering

the received signal vector as presented in (18), the sum rate [29] that DSD can offer is defined as

$$R = \sum_{n=1}^{N} \log_2 \frac{\det(\mathbf{K}_{\mathbf{y}_n})}{\det(\mathbf{K}_{\mathbf{n}_n})}, \tag{26}$$

where $\mathbf{K}_{\mathbf{y}_n}$ and $\mathbf{K}_{\mathbf{n}_n}$ are the autocorrelation matrix of the equivalent received signal vector and the equivalent noise vector of the nth user class, respectively. It is easy to show that (26), can be expressed as

$$R = \sum_{n=1}^{N} \log_2 \det \left(\mathbf{I} + \mathbf{B}_n \mathbf{B}_n^{\mathcal{H}}\right), \tag{27}$$

where $\mathbf{B}_n = \mathbf{K}_{\mathbf{n}_n}^{-1/2} \check{\mathbf{H}}_n \mathbf{K}_{\mathbf{s}_n}^{1/2} \in \mathbb{C}^{N_{t_n} \times N_{t_n}}$. As $\mathbf{B}_n \mathbf{B}_n^{\mathcal{H}}$ is a Hermitian symmetric positive-definite square matrix, we have

$$\mathbf{B}_n \mathbf{B}_n^{\mathcal{H}} = \bar{\mathbf{Q}}_n \boldsymbol{\Lambda}_n \bar{\mathbf{Q}}_n^{\mathcal{H}}, \tag{28}$$

where $\bar{\mathbf{Q}}_n$ is a square unitary matrix, $\bar{\mathbf{Q}}_n \bar{\mathbf{Q}}_n^{\mathcal{H}} = \mathbf{I}$, and $\boldsymbol{\Lambda}_n$ is a diagonal matrix whose diagonal elements are the eigenvalues of the matrix $\mathbf{B}_n \mathbf{B}_n^{\mathcal{H}}$. Then, the reliable sum rate that the system can offer is

$$R = \sum_{n=1}^{N} \sum_{i=1}^{N_{t_n}} \log_2 \left(1 + \lambda_{i,n}\right). \tag{29}$$

Note that the eigenvalues $\lambda_{i,n}$ in (29) can be obtained computing the eigenvalues of $\mathbf{B}_n^{\mathcal{H}} \mathbf{B}_n$. As mentioned before, for notational simplicity we assume that $\mathbf{K}_{\mathbf{s}_n} = \sigma_s^2 \mathbf{I}$. When the DSD algorithm is applied, the equivalent noise vector for the nth user class $\mathbf{n}_n = \mathbf{Q}_n \sum_{\substack{m=1 \\ m \neq n}}^{N} \mathbf{H}_m \mathbf{s}_m + \mathbf{Q}_n \mathbf{n} \in \mathbb{C}^{N_{t_n} \times 1}$ is not white due to the residual inter-user class interference. Then its autocorrelation matrix is given by

$$\mathbf{K}_{\mathbf{n}_n} = \mathbb{E}[\,\mathbf{n}_n \mathbf{n}_n^{\mathcal{H}}] = \sigma_s^2 \mathbf{Q}_n \left(\sum_{\substack{m=1 \\ m \neq n}}^{N} \mathbf{H}_m \mathbf{H}_m^{\mathcal{H}}\right) \mathbf{Q}_n^{\mathcal{H}} + \sigma_n^2 \mathbf{I}. \tag{30}$$

Finally, the eigenvalues $\lambda_{i,n}$ in (29) are obtained from the eigenvalues of matrix $\mathbf{B}_n^{\mathcal{H}} \mathbf{B}_n$ given by

$$\mathbf{B}_n^{\mathcal{H}} \mathbf{B}_n = \frac{\sigma_s^2}{\sigma_n^2} \check{\mathbf{H}}_n^{\mathcal{H}} \left[\frac{\sigma_s^2}{\sigma_n^2} \mathbf{Q}_n \left(\sum_{\substack{m=1 \\ m \neq n}}^{N} \mathbf{H}_m \mathbf{H}_m^{\mathcal{H}}\right) \mathbf{Q}_n^{\mathcal{H}} + \mathbf{I}\right]^{-1} \check{\mathbf{H}}_n. \tag{31}$$

To compute the sum rate for the received signal vector in (9) when the detection is perform for all user classes together, we suppress the index n from the above analysis and considering that $\mathbf{K}_{\mathbf{s}} = \sigma_s^2 \mathbf{I}$ and $\mathbf{K}_{\mathbf{n}} = \sigma_n^2 \mathbf{I}$, we get the well-known expression:

$$\bar{R} = \sum_{i=1}^{N_t} \log_2(1 + \lambda_i), \tag{32}$$

where the values λ_i are the eigenvalues of the matrix $\mathbf{B}^{\mathcal{H}} \mathbf{B} = \frac{\sigma_s^2}{\sigma_n^2} \mathbf{H}^{\mathcal{H}} \mathbf{H}$ [30]. In Appendix 1, we show that, as well as the sum rate in (32) when all user classes are detected together, the sum rate in (29) for the proposed algorithm is independent of the detection procedure. However, the lower bound on the achievable uplink sum rate obtained by using linear detectors is different for each detector [31]. In order to analyze the behavior of the lower bound on the sum rate for the proposed scheme, we consider that a linear detector according to (21) is applied to the equivalent received signal vector (18) to detect the transmitted symbol vector of the user class n, then we have

$$\tilde{\mathbf{y}}_n = \mathbf{W}_n \mathbf{y}_n = \mathbf{W}_n \check{\mathbf{H}}_n \mathbf{s}_n + \mathbf{W}_n \mathbf{n}_n. \tag{33}$$

Taking the kth element of $\tilde{\mathbf{y}}_n$ we have

$$\tilde{y}_{k,n} = \mathbf{w}_{k,n} \check{\mathbf{h}}_{k,n} s_{k,n} + \sum_{\substack{j=1 \\ j \neq k}}^{|C_n|} \mathbf{w}_{k,n} \check{\mathbf{h}}_{j,n} s_{j,n} + \mathbf{w}_{k,n} \mathbf{n}_n, \tag{34}$$

where $\mathbf{w}_{k,n}$ is the kth row of $\mathbf{W}_n$. Modeling the noise interference, the inter-user class interference and the inter-user interference in the user class n in (34) as additive Gaussian noise independent of $s_{k,n}$, considering (30) and that the channel is ergodic so that each codeword spans over a large number of realizations, we obtain the lower bound on the achievable rate for the DSD algorithm with linear detectors as described in (35).

