Distributed Downlink Multi-Cell Processing Requiring Reduced-Rate Backhaul Data Exchange

Rong Zhang\textsuperscript{1}, Krishnamurthy. Giridhar\textsuperscript{2} and Lajos Hanzo\textsuperscript{1}
\textsuperscript{1} School of ECS., Univ. of Southampton, SO17 1BJ, UK.
\textsuperscript{2} Department of Electrical Engineering, I.I.T. Madras, INDIA.
Email: rz, lh@ecs.soton.ac.uk, http://www-mobile.ecs.soton.ac.uk

Abstract—Different-complexity Multi-Cell Processing (MCP) schemes employing Distributed Signal-to-Interference-Leakage-plus-Noise-Ratio (SILNR) precoding techniques are proposed, which require reduced backhaul data exchange in comparison to the conventional MCP structure. Our results demonstrate that the proposed structures are capable of increasing the achievable cell-edge throughput and offering different geographic rate profile distributions as well as meeting different delay requirements.

I. INTRODUCTION

Multiple Input Multiple Output (MIMO) system based Multi-Cell Processing (MCP)\cite{1}, \cite{2} constitutes a promising enabler for improving the throughput of cell-edge Mobile Stations (MSs) by jointly pre-processing the data of all the cooperating Base Stations (BSs) involved. Moreover, MCP requires the Channel State Information (CSI) of all the links to all supported MSs at all transmitters in order to facilitate Down-link (DL) joint precoding. This is typically achieved by a Central Unit (CU), which connects all the cooperating BSs considered via a limited-rate backhaul.

In the context of linear precoding, the so-called Signal-to-Interference-Leakage-plus-Noise-Ratio (SILNR) maximisation technique\cite{3}, which carefully balances the received signal power of the target MS against the interference power imposed on the remaining MSs, is capable of combining the benefits of both the egoistic Beam-Forming (BF) technique and of the altruistic Zero-Forcing (ZF) technique\cite{4}. The SILNR maximisation technique may be implemented in either a joint or distributed fashion. For the joint SILNR (JSILNR)\cite{3}, the precoding matrix of all BSs cooperating for transmitting to a particular MS has to be determined jointly at the CU with the aid of all the MIMO channels involved. This is quite a demanding requirement, necessitating the backhaul exchange of both the data streams of all the MSs as well as the DL CSI of all the MSs. By contrast, for the distributed SILNR (DSILNR)\cite{5}, the BS’s precoding matrix calculated for transmission to a particular MS may be determined locally at each individual BS, where only the channels spanning from the corresponding BS to the MSs are involved in the precoding matrix calculation. Hence, the DSILNR technique substantially reduces the signalling requirements from a CSI exchange perspective.

However, the challenges of MCP relying on reduced data - rather than CSI - exchange have not been explored in the open literature. Hence, for the sake of reducing the burdens imposed on practical limited-rate backhaul design, we propose a range of reduced-complexity MCP structures employing DSILNR relying on a reduced amount of data exchange and investigate their achievable outage rate and delay-limited performance.

We organise our paper as follows. In Section II, we provide a unified system model for Single Cell Processing (SCP) and MCP. Then, various low-complexity MCP structures are proposed along with their precoding techniques employed. In Section III, we characterise the attainable performance of our MCP structures. Finally, we conclude in Section IV.

Notation: Throughout the paper, lower (upper) case boldface letters represent column vectors (matrices). The superscript $(\cdot)^T$ denotes transposition and $(\cdot)^H$ represents the conjugate transpose. In addition, I represents the identity matrix.

II. SYSTEM DESCRIPTION

A. Unified Model

Let us first introduce the cellular topology of Fig 1, where the hexagonal three-sector cellular model associated with a unity frequency reuse is employed. Let $B_c$ denote the set hosting the BSs involved in the cooperative transmission. Practically, $|B_c|=3$ denotes a realistic scenario, where $|\cdot|$ represents the cardinality of a set. These three adjacent BSs form a joint cooperative transmission site, as indicated by the hexagonal area surrounded by the dashed line in Fig 1. Hence the cell-edge of the conventional cells effectively becomes the cell-centre of the newly formed cooperative site, as indicated by the shaded circle in Fig 1. These cooperative BSs are connected to a common CU via a limited-rate backhaul.

