Distributed Dimension Reduction for Distributed Massive MIMO C-RAN with Finite Fronthaul Capacity

Fred Wiffen†∗, Woon Hau Chin†∗ and Angela Doufexi∗
∗Department of Electrical & Electronic Engineering, University of Bristol, U.K.
†Bristol Research & Innovation Laboratory, Toshiba Europe Limited, U.K.
‡Viavi Solutions, Stevenage, U.K.
Email: fred.wiffen@bristol.ac.uk

Abstract—The use of a large excess of service antennas brings a variety of performance benefits to distributed MIMO C-RAN, but the corresponding high fronthaul data loads can be problematic in practical systems with limited fronthaul capacity. In this work we propose the use of lossy dimension reduction, applied locally at each remote radio head (RRH), to reduce this fronthaul traffic. We first consider the uplink, and the case where each RRH applies a linear dimension reduction filter to its multi-antenna received signal vector. It is shown that under a joint mutual information criterion, the optimal dimension reduction filters are given by a variant of the conditional Karhunen-Loeve transform, with a stationary point found using block co-ordinate ascent. These filters are then modified such that each RRH can calculate its own dimension reduction filter in a decentralised manner, using knowledge only of its own instantaneous channel and network slow fading coefficients. We then show that in TDD systems these dimension reduction filters can be re-used as part of a two-stage reduced dimension downlink precoding scheme. Analysis and numerical results demonstrate that the proposed approach can significantly reduce both uplink and downlink fronthaul traffic whilst incurring very little loss in MIMO performance.

Index Terms—Distributed MIMO, Massive MIMO, Dimension Reduction, Fronthaul Compression, MIMO Relay.

I. INTRODUCTION

DISTRIBUTED multi-user MIMO systems use multiple geographically distributed remote radio head (RRHs) to co-operatively serve multiple users within an extended coverage area. This distribution of the base station (BS) antennas provides macro-diversity to reduce variations in pathloss, improving uniformity of service compared to centralised MIMO configurations. It has seen increased interest in recent years due to the growing popularity of the cloud radio access network (C-RAN) architecture — where multiple RRHs are connected via fronthaul links to a shared central processor (CP), enabling joint MIMO processing of all signals.

However, the use of a large number of antennas brings a proportionate increase in the quantity of data that must be transferred over the fronthaul connections. This can present a challenge for practical systems which may rely on limited capacity fronthaul links based on wireless point-to-point, ethernet, or shared fibre [1].

In this work we consider the use of lossy distributed dimension reduction to reduce the amount of fronthaul data in distributed massive MIMO C-RAN systems. Whilst there has been interest in distributed dimension reduction techniques within the broader literature, its explicit application to distributed MIMO systems has yet to be studied in detail.

We propose to apply linear dimension reduction at each RRH to produce a low dimension effective MIMO system that exploits the inherent sparsity of massive MIMO to approximate the performance of the large-scale distributed MIMO system whilst reducing the quantity of data that must be transferred over the fronthaul connections. We show that this approach can be employed on both the uplink & downlink, and constitutes a bidirectional approach to fronthaul load reduction that is well suited to time division duplex (TDD) systems.

A. Related Work

Dimension reduction plays a role in many data compression/schemes, and exploits the fact that many high dimension signals have a sparse representation in an alternative signal basis and can therefore be represented, perfectly or approximately, using a reduced number of coefficients.

In [2] a number of problems related to finding reduced dimension distributed representations of an observed signal vector are investigated. It is shown that when linear dimension reduction is applied locally at a single node, the MMSE-optimal dimension reduction filter is given by the principal eigenvectors of the conditional covariance matrix at that node. They call this the ‘conditional’ Karhunen-Loeve transform due to its connection to the well known Karhunen-Loeve transform (the eigenvectors of the marginal covariance matrix [3]).

The work in [4] addresses the related problem of finding the reduced dimension signals from which some other correlated quantity is linearly estimated. The MMSE-optimal filters are shown to depend on various second order signal statistics, with a stationary point found using block coordinate ascent.
Some data reduction strategies using forms of dimension reduction have previously been proposed for distributed MIMO C-RAN systems. On the uplink, a simple dimension reduction strategy is proposed for reducing fronthaul data in [5], where each RRH takes simple unweighted sums of the signals received at different antennas in different time/frequency slots (i.e. binary dimension reduction filters). In our previous work in [6], we propose an uplink scheme that achieves dimension reduction at each RRH by matched filtering the received signal using a subset of the local user channel vectors. The channel vector subsets for all RRHs are selected jointly at the CP using full CSI to achieve a good overall signal representation.