Then, the lower bound on the achievable rate for the entire system is given by

$$R_{k,n} = \mathbb{E}\left[\log_2\left(1 + \frac{\sigma_s^2 \mid \mathbf{w}_{k,n} \check{\mathbf{h}}_{k,n} \mid^2}{\sigma_s^2 \sum_{\substack{j=1 \\ j \neq k}}^{|C_n|} \left|\mathbf{w}_{k,n} \check{\mathbf{h}}_{j,n}\right|^2 + \left|\mathbf{w}_{k,n} \left[\sigma_s^2 \mathbf{Q}_n \left(\sum_{\substack{m=1 \\ m \neq n}}^{N} \mathbf{H}_m \mathbf{H}_m^{\mathcal{H}}\right) \mathbf{Q}_n^{\mathcal{H}} + \sigma_n^2 \mathbf{I}\right] \mathbf{w}_{k,n}^{\mathcal{H}}\right|}\right)\right], \tag{35}$$

Arévalo *et al. EURASIP Journal on Wireless Communications and Networking* (2017) 2017:131

Page 8 of 15

$$R = \sum_{n=1}^{N} \sum_{k=1}^{|C_n|} R_{k,n}. \tag{36}$$

For the SIC receiver, each stream is filtered by a linear detector and then, its contribution is subtracted from the received signal to improve the subsequent detection. For each layer the linear filter is recalculated. The performance of SIC detectors can be improved if we choose the cancellation order as a function of the SINR at the output of the linear detector in each layer. The lower bound for the sum rate of the proposed algorithm when a SIC detector is used for each user class, could be calculated updating the expression (35) in each layer, i.e., the values of $\mathbf{w}_{k,n}$ are recalculated for each detected stream.

5 Numerical results

In this section, we evaluate the performance of the proposed algorithm with different detectors in terms of the sum rate and the BER via simulations. Moreover, the computational complexity of the proposed and existing algorithms is also evaluated.

5.1 Sum rate

To evaluate the analytic results obtained in Section 4, the sum rate and the lower bounds for the proposed algorithm with different detection schemes will be evaluated

considering CAS and DAS configuration assuming perfect CSI. For the CAS configuration, we employ $L_k = 0.7$, $\tau = 2$, the distance d_k to the BS is obtained from a uniform discrete random variable distributed between 0.1 and 0.99, the shadowing spread is $\sigma_k = 3$ dB and the transmit and receive correlation coefficients are $\rho_{rx} = 0.2$ and $\rho_{tx} = 0.4$ (when $N_{t_{k,n}} > 1$), respectively. For DAS configurations, we consider a densely populated cell, where a fraction of the active users are in the center of the cell and the remaining users are in other locations of the cell. We explore different particular values for the fraction of users in the center and around the cell. Based on that, we choose specific values for N_B, D, and Q. For the DAS configuration, we also consider $L_{k,j}$ taken from a uniform random variable distributed between 0.7 and 1, $\tau = 2$, the distance $d_{k,j}$ for each link to an antenna is obtained from a uniform discrete random variable distributed between 0.1 and 0.5, the shadowing spread is $\sigma_{k,j} = 3$ dB and the transmit and receive correlation coefficients are $\rho_{rx} = 0.2$ and $\rho_{tx} = 0.4$, respectively.

In Fig. 2 we evaluate the sum rate in two different scenarios for the user requirements. In both cases, we fix the SNR = 10 dB and increase the number of receive antennas. For the DAS configuration, we consider $N_B = 1/2N_r$ antennas at the BS. We also consider $D = 4$ arrays of antennas distributed around the cell, each equipped
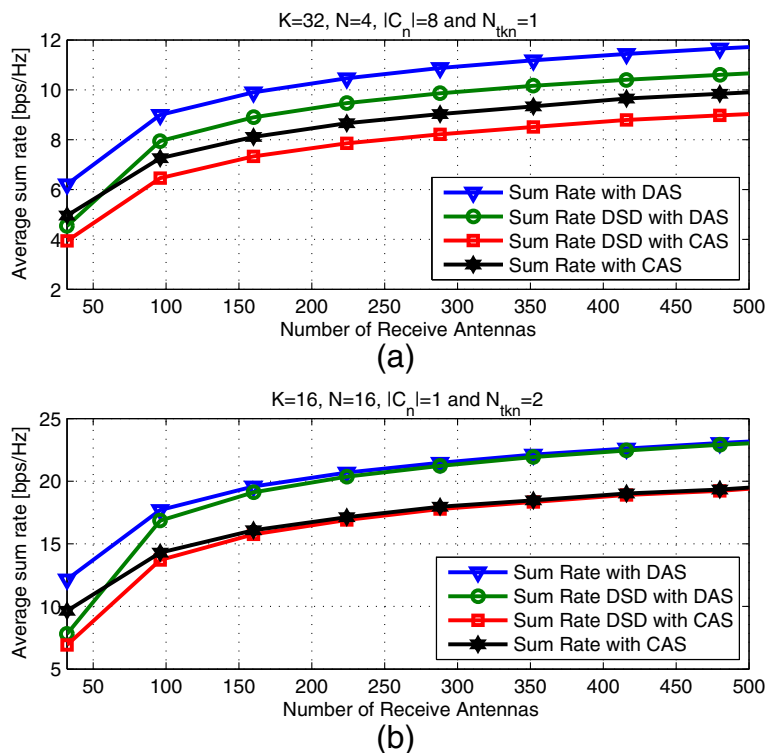