Consider a general cooperative scenario constituted by $|B_c|=N_b$ BSs, where each BS is equipped with $N_t$ transmit antennas. Let us assume that a total of $N_u = N_b$ MSs - each equipped with $N_r$ receive antennas - are involved in the cooperative scenario, where each of the $N_u$ MSs roams within the coverage area of a different BS, which is often referred to as its anchor BS, as portrayed in Fig 1. Furthermore, we assume that each MS has to receive $N_s$ multiplexed DL data streams. Hence, the cooperative scenario may be described by the parameter combination $\{N_b, N_t, N_u, N_r, N_s\}$. Additionally, we let $N_T = (N_b \times N_t)$, $N_R = (N_u \times N_r)$ and $N_S = \ldots$
Fig. 1. The cellular topology considered and multicell processing

$$(N_u \times N_s)$$ denote the total number of BS transmitter antennas, MS receiver antennas and MS data streams, respectively.

Hence, a unified discrete-time model for the signal received by MS$_j$ for both SCP and MCP may be written as:

$$y_j = H_j G E x + n_j, \quad (1)$$

where $y_j \in \mathbb{C}^{N_r}$ and $n_j \in \mathbb{C}^{N_r}$ denote the received signal vector of MS$_j$ and the circularly symmetric complex Gaussian noise vector having a covariance matrix of $R_n = N_0 I$, respectively. Furthermore, $x = [x_1^T, x_2^T, \ldots, x_{N_s}^T]^T$ hosts all data streams of all MSs, where $x_i \in \mathbb{C}^{N_r}, j \in [1, N_u]$ denotes the $N_s$ data streams destined for MS$_j$ that have i.i.d. zero-mean unit-variance complex Gaussian entries, which are assumed to be independent of both the noise and of the channel. The entries of $x$ obey $\mathbb{E}[x_i x_j^H] = I$, where $\mathbb{E}[]$ stands for the expectation operation. Furthermore, $H_{j \rightarrow i} \in \mathbb{C}^{N_r \times N_T}$ represents the equivalent MIMO channel matrix and similarly, $G_{E} \in \mathbb{C}^{N_T \times N_S}$ is the equivalent linear precoding matrix.

More explicitly, we let $H_{j \rightarrow i} \in \mathbb{C}^{N_r \times N_t}, i \in [1, N_u], j \in [1, N_s]$ denote the linear precoding matrix between MS$_j$ and BS$_i$. We then let $G_{i,j} \in \mathbb{C}^{N_T \times N_T}$, denote the linear precoding matrix employed at BS$_i$, which was calculated for transmission to MS$_j$. Hence, the global MIMO channel matrix $H \in \mathbb{C}^{N_r \times N_T}$ and the global linear precoding matrix $G \in \mathbb{C}^{N_T \times N_S}$ may be expressed as:

$$H = \begin{bmatrix} H_{1,1} & \cdots & H_{1,N_u} \\ \vdots & \ddots & \vdots \\ H_{N_u,1} & \cdots & H_{N_u,N_u} \end{bmatrix}, \quad G = \begin{bmatrix} G_{1,1} & \cdots & G_{1,N_s} \\ \vdots & \ddots & \vdots \\ G_{N_s,1} & \cdots & G_{N_s,N_s} \end{bmatrix}.$$  

The equivalent MIMO channel matrix $H_{E}$ of Eq (1) may then be expressed as $H_{E} = H_{j \rightarrow i} G_{E}$, where $H_{j \rightarrow i}$ denotes the jth row of $H$. According to the equivalent linear precoding matrix $G_{E}$ of Eq (1), for SCP, we may write $G_{E} = \text{diag}([G])$, which is a diagonal matrix hosting the diagonal entries of matrix $G$. By contrast, for MCP, we have $G_{E} = G$, which subsumes the various MCP structures to be introduced in the next section. Hence the discrete-time signal received at MS$_j$ for SCP and MCP may be respectively written as:

$$y_j^S = H_{j} G_{j} x_j + \sum_{i \in B_{i-1}} H_{j,i} G_{i} x_i + n_j, \quad (2)$$

$$y_j^m = H_{j \rightarrow i} G_{v,j} x_j + \sum_{k=1, k \neq j}^{N_s} H_{j \rightarrow k} G_{v,k} x_k + n_j, \quad (3)$$

where $G_{v,j}$ denotes the jth column of matrix $G$ and $B_{i-1}$ hosts the BSs within the cooperative site, excluding BS$_j$.