On the downlink, the popular sparse beamforming approach first outlined in [7] can be seen as a form of dimension reduction for fronthaul data reduction. Dimension reduction is also applied at a network level in [8], where subsets of RRHs are deactivated to improve network energy efficiency. Despite its clear applicability, there has been little direct investigations into the optimisation of distributed dimension reduction for the distributed MIMO C-RAN context. Our work here seeks to begin to address this gap in the literature.

B. Paper Structure

The paper is structured as followed: Section II defines the system model for the distributed MIMO C-RAN uplink and downlink, before Section III provides an introductory discussion on the use of distributed dimension reduction on the distributed massive MIMO C-RAN uplink. Section IV addresses & analyses the problem of jointly optimising the dimension reduction filters under a joint mutual information criteria. An alternative fully decentralised approach in which each RRH calculates its own dimension reduction filter using partial CSI is proposed in Section V, before Section VI discusses some practical aspects of distributed dimension reduction. Section VII extends the dimension reduction concept to the distributed MIMO C-RAN downlink. Finally, Section VIII provides numerical results demonstrating the performance of the proposed approaches.

C. Notation

We use the notations x, x and X to represent scalar, vector and matrix quantities respectively. The conjugate, transpose and conjugate transpose of X is represented X*, XT and X†. [X]i,j represents the entry of X at row i, column j. Italics, X, are used in places to emphasise where a quantity is treated as a random variable. Trace, determinant, 2-norm and expectation are denoted tr(.), det(.), ‖.‖ and E{.}.

II. System Model

We consider a C-RAN system in which K single antenna users are co-operatively served by L geographically distributed multi-antenna RRHs, each equipped with M antennas. We focus specifically on massive MIMO C-RAN systems, in which the total number of base station antennas is significantly larger than the number of users being served, ML ≫ K.

The RRHs are connected to a single shared central processor (CP) via fronthaul connections, and uplink symbol detection & downlink symbol precoding are performed at the CP using signals from/for all RRHs.

We assume that the network operates in time division duplex mode, such that the uplink and downlink channels are reciprocal and can be estimated using uplink pilots transmitted by the users. A detailed discussion on CSI requirements is provided in Section VI. All numerical results provided in this work use the configurations described in Section VIII.

A. Distributed MIMO Uplink

The uplink signal received locally at RRH l, yl ∈ CM is

\[ y_l = \sqrt{\rho} H_l x + \eta_l, \]

where \( \rho \) is the user SNR, \( x = [x_1 \ x_2 \ \ldots \ x_K]^T \) contains the K unit variance transmit symbols, \( x_k \sim \mathcal{CN}(0, 1) \), and \( \eta_l \sim \mathcal{CN}(0, I_M) \) is receiver noise. The uplink channel matrix, \( H_l \in \mathbb{C}^{M \times K} \), is structured

\[ H_l = [\sqrt{p_1} h_{l,1} \ \ldots \ \sqrt{p_K} h_{l,K}] = \bar{H}_l P^{1/2} \]

where \( h_{l,k} \in \mathbb{C}^M \) is the channel between user k and RRH l, and \( p_k = ||P||_{k,k} \) the power control coefficient for user k (absorbed into the channel matrix to simplify notation). It is assumed that each \( H_l \) is full rank, i.e. \( r = \text{rank}(H_l) = \min(M, K) \), and that,

\[ \frac{1}{M} E\{||h_{l,k}||^2\} = \beta_{l,k}, \]

where \( \beta_{l,k} \) is the pathloss (slow fading) coefficient for the channel between user k and RRH l.

B. Distributed MIMO Downlink

For the downlink we consider a TDD configuration where the downlink channel is reciprocal to the uplink channel and the received signal at user k is

\[ y_k = \sum_{l=1}^{L} h_{l,k}^† s_l + \eta, \]

where \( s_l \in \mathbb{C}^M \) is the precoded signal transmitted by RRH l and \( \eta \sim \mathcal{CN}(0, 1) \) receiver noise. This can be written more compactly in vector form

\[ y = \sum_{l=1}^{K} \bar{H}_l^† s_l + \eta. \]

We assume that the transmissions from each RRH are subject to individual & identical power constraints,

\[ E\{s_l^† s_l\} \leq P. \]

III. Uplink Distributed Dimension Reduction

We begin by considering the distributed MIMO uplink. In our proposed scheme, shown in Figure 1, each multi-antenna RRH in the network receives an M-dimension signal,

\[ y_l = \sqrt{\rho} H_l x + \eta_l, \]
MIMO channel, be treated as the output of an equivalent reduced dimension signal produced at each RRH can then be seen as a generalisation of this for distributed systems, where \( A_l \) is allowed to take any value. MIMO signals to create a reduced dimension system that approximates the performance of the full dimension system.