Fig. 2 a Sum rate vs number of receive antennas. $N = 4$, | C_n |= 8 users per class, $N_{t_{k,n}} = 1$ antenna per user. **b** Sum rate vs number of receive antennas. $N = 16$, | C_n |= 1 user per class, $N_{t_{k,n}} = 2$ antennas per user

with $Q = 1/8N_r$ antennas. For Fig. 2a, we consider $N = 4$ classes of users with $| C_n |= 8$ users each and $N_{t_{k,n}} = 1$ antenna per user. We can see that the sum rate of the proposed DSD algorithm is close to the sum rate of the traditional MIMO system and with a low computational complexity on the detection procedures as will be shown in the next subsection. For Fig. 2b, we consider 16 active users in the system and that we need to detect each user independently, i.e., $N = 16$ classes of users with $| C_n |= 1$ user at each class and $N_{t_{k,n}} = 2$ antennas per user. Under these conditions, the sum rate of the proposed scheme reaches the sum rate of the traditional MIMO system, especially for a large number of receive antennas. From the plot in Fig. 2, we can see that the sum rate for DAS is higher than that for the CAS configuration.

In Fig. 3, we compare the lower bound on the sum rate for different detectors such as, ZF, MMSE, and SIC. We consider the DAS configuration under the same conditions as in the previous experiment. For Fig. 3a, we consider $N = 3$ classes of users with $| C_n |= 10$ users at each class and $N_{t_{k,n}} = 1$ antenna per user. We can see from the plot that, similarly to the traditional MIMO systems, the lower bound on the sum rate for ZF and MMSE achieves the sum rate when N_r grows. For Fig. 3b, we consider $N = 2$ classes of users with $| C_n |= 16$ users at each class and $N_{t_{k,n}} = 1$ antennas per user. We can see

that the SIC-MMSE achieves the sum rate and it could be considered optimal in terms of sum rate.

In Fig. 4, we compare the lower bound sum rate versus SNR. We consider $N = 8$ classes of users with $| C_n |= 1$ user in each class and $N_{t_{k,n}} = 8$ antennas per user transmitting with high correlation between antennas $\rho_{tx} = 0.85$. We consider the DAS configuration with $N_B = 96$, $D = 4$, and $Q = 8$. We can see from the plot that the lower bounds for the proposed algorithms are very close to the lower bounds when the detection procedure is carried out together for all users. Since ZF DSD separates the user classes by computing a MMSE matrix which takes the noise component into account, it will be the only detector that outperform its coupled counter part.

Except for the ZF-DSD case, there is a small loss in the performance for DSD schemes. However, with the current processor technologies, to use a SIC MMSE in a 256×256 massive MIMO system could be infeasible. By dividing the users into classes of users, the computational complexity in the detection procedure decreases significantly and makes the use of the SIC-MMSE and more complex detectors feasible.

5.2 Computational complexity analysis

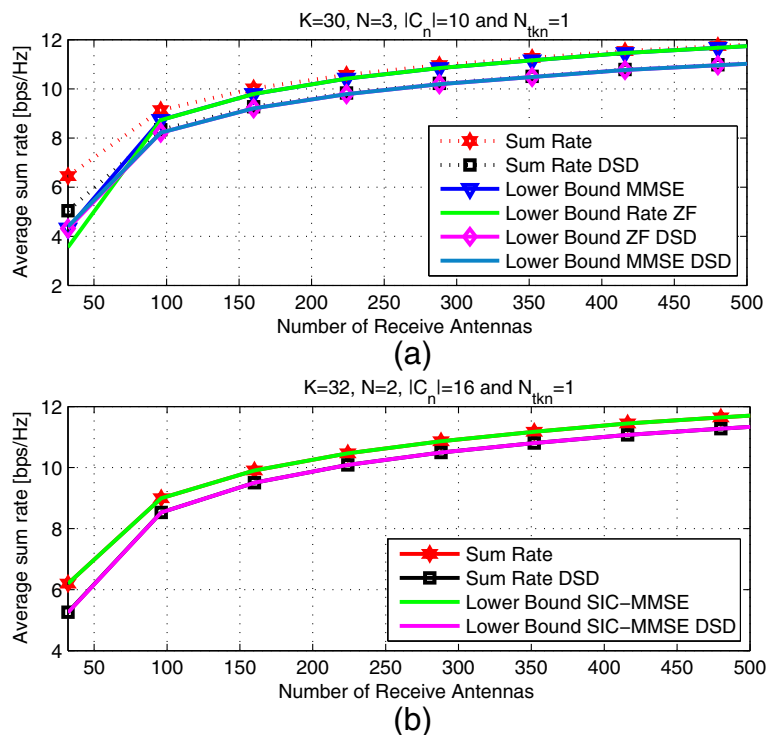In this subsection, the computational complexity of the proposed algorithm is evaluated and compared with the

Fig. 3 a Sum rate vs number of receive antennas. $N = 3$, $| C_n |= 10$ users per class, $N_{t_{k,n}} = 1$ antenna per user. **b** Sum rate vs number of receive antennas. $N = 2$, $| C_n |= 16$ users per class, $N_{t_{k,n}} = 1$ antenna per user
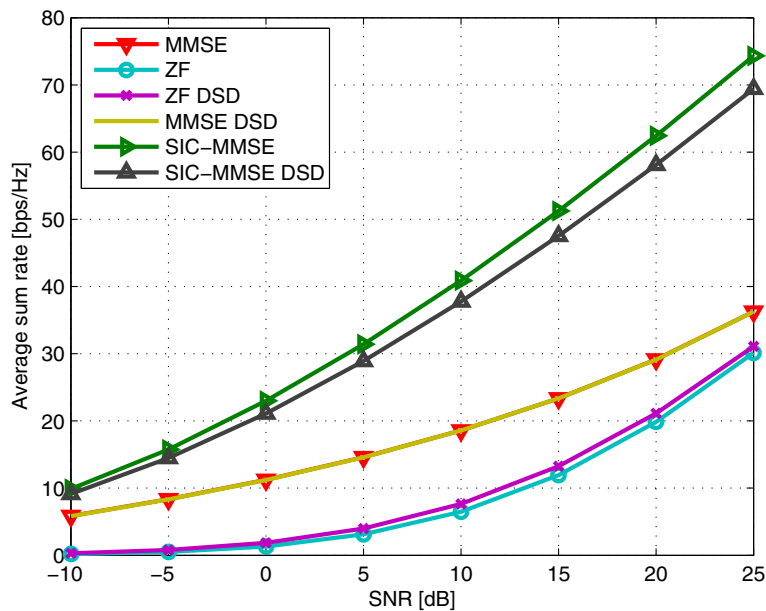
Arévalo *et al. EURASIP Journal on Wireless Communications and Networking* (2017) 2017:131