B. Conventional MCP Structure and Distributed SILNR

In conventional MCP as seen at the left of Fig 2, each BS is responsible for the co-processing of all MSs’ data streams $x$, which requires a substantial amount of data exchange before their precoded DL transmission ensues. Hence the corresponding global linear precoding matrix $G$ has only non-zero entries $G_{i,j}, i \in B_c, j \in [1, N_u]$ and this structure may be referred to as Equally Loaded (EL) cooperative BS structure.

The linear SILNR maximisation precoding technique [3] aims for maximising the signal power received at the intended MS and at the same time for minimising the interference imposed on all other MSs. In the context of DSILNR [5], this maximisation takes place locally at each individual BS, where the SILNR $\gamma$ at MS$_j$ is given by:

$$\gamma = \frac{\text{Tr}[G_{i,j} H_{i,j}^H G_{i,j}]}{\text{Tr}[G_{i,j}^H (N_0 N_I + \sum_{k=\chi, j}^{\infty} H_{k,i}^H H_{k,i}) G_{i,j}]}.$$  

where the numerator denotes the signal power received at MS$_j$ benefiting from the transmission at BS$_i$, while the denominator represents the interference leakage power imposed on all other MSs due to the transmission of BS$_i$ to MS$_j$ with $[\chi_u, \chi_l] = [N_u, 1]$. Hence the optimisation problem is:

$$G_{i,j} = \arg \max_{G_{i,j}} \gamma.$$  

The above maximisation problem of Eq (5) requires the knowledge of the power allocation $P_{i,j}$. To satisfy the per-BS power constraint $P_{BS}$, which is assumed to be the same for all BSs, we adopt the power allocation strategy of [5]:

$$P_{i,j} = \frac{\text{Tr}[H_{i,j}^H H_{i,j}]}{\text{Tr}[\sum_{k=\chi}^{\chi_u} H_{k,i}^H H_{k,i}]} P_{BS},$$  

where the numerator denotes the channel gain associated with the transmission of BS$_i$ to MS$_j$, while the denominator represents the aggregate channel gain of the transmission of BS$_i$ to all MSs. Again, we have $[\chi_u, \chi_l] = [N_u, 1]$ for the conventional MCP structure. Once the power allocation was determined, the optimisation problem of Eq (5) may be further decoupled into individual optimisation steps by constraining $G_{i,j}$ to be an orthonormal matrix, as discussed in [3]. Hence we may obtain the optimisation solution as $G_{i,j} = \text{eigv}(\text{B}^{-1} \text{A})$, representing the eigenvectors corresponding to the $N_u$ largest eigenvalues of $\text{B}^{-1} \text{A}$, with $\text{B} = N_0 N_t I / P_{BS} + \sum_{k=1, k \neq j}^{N_s} H_{k,i}^H H_{k,i}$ and $\text{A} = H_{i,i}^H H_{i,i}$. Finally, the resultant linear precoding matrix $G_{i,j}$ is multiplied by $P_{i,j}/N_s$, where we assumed that the power $P_{i,j}$ is equally distributed amongst the $N_s$ DL data streams to be transmitted to MS$_j$.