At a basic level, we note that providing the overall signal dimension is greater than the number of users, \( NL \geq K \), then randomly chosen dimension reduction filters are generally sufficient to produce an overall reduced dimension channel, \( \mathbf{H}_c \), with full rank, and hence achieve the full MIMO multiplexing gain of \( K \). However, random filters do not capture the diversity gains provided by a system with a large overall excess of antennas, and can result in a poorly conditioned system.\(^1\)

On the other hand, it is clear that if the dimension reduction filters are optimised with respect to the channel, the reduced dimension MIMO system should be able to outperform a smaller distributed MIMO system with only \( N \) antennas deployed at each RRH – since this can be modelled as a special case of linear dimension reduction where \( A_l \) contains the first \( N \) columns of \( \mathbf{I}_M \). For example, we could deploy \( M \) antennas and choose \( A_l \) to be the ‘best’ \( N \) columns of \( \mathbf{I}_M \). This is the antenna selection method that has been widely studied for centralised MIMO systems and is known to be able to achieve a significant proportion of the system capacity using a reduced number of antennas.\(^2\) The work here can be seen as a generalisation of this for distributed systems, where \( A_l \) is allowed to take any value.

### IV. Filter Design Under Joint Mutual Information Criteria

Here, we consider the problem of finding the dimension reduction filters that maximise the joint mutual information between the reduced dimension signals and the user symbols,

\[
\begin{align*}
\text{maximise} & \quad I(\mathbf{z}_1, \ldots, \mathbf{z}_L; \mathbf{x}) \\
\text{subject to} & \quad A_l^\dagger A_l = \mathbf{I}_N.
\end{align*}
\]

This is equivalent to finding the dimension reduction filters that maximise the total amount of information that the set of reduced dimension signals jointly capture about the transmitted symbols, given by

\[
I(\mathbf{z}_1, \ldots, \mathbf{z}_L; \mathbf{x}) = \log_2 \det (\mathbf{I}_K + \rho \sum_{l=1}^L \mathbf{H}_l^\dagger A_l^\dagger A_l \mathbf{H}_l),
\]

It also represents the maximum sum rate of the reduced dimension MIMO system under optimal joint symbol detection. Unfortunately \(^1\) is non-convex in the \( A_l \), making find a global maximum challenging. Here, we instead show that a stationary point can be found using block coordinate ascent. We begin by considering the problem of finding the optimal dimension reduction filter, \( A_l \) at RRH \( l \), under the assumption that the other dimension reduction filters, \( A_1, \ldots, A_{l-1}, A_{l+1}, \ldots, A_L \), are fixed.

\(^1\) Using the QR decomposition, an arbitrary filter, \( \mathbf{A}_l \in \mathbb{C}^{M \times N} \), with rank \( N \) can be represented, \( \mathbf{A}_l = \mathbf{A}_l \mathbf{R}_l \), where \( \mathbf{A}_l \in \mathbb{C}^{M \times N} \) is semi-orthogonal and \( \mathbf{R}_l \in \mathbb{C}^{N \times N} \) is invertible. Since \( \mathbf{R}_l \) is invertible it does not affect the information content of \( \mathbf{z}_l = \mathbf{A}_l^\dagger \mathbf{y}_l \).

\(^2\) For the independent Rayleigh fading case it can be shown that using random dimension filters is equivalent to simply reducing the number of antennas at each RRH from \( M \) to \( N \).
A. Optimisation at Single RRH

Using the chain rule, the joint mutual information can be expanded

\[ I(z_1, \ldots, z_L; x) = I(z_f^i; x) + I(z_l; x|z_f^i) \] (15)

where \( z_f^i = \{z_1, \ldots, z_{i-1}, z_{i+1}, \ldots, z_L \} \). In this expansion the first term, \( I(z_f^i; x) \), represents the joint mutual information when the CP has access to the reduced dimension signals from the other \( L-1 \) RRHs. The second term, \( I(z_l; x|z_f^i) \), represents the conditional information contribution of RRH \( l \) – i.e. the increase when the CP also has access to \( z_l \). Applying the well-known matrix determinant lemma, this can be shown to be

\[ I(z_l; x|z_f^i) = \log_2 \det \left( I_N + \rho A_l^H Q_l H_l^H A_l \right), \] (16)

where

\[ Q_l = \left( I_K + \rho \sum_{j \neq l} H_j^H A_j A_j^H H_j \right)^{-1}. \] (17)

Since the first term in (15) does not depend on \( A_l \), the optimal dimension reduction filter for RRH \( l \) is the solution to

\[
\begin{align*}
\text{maximise} \quad & \log_2 \det \left( I_N + \rho A_l^H Q_l H_l^H A_l \right) \\
\text{subject to} \quad & A_l^H A_l = I_N
\end{align*}
\] (18)

By the Poincare separation theorem [10], this has a single global maximum that is achieved when the columns of \( A_l \) are the eigenvectors corresponding to the \( N \) largest eigenvalues of \( H_l Q_l H_l^H \),

\[ A_l = N \text{ principal eigenvectors of } H_l Q_l H_l^H. \] (19)