Page 10 of 15

Fig. 4 Sum rate lower bound vs SNR. $N = 8$, $| C_n | = 1$ user per class, $N_{t_{k,n}} = 8$ antennas per user. We consider DAS configuration with $N_B = 96$, $D = 4$, and $Q = 8$

traditional coupled detection schemes, when all user classes are detected together, by counting the number of floating point operations (FLOPs) per received vector **y**. Different detection schemes are considered such as MMSE, SIC, and MB-SIC. The SIC-based receivers all use MMSE detection. Furthermore, the single-branch SIC and the first branch of the MB-SIC employ norm-based ordering. We consider QPSK modulation; however, the computational cost in these detectors does not change

significantly with the modulation order. The number of FLOPs for the complex QR decomposition of an $N_{t_n} \times N_r$ matrix is given in [32] as $16 \left(N_r^2 N_{t_n} - N_{t_n}^2 N_r + 1/3 N_{t_n}^3 \right)$. To compute the number of FLOPs required for the remaining operations, we use the Light-speed Matlab toolbox [33].

Figure 5 shows the computational complexity versus the number of user classes for different detection algorithms. We consider $K = 100$ active users, $N_{t_{k,n}} = 2$
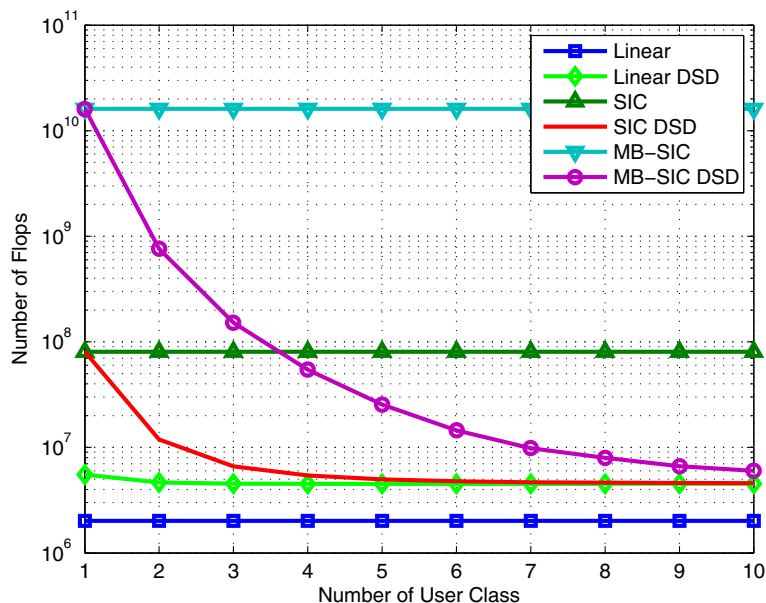

Fig. 5 Computational complexity versus number of user class, $K = 100$ active users, $N_{t_{k,n}} = 2$ antennas per user, $N_r = 200$ receive antennas

Arévalo *et al. EURASIP Journal on Wireless Communications and Networking* (2017) 2017:131
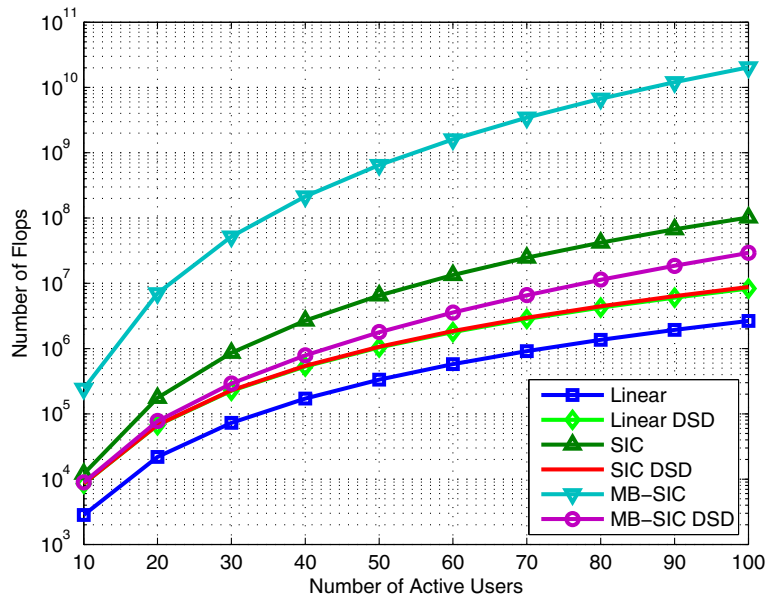
Page 11 of 15

Fig. 6 Computational complexity versus number of active user, $N = 5$ classes of user, $N_{t_{k,n}} = 2$ antennas per user, $N_t = KN_{t_{k,n}}$ transmit antennas and $N_r = 3N_t$ receive antennas

transmit antennas per user and $N_r = 200$ receive antennas distributed around the cell. We consider an increasing number of classes of users, when K is not divisible by the number of classes, the number of active users is set to a smaller value so as to allow the division in N classes. We can see from the figure that the complexity of the SIC and the MB-SIC detectors with the DSD algorithm is lower than the SIC and the MB-SIC coupled detectors,

respectively. Furthermore, for receivers with DSD, the computational complexity is reduced as the number of user classes is increased. This fact represents an important advantage for receivers with DSD, due to the fact that it allows the use of more complex detectors for each user class according to its data requirements.