C. Reduced-Complexity MCP Structures

1) UEL-I: In the first reduced-complexity structure considered, BS$_i$ is only responsible for transmission to MSs within its own cell, as well as for transmissions to all the MSs roaming within the previous $(i-1)$ BSs, where the data streams $[x_1, \ldots, x_{i-1}]$ of all the $(i-1)$ MSs have to be known.
at BS₂. Hence, this may be referred to as the UnEqually Loaded structure of type one (UEL-I). For example, when considering |B₉| = 3 as seen in the middle of Fig 2, BS₁ is only responsible for transmissions to its own cell-edge MS₁, while BS₂ is responsible for transmissions to its own cell-edge MS₂ and to MS₁ and finally, BS₃ jointly transmits to all MS₁, MS₂, and MS₃ have the same configuration as that of UEL-I, with the minor change that the data streams x₁, x₂, and x₃ jointly transmit to MS₁, MS₂, and MS₃ have to be available at BS₃. Hence, the global linear precoding matrix G becomes a lower-triangular matrix. The DSILNR precoding technique employed remains similar to that discussed in Section II-B, with the slight change that the parameters of Eq (6) become [xᵢ, χᵢ] = [i, 1].

2) UEL-II: In the second reduced-complexity structure advocated, BS₁ is only responsible for transmissions to MSs in its own cell as well as for transmissions to MSs belonging to BS₁, where only the data streams xᵢ−₁ have to be available at BS₁. This may be referred to as the UnEqually Loaded structure of type two (UEL-II). For example, when considering |B₉| = 3 as seen at the right of Fig 2, BS₁ and BS₂ have the same configuration as that of UEL-I, while BS₃ is only responsible for transmissions to MS₂ and MS₃, where only the data streams x₂ of MS₂ have to be available at BS₃. Hence, the global linear precoding matrix G becomes a zig-zag shaped matrix. The DSILNR precoding technique essentially remains the same as discussed in Section II-B, where the parameters of Eq (6) become [xᵢ, χᵢ] = [i, max(i − 1, 1)].

3) IL-I and IL-II: The third reduced-complexity structure is similar to the UEL-I structure, with the slight difference that the MCP is operated progressively, where the participation of BS does not affect the transmission precoding matrix employed at the BSs that have already ongoing transmissions. For example, when BS₃ participates in the MCP, then the precoding matrices already employed at BS₁ and BS₂ remain unchanged. Hence, this structure may be referred to as the Incrementally Loaded structure of type one (IL-I). Note that this structure has the same data stream exchange requirement as that of the UEL-I structure, while imposing only marginal modifications on the network. Likewise, we may define the Incrementally Loaded structure of type two (IL-II) as the progressive version of the UEL-II structure. For the other schemes, the DSILNR precoding technique remains the same, as that discussed in Section II-B, with the minor change that the parameters of both Eq (6) and Eq (4) become [xᵢ, χᵢ] = [i, 1] for the IL-I structure and [xᵢ, χᵢ] = [i, max(i − 1, 1)] for the IL-II structure, respectively.

### III. Performance Evaluation

In this section, we characterise the various MCP structures employing the DSILNR precoding technique in both the fully-loaded scenario (NT = NR) of \{N₁, N₂, N₃, N₄, N₅\} = \{3, 2, 3, 2, 1\} and the lightly-loaded scenario (NT > NR) of \{N₁, N₂, N₃, N₄, N₅\} = \{2, 3, 3, 2, 1\}.

We considered the so-called Urban Micro setup [6], where the BS-to-BS distance was defined as D = 1000m. The MIMO channel of each BS-MS pair is constituted by three components, i.e., by 

\[ h_{ij} = (A_{ij}^1, A_{ij}^2, 1)^{l/2} h_{ij}, \]

where \( h_{ij} \) represents the fast fading component, which is assumed to be frequency-flat with zero-mean and unity-variance complex Gaussian entries, \( A_{ij}^1 \) is the lornormal shadow fading component, where \( \xi \) is generated by a zero-mean real-valued Gaussian random variable having a standard derivation of \( \sigma_s = 8dB \). The pathloss model is given by \( d^{-\alpha} \), where \( d \) denotes the BS-MS distance in meter and \( \alpha [\alpha, \beta] = [-3, 1.35 \times 10^7] \) [7]. Furthermore, we let the Signal-To-Noise Ratio (SNR) at the cell-edge area be 25 dB.