In Figs. 6 and 7, we plot the required number of FLOPs versus the number of active users and versus the number
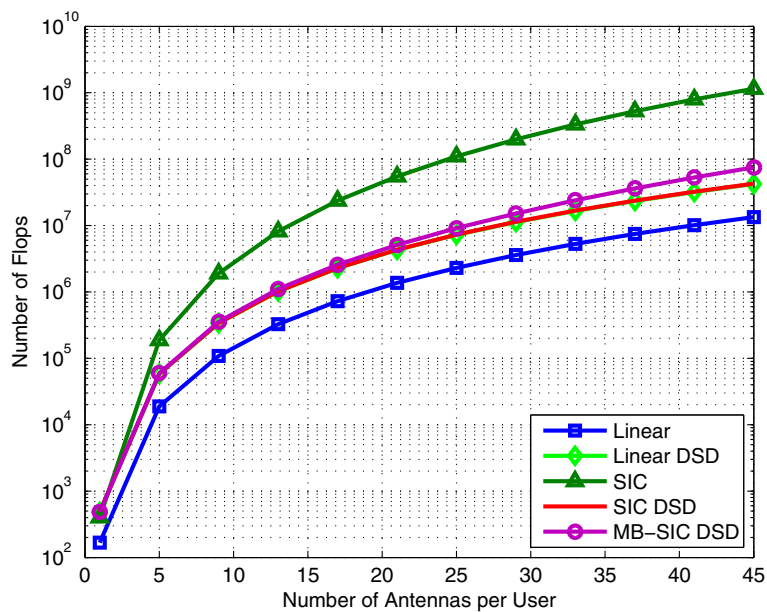

Fig. 7 Computational complexity versus number of antennas per user, $K = 10$ active user, $N = 10$ classes of users, $N_t = KN_{t_{k,n}}$ transmit antennas and $N_r = 2N_t$ receive antennas

of antennas per user, respectively. For Fig. 6 we consider $N = 5$ classes of users, $N_{t_{k,n}} = 2$ transmit antennas per user and $N_r = 3KN_{t_{k,n}}$ receive antennas. The MB-SIC and SIC detectors with DSD have a lower computational cost than the coupled SIC detector. For this 5G context, with a high number of antennas, efficient coupled detectors are not feasible to be implemented; however, if a specific user class requires the benefits of complex detectors, the DSD algorithm reduces the cost so that more complex detectors could be applied as illustrated with MB-SIC in Fig. 6. For the results in Fig. 7, we consider that we have $K = 10$ active users and that we need to detect each user independently, i.e., $N = 10$. The number of transmit antennas per user $N_{t_{k,n}}$ is increased. We also consider that the number of receive antennas distributed around the cell is given by $N_r = 2KN_{t_{k,n}}$. The MB-SIC and SIC detectors with DSD have a significantly lower complexity when compared with the SIC detector where all users are coupled.

It is worth noting that the curves displayed in Figs. 5, 6, and 7 will have a substantial decrease if the channel does not change over a time period due to quasi static channels. In this case, the equivalent channel matrices for each user class are stored for subsequent use. It would increase the gap, in terms of the computational cost, for the detection schemes using the DSD algorithm.

5.3 BER performance
In this subsection, the BER performance of the proposed algorithm is evaluated using different detectors which includes linear, SIC, and MB-SIC with linear MMSE receive filters. The SIC detector of [8] uses a norm-based cancellation ordering, the MB-SIC of [10] employs

a fixed number of branches, equal to the total number of transmit antennas per user class N_{t_n} for the DSD schemes, and norm-based ordering in its first branch. A massive MIMO system operating in heterogeneous networks with K active users is considered. We also consider the DAS configuration where the $N_r = N_B + DQ$ receive antennas are distributed around the cell in D radio heads with Q antennas each and the remaining N_B antennas are located at the BS. We consider QPSK modulation. The SNR per transmitted information bit is defined as

$$\mathrm{SNR} = 10\log_{10}\frac{N_t\sigma_s^2}{M\sigma_n^2}, \tag{37}$$

where σ_s^2 is the common variance of the transmitted symbols, σ_n^2 is the noise variance at the receiver and M is the number of transmitted bits per symbol. The numerical results correspond to an average of 3000 simulations runs, with $500N_t$ symbols transmitted per run. For the N_B antennas at the BS, we employ $L_k = 0.3$, $\tau = 2$, the distance to the users is obtained from a uniform discrete random variable distributed between 0.4 and 0.7, the shadowing spread is $\sigma_k = 1$ dB and the transmit and receive correlation coefficients are equal to $\rho_{rx} = 0.4$. For the R remote arrays of antennas, we use $L_{k,j}$ taken from a uniform random variable distributed between 0.3 and 0.5, the shadowing spread $\sigma_{k,j} = 1$ dB and the receive correlation coefficients are equal to $\rho_{rx} = 0.5$. When the number of transmit antennas at the users is $N_{t_{k,n}} > 1$, the transmit correlation coefficient is equal to $\rho_{tx} = 0.55$.

For the experiment in Fig. 8, we consider $K = 12$ user devices, where each user is equipped with $N_{t_{k,n}} = 3$ transmit antennas and $N = 3$ classes of users. The system
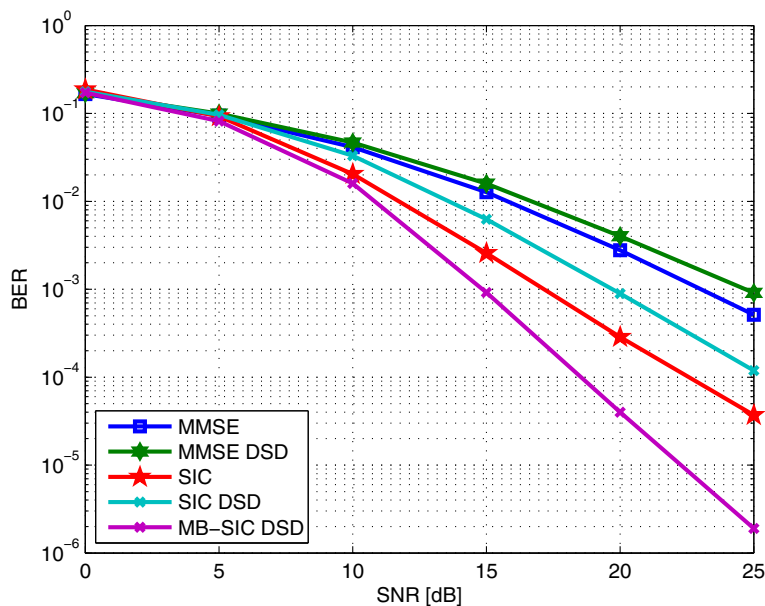

Fig. 8 BER versus SNR with $K = 12$ active users, $N = 3$ classes of users, $N_{t_{k,n}} = 3$ transmit antennas per user and $N_r = 36$ receive antennas
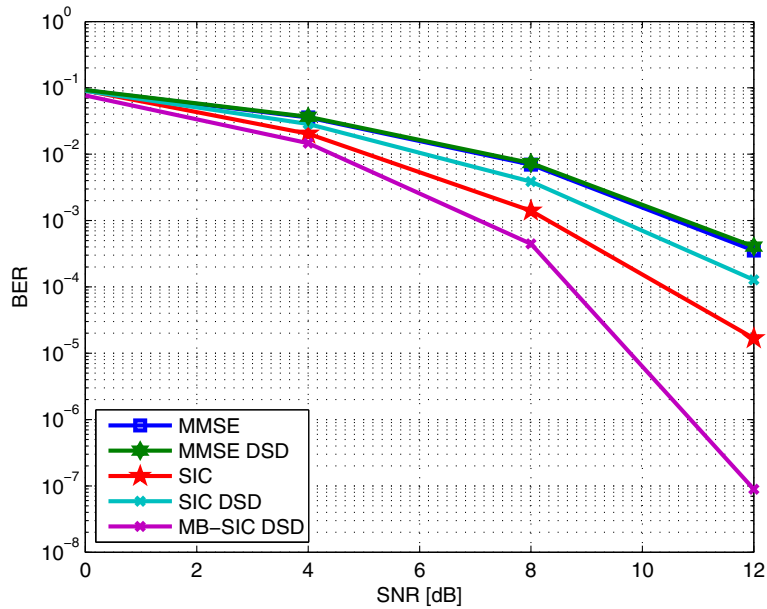
Fig. 9 BER versus SNR with $K = 8$ active users, $N_{t_{k,n}} = 8$ transmit antennas per user, $N = 8$ classes of users, and $N_r = 128$ receive antennas

employs perfect channel state information and QPSK modulation. For the DAS configuration, we consider $N_B = 8$ receive antennas at the BS, $D = 4$ remote radio heads, and $Q = 7$ receive antennas per remote radio head. We can see from the figure that the decoupled SIC detection presents a performance close to the coupled SIC detector with a difference around 2 dB in the high SNR region. In addition the decoupled SIC detector has a

drastic reduction in the computational cost when compared with its coupled version. The result also indicates a remarkable superiority in the performance for the MB-SIC receiver with the DSD scheme over the coupled SIC detector which also has a lower computational complexity than the coupled SIC detector.

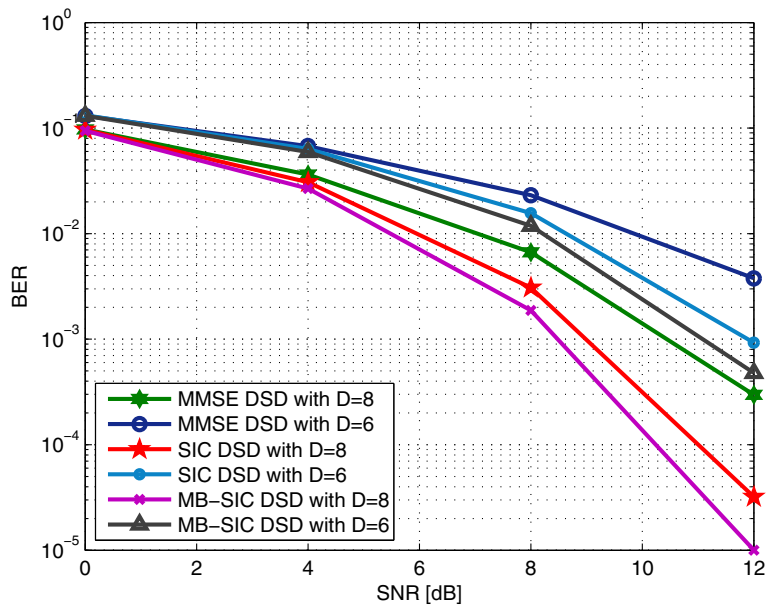In the next example, we evaluate the performance of the proposed scheme considering $K = 8$ active users,

Fig. 10 BER versus SNR with $K = 64$ active users, $N_{t_{k,n}} = 1$ transmit antenna per user, $N = 4$ classes of users, $N_B = 34$, D remote antennas arrays with $Q = 7$ receive antennas per array

each user transmitting with $N_{t_{k,n}} = 8$ antennas. We also consider that we need to detect each user independently of each other, i.e., $N = 8$ classes of users with $\mid C_n \mid = 1$ user per class. For the DAS configuration, we consider $N_B = 64$ receive antennas at the BS, $D = 8$ remote radio heads, and $Q = 8$ receive antennas per remote radio head. The Fig. 9 indicates that the performance of the SIC detector with DSD is close to the SIC detector with a lower computational complexity. Note that the results for MB-SIC with DSD show a very good performance with a computational complexity much lower than the SIC detector without DSD.