In each simulation, \( N_s = 3 \) MS locations are generated randomly and independently, where each of the MSs is randomly uniformly positioned in the area determined by the radius range of \( r \) of \( (0, 0.1D) \) and angle range of 120°, corresponding to the sector covered by its anchor BS. Furthermore, we investigate a total of 100 independent simulations, where each simulation was terminated after 5000 simulation runs.

#### A. Achievable Outage Rate

1) Performance Metric: We plot the 'outage rate', which is defined as the Cumulative Distribution Function (CDF) of the ergodic achievable rate. We consider an equal power allocation in the SCP scenario, namely \( G_{i,j} = \sqrt{P_{BS}/N_s} \), when a single-user detector is employed as a benchmark. The achievable rate of MSᵢ of the SCP and MCP schemes are given as:

\[ R_j^o = \log | I + \frac{H_{j,j} G_{j,j} G_j H_j^H}{N_0 I + \sum_{i \in \{B_{i,j} \}} H_{j,j} G_{i,j} G_i H_i^H}|, \]

\[ R_j^{o n} = \log | I + \frac{H_{j,j} G_{j,j} G_j H_j^H}{N_0 I + \sum_{k=1}^{N_s} \sum_{k,j} H_{j,j} G_{j,j} G_{k,j} H_{k,j}^H}|. \]

2) Performance Investigation: Fig 3 shows the outage rate of each of the three MSs for both the conventional and the proposed structures in the fully-loaded scenario and lightly-loaded scenario, respectively. It can be seen for both scenarios that all three MSs supported by all structures considered achieve a higher outage rate than that of the SCP arrangement marked by the asterisks. In general, MS₂ characterised in the middle subplot of Fig 3 exhibits a similar outage rate for all structures considered, while MS₁ and MS₃ exhibit quite a different outage rate behaviour for the various structures.

As seen in the top subplot of both Fig 3(a) and Fig 3(b) for MS₁, the proposed type one structures are indicated by the solid line for UEL-I and the dash-dot line for IL-I, both of which achieve a outage rate improvement over the conventional structure indicated by the crosses. However, the proposed type two structures indicated by the dashed
line for UEL-II and the dotted line for IL-II only achieve marginal or no outage rate improvement, when compared to the conventional structure. The outage rate improvement of MS₁ in the proposed structures is a benefit of the higher total transmission power radiated from BS₁,₂ for their transmission to MS₁ compared to the conventional structure.

On the other hand, observe in the bottom subplot of both Fig 3(a) and Fig 3(b) for MS₃, all proposed structures suffer from an outage rate loss, when compared to the conventional structure. Since in contrast to the conventional structure, where MS₃ benefits from multiple BSs’ transmission, only the anchor BS₃ is responsible for transmission to MS₃ in the proposed structures and only a fraction of the total transmission power is available for BS₃ for the transmission to MS₃. Importantly, the achievable outage rate remains still higher than that of the SCP, where the full transmit power is available at BS₃ for transmissions to MS₃. This is because although the transmit power available at BS₃ for transmission to MS₃ is reduced in MCP in comparison to the SCP scenario, the interference imposed by the other BSs on MS₃ is reduced more substantially, as a benefit of the DSILNR precoding technique employed at all BSs. This also explains the fact that the UEL-I and UEL-II structures achieve a higher outage rate than the IL-I and IL-II structures, since the precoding techniques employed in the latter two approaches do not take into account any of the newly activated cooperative BSs, which in turn does not facilitate the control of the interference imposed on those newly initiated MSs.

B. Delay Limited Performance

We now investigate the delay limited performance for both the best-supported MS₁ and the worst-supported MS₃.

1) Performance Metric: Given a system bandwidth $B$, a packet’s transmission may be modelled by a queue and a wireless link using the M/G/1 model of [8], which has a Poissonian source packet arrival process having an arrival rate of $\lambda$, a general i.i.d. packet delivery time $T$ and a single server. Let us denote the maximum average tolerable delay by a particular service as $D$, which should be no less than the packet delivery time $T$ plus the queueing delay $D_q$. For simplification, we do not consider applications that have hard delay requirements and ignore the propagation delay, the data exchange delay and the feedback delay. Then we have $T + D_q \leq D$, which may be quantified by the Pollaczek-Khintchine formula [8]:

$$\mathbb{E}[T] + \lambda[\mathbb{E}[T]^2 + \sigma_T^2]/[2 - 2\lambda\mathbb{E}[T]] \leq D, \quad (9)$$

where $\mathbb{E}[T]$ and $\sigma_T^2$ denote the Average Successful Packet Delivery Time (ASPDT) and its variance, respectively. We assume encountering i.i.d. packet error events and assume that the corrupted packets are retransmitted by an Automatic Repeat reQuest (ARQ) scheme, which keeps retransmitting a given packet, until its successful reception is declared. Hence the probability of the $l$th transmission of a packet succeeding is given by $Pr(l) = f(\gamma)[1 - f(\gamma)]^{l-1}$, where $f(\gamma)$ is defined as the Packet Success Ratio (PSR). In this paper, we consider Gray-mapped square 2$^3$-QAM packets having a length of $N$ bits, where $b$ represents the number of Bits Per Symbol (BPS).

Hence the PSR is given by [9]:

$$Pr(l) = f(\gamma)[1 - f(\gamma)]^{l-1}, \quad (11)$$

where the packet duration is $\tau = N/bR$ and $R$ represents the symbol rate. By substituting Eq (10) and Eq (11) into Eq (9), we arrive at the following constraint:

$$N/bR\lambda + \lambda N/bR - \lambda N^2/2\sigma_T^2 \leq f(\gamma). \quad (12)$$
Let us define the normalised transmission rate as \( R_n = R/B \), the normalised delay as \( D_n = DB \) and the normalised service arrival rate as \( \lambda_n = \lambda/B \). Given \( D_n \), \( b \), \( N \) and \( \lambda_n \), we may find a set of combined solutions \([b, R_n]\) for Eq (12), when the equality holds subject to the following constraints. Firstly, the average delay is higher than the packet duration, hence we have \( D > \tau \). Secondly, the service arrival rate considered must not be higher than the packet service rate \( \mu = 1/E[T] \) delivered by the system, hence we have \( \lambda < \mu \).

Our goal is to find the minimum normalised transmission rate \( R_n \) so that a given service maximum tolerable normalised delay \( D_n \) is not exceeded, which is characterised in Fig 4 for both the fully-loaded and lightly-loaded scenarios in both the conventional and the proposed structures, when the normalised service arrival rate is set to \( \lambda_n = 0.005 \) and the packet length is set to \( N = 240 \) bits.

2) Delay Investigation: In the fully-loaded scenario characterised in Fig 4(a) we observe for MS1 that the proposed structures require a consistently lower normalised transmission rate \( R_n \) for maintaining a given maximum delay \( D_n \) than the conventional structure. When the normalised transmission rate \( R_n \) approaches its maximum value of unity, the system became capable of operating at a lower delay, hence the proposed structures support a tighter delay constraint. On the other hand, for MS3, only the UEL-II structure is capable of imposing as low a delay as the conventional structure, and this is at the cost of requiring a higher normalised transmission rate \( R_n \).

When considering the lightly-loaded scenario of Fig 4(b) and MS1, both of the proposed type one structures operate at a lower delay than the conventional EL structure, while both proposed type two structures achieved a similar delay at the cost of consistently requiring a higher normalised transmission rate \( R_n \). On the other hand, we observe for MS3 that although the outage rate performance of the proposed structures is inferior in comparison to that of the conventional structure, as seen in the bottom subplot of Fig 3(b), the delay imposed by both the UEL-I and UEL-II structures is the same as that of the conventional structure, which is achieved at the cost of requiring an increased normalised transmission rate \( R_n \).

IV. CONCLUSION

Reduced-complexity MCP structures associated with reduced back-haul data exchange requirements and employing the DSILNR precoding technique were proposed in the context of a unified system model. Our results demonstrated that the proposed structures achieve a higher outage rate than the SCP benchmark for all the MSs supported. Naturally, the proposed MCP structures cannot be expected to outperform the conventional MCP structure in all scenarios, but they tend to offer a different geographical rate distribution and different delays, while requiring reduced back-haul data exchange.

REFERENCES