In Fig. 10 we evaluate the performance of the proposed scheme with MMSE, SIC, and MB-SIC detection. We consider $K = 64$ active users, $N = 4$ classes of users, $\mid C_n \mid = 16$ users per class, and $N_{t_{k,n}} = 1$ transmit antenna per user. For the DAS configuration, we consider $N_B = 34$ receive antennas at the BS and $Q = 7$ receive antennas per remote radio head. To show the behavior of the BER performance with different numbers of distributed antennas, we consider two configurations of remote radios heads, $D = 8$ and $D = 6$. As expected, the results shows that when the number of RRHs is increased, the BER performance is improved due to the low propagation effects caused by the short distances between users and some remote antenna arrays. We also can see from the figure that the SIC and the MB-SIC detector with DSD offer an excellent BER performance with a low computational cost.

6 Conclusions

In this paper, a mathematical signal model and the DSD algorithm for the uplink of massive MIMO systems operating in heterogeneous cellular networks with different classes of users using CAS and DAS configurations have been presented. The proposed algorithm allows one to separate the received signals for each category of users efficiently into independent parallel single user class signals at the receiver side, applying a common channel inversion and QR decomposition and assuming that the channel matrix was previously estimated. With the proposed scheme, it is possible to handle different classes of users in heterogeneous networks and to use different modulation and/or detection schemes for each class according to its data service requirements. The main advantage of DSD is the reduction in the computational cost of efficient detection schemes that, for its high computational complexity, are not feasible for implementation when the signals received from all active users are coupled.

Appendix 1

Let us consider that a linear detector according to (21) is applied to the equivalent received signal vector (18) to detect the transmitted symbol vector of the nth user class

as in (33). If we define the matrix $\bar{\mathbf{A}}_n = \mathbf{W}_n \check{\mathbf{H}}_n$ and the vector $\bar{\mathbf{n}}_n = \mathbf{W}_n \mathbf{n}_n$ we can rewrite (33) as

$$\tilde{\mathbf{y}}_n = \bar{\mathbf{A}}_n \mathbf{s}_n + \bar{\mathbf{n}}_n, \tag{38}$$

then, in analogy with the analysis in Section 4, the sum rate for the nth user class after the detection is given by

$$R_n = \sum_{i=1}^{N_{t_n}} \log_2(1 + \bar{\lambda}_i), \tag{39}$$

where the values $\bar{\lambda}_i$ are the eigenvalues of the matrix $\bar{\mathbf{B}}_n^{\mathcal{H}} \bar{\mathbf{B}}_n$ with $\bar{\mathbf{B}}_n = \mathbf{K}_{\bar{\mathbf{n}}_n}^{-1/2} \bar{\mathbf{A}}_n \mathbf{K}_{\mathbf{s}_n}^{1/2}$. Then

$$\bar{\mathbf{B}}_n^{\mathcal{H}} \bar{\mathbf{B}}_n = \mathbf{K}_{\mathbf{s}_n}^{1/2} \bar{\mathbf{A}}_n^{\mathcal{H}} \mathbf{K}_{\bar{\mathbf{n}}_n}^{-1} \bar{\mathbf{A}}_n \mathbf{K}_{\mathbf{s}_n}^{1/2}, \tag{40}$$

where $\mathbf{K}_{\mathbf{s}_n} = \sigma_s^2 \mathbf{I}$ and $\mathbf{K}_{\bar{\mathbf{n}}_n} = \mathbf{W}_n \mathbf{K}_{\mathbf{n}_n} \mathbf{W}_n^{\mathcal{H}}$. Assuming that $\mathbf{K}_{\bar{\mathbf{n}}_n}$ has an inverse, we finally obtain

$$
\begin{aligned}
\bar{\mathbf{B}}_n^{\mathcal{H}} \bar{\mathbf{B}}_n &= \sigma_s^2 \bar{\mathbf{A}}_n^{\mathcal{H}} \left(\mathbf{W}_n \mathbf{K}_{\mathbf{n}_n} \mathbf{W}_n^{\mathcal{H}} \right)^{-1} \bar{\mathbf{A}}_n \\
&= \sigma_s^2 \check{\mathbf{H}}_n^{\mathcal{H}} \mathbf{W}_n^{\mathcal{H}} \left(\mathbf{W}_n \mathbf{K}_{\mathbf{n}_n} \mathbf{W}_n^{\mathcal{H}} \right)^{-1} \mathbf{W}_n \check{\mathbf{H}}_n \\
&= \sigma_s^2 \check{\mathbf{H}}_n^{\mathcal{H}} \mathbf{K}_{\mathbf{n}_n}^{-1} \check{\mathbf{H}}_n.
\end{aligned} \tag{41}
$$

Note that (41) and (31) will yield the same result which proves that the sum rate for the DSD algorithm is independent of the linear detection procedure.

Author details
[1]Center for Studies in Telecommunications (CETUC), Pontifical Catholic University of Rio de Janeiro (PUC-Rio), Rio de Janeiro, Brazil. [2]Communications Research Laboratory, Ilmenau University of Technology, Ilmenau, Germany.

References
1. TL Marzetta, Noncooperative cellular wireless with unlimited numbers of base station antennas. IEEE Trans. Wireless Commun. **9**(11), 3590–3600 (2010)
2. EG Larsson, O Edfors, F Tufvesson, TL Marzetta, Massive MIMO for next generation wireless systems. IEEE Commun. Mag. **52**(2), 186–195 (2014)
3. F Rusek, D Persson, B Lau, E Larsson, T Marzetta, O Edfors, F Tufvesson, Scaling up MIMO: Opportunities, and challenges with very large arrays. IEEE Signal Process. Mag. **30**(1), 40–60 (2013)
4. RC de Lamare, Massive MIMO systems: Signal processing challenges and future trends. URSI Radio Science Bulletin (2013)
5. T Marzetta, Noncooperative cellular wireless with unlimited numbers of base station antennas. IEEE Trans. Wireless Commun. **9**(11), 3590–3600 (2010)
6. EG Larsson, F Tufvesson, O Edfors, TL Marzetta, Massive MIMO for next generation wireless systems. IEEE Commun. Mag. **52**(2), 186–195 (2014)
7. 3GPP-LTE, Technical specification group radio access network: evolved universal terrestrial radio access (E-UTRA): further advancements for E-UTRA physical layer aspects (release 13). 3GPP TR136.942 (2016)

Arévalo *et al. EURASIP Journal on Wireless Communications and Networking*   (2017) 2017:131

Page 15 of 15

8. GD Golden, CJ Foschini, RA Valenzuela, PW Wolniansky, Detection algorithm and initial laboratory results using V-BLAST space-time communication architecture. Electron. Lett. **35**(1), 14–16 (1999). doi:10.1049/el:19990058

9. RC de Lamare, R Sampaio-Neto, Minimum mean-squared error iterative successive parallel arbitrated decision feedback detectors for DS-CDMA systems. IEEE Trans. Commun. **56**(5), 778–789 (2008)

10. RC de Lamare, Adaptive and iterative multi-branch MMSE decision feedback detection algorithms for multi-antenna systems. IEEE Trans. Wirel. Commun. **12**(10), 5294–5308 (2013). doi:10.1109/TWC.2013.092013.130233

11. P Li, RC de Lamare, R Fa, Multiple feedback successive interference cancellation detection for multiuser MIMO systems. IEEE Trans. Wirel. Commun. **10**(8), 2434–2439 (2011)

12. P Demestichas, A Georgakopoulos, D Karvounas, K Tsagkaris, V Stavroulaki, J Lu, C Xiong, J Yao, 5G on the horizon: Key challenges for the radio-access network. IEEE Veh. Technol. Mag. **8**(3), 47–53 (2013)

13. H Pervaiz, L Musavian, Q Ni, in *2015 IEEE International Conference on Communication Workshop (ICCW)*. Area energy and area spectrum efficiency trade-off in 5G heterogeneous networks, (2015), pp. 1178–1183

14. W Roh, A Paulraj, MIMO channel capacity for the distributed antenna systems. IEEE Veh. Technol. Conf. **3**, 1520–1524 (2002)

15. D Wang, J Wang, X You, Y Wang, M Chen, X Hou, Spectral efficiency of distributed MIMO systems. IEEE J. Selected Areas Commun. **31**(10), 2112–2127 (2002)

16. F Römer, M Fuchs, M Haardt. Distributed MIMO systems with spatial reuse for highspeed indoor mobile radio access. 20th Meeting of the Wireless World Research Forum (WWRF), Ottawa, Canada, (2008)

17. J Andrews, S Buzzi, W Choi, S Hanly, A Lozano, A Soong, J Zhang, What will 5g be? IEEE J. Selected Areas Commun. **32**(6), 1065–1082 (2014)

18. J Kermoal, L Schumacher, K Perdersen, et al, A stochastic MIMO radio channel model with experimental validation. IEEE J. Selected Areas Commun. **20**(6), 1211–1226 (2002)

19. QH Spencer, AL Swindlehurst, M Haardt, Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels. IEEE Trans. Signal Process. **52**(2), 461–471 (2004)

20. LU Choi, RD Murch, A transmit preprocessing technique for multiuser MIMO systems using a decomposition approach. IEEE Trans. Wireless Commun. **3**(1), 20–24 (2004)

21. V Stankovic, M Haardt, Generalized design of multi-user MIMO precoding matrices. IEEE Trans. Wireless Commun. **7**(3), 953–961 (2008)

22. CB Chae, S Shim, RW Heath, Block diagonalized vector perturbation for multiuser MIMO systems. IEEE Trans. Wireless Commun. **7**(11), 4051–4057 (2008)

23. L Arévalo, RC de Lamare, M Haardt, R Sampaio-Neto, in *2015 IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*. Uplink block diagonalization for massive MIMO-OFDM systems with distributed antennas, (2015)

24. V Stankovic, M Haardt. Improved diversity on the uplink of multi-user MIMO systems. 2005 European Conference on Wireless Technology (EuWiT), (2005), pp. 113–116

25. K Zu, RC de Lamare, M Haardt, Generalized design of low-complexity block diagonalization type precoding algorithms for multiuser MIMO systems. IEEE Trans. Commun. **61**(10), 4232–4242 (2013)

26. H Sung, S Lee, I Lee, Generalized channel inversion methods for multiuser MIMO systems. IEEE Trans. Commun. **57**(11), 3489–3499 (2009)

27. L Arévalo, RC de Lamare, K Zu, R Sampaio-Neto. Multi-branch lattice reduction successive interference cancellation detection for multiuser MIMO systems. IEEE International Symposium on Wireless Communications Systems, (2014), pp. 219–223

28. A Paulraj, R Nabar, D Gore, *Introduction to space-time wireless communications*. (Cambridge University Press, 2003)

29. CK Wen, S Jin, KK Wong, On the Sum-Rate of Multiuser MIMO Uplink Channels with Jointly-Correlated Rician Fading. IEEE Trans. Commun. **59**(10), 2883–2895 (2011)

30. DNC Tse, P Viswanath, *Fundamentals of wireless communications*. (Cambridge University Press, 2005)

31. HQ Ngo, EG Larsson, TL Marzetta, Energy and spectral efficiency of very large multiuser MIMO systems. IEEE Trans. Commun. **61**(4), 1436–1449 (2013)

32. G Golub, CV Loan, *Matrix computations*. (The Johns Hopkins University Press, 1996)

33. T Minka, *The Lightspeed Matlab toolbox, efficient operations for Matlab programming, version 2.2*. (Microsoft Corp, 2007)