The optimal dimension reduction filter at RRH \( l \) contains the \( N \) principal eigenvectors of the conditional covariance matrix \( \mathbb{E}\{y_i y_j^H | z_f^i \} = I_M + \rho H_l Q_l H_l^H \), conditioned with respect to the signals supplied by the other RRHs. This is an instance of the conditional Karhunen-Loeve transform (CKLT) as previously defined in [12]. For the special case of a single RRH \( (L = 1) \) we have \( Q_l = I_K \), and the dimension reduction filter is the Karhunen-Loeve transform (KLT) – the filter that also maximises signal power, \( \mathbb{E}\{||z_l||^2\} \).

B. Block Coordinate Ascent

Since the CKLT achieves the global maximum to (18), the filters can be jointly optimised using block coordinate ascent by successively updating the CKLT at each RRH in turn. Since the objective function (14) is bounded and increases monotonically with each update, convergence to a stationary point is assured. In our simulations here we initialise the coordinate ascent using the KLT filters.

C. Insights

The loss in information due to dimension reduction is

\[ \Delta = I(y_1, \ldots, y_L; x) - I(z_1, \ldots, z_L; x) \] (20)

\[ = \log_2 \frac{\det(I_K + \rho \sum_{j \neq l} H_j^H H_j)}{\det(I_K + \rho \sum_{i=1}^L H_i^H D_i H_i)}. \] (21)

 Providing the reduced dimension system is full rank, the overall loss is bounded at high SNR,

\[ \lim_{\rho \to \infty} \Delta = \log_2 \frac{\det(\sum_{i=1}^L H_i^H H_i)}{\det(\sum_{i=1}^L H_i^H D_i H_i)}, \] (22)

and represents a vanishing fraction of the joint mutual information captured by the reduced dimensions signals. This can be seen in Figure 2 for a system with \( L = 4, M = 8, K = 8 \). Note that with \( N \geq 3 \), the overall loss in information due to applying dimension reduction is very small. There is a clear benefit to using a larger number of antennas at each RRH and then applying dimension reduction compared to simply deploying a reduced number of antennas.

V. FULLY DECENTRALISED DIMENSION REDUCTION USING PARTIAL CSI

The method in Section IV for jointly optimising the uplink dimension reduction filters requires full network CSI, and therefore in practice must be implemented centrally at the CP. We now investigate an alternative decentralised approach, where each RRH locally calculates its own uplink filter using full knowledge of its own channel matrix, \( H_l \), whilst treating the channel matrices for the other \( L-1 \) RRHs as unknown random quantities, \( H_j \).

We can adapt the joint mutual information maximisation method from Section IV by choosing \( A_l \) instead to maximise the expected value of the quantity

\[
\begin{align*}
\text{maximise} \quad & \mathbb{E}\{\log_2 \det( I_N + \rho A_l^H Q_l H_l^H A_l )\} \\
\text{subject to} \quad & A_l^H A_l = I_N,
\end{align*}
\] (23)

where \( Q_l \) is now a random variable that depends on the random channel and filter realisations at the other RRHs

\[ Q_l = \left( I_K + \rho \sum_{j \neq l} H_j^H A_j A_j^H H_j \right)^{-1}. \] (24)

In this expression the optimal \( A_l \) depends on both the other \( A_j \) and the probability distributions for the \( H_j \), making finding
The dimension reduction filters can then be calculated as shown in (26). This lower bound is established in two stages:

- First, we use the lower bound in (25) to eliminate the dependence on the dimension reduction filters at the other RRHs, by noting that $A_l A_l^\dagger$ is an orthogonal projection. This is equivalent to designing the dimension reduction filter at RRH $l$ under the condition that the other $L-1$ RRHs each supply their full signal to the CP.

- Second, we exploit the convexity of (25) and apply Jensen’s inequality to produce an objective function that depends only on the channel statistics $E\{H_j^\dagger H_j\}$.

Under the mild assumption that the different user channel realisations vary independently, $E\{h_{ij,k}^*) h_{ij,j}\} = 0$, $E\{H_j^\dagger H_j\} = \Psi_j$ is a diagonal matrix that depends only on the user power control and pathloss coefficients,

$$\Psi_j = p_k E\{|h_{ij,k}|^2\} = p_k^2 \beta_{ij,k} M$$  

(27)

The dimension reduction filters can then be calculated as

$$A_l = N \text{ princ. eigenv. of } H_l (I_K + \rho \sum_{j \neq l} \Psi_j)^{-1} H_l^\dagger.$$  

(28)

We refer to this filter as the D-CKLT. In contrast with the CKLT, with the D-CKLT dependencies between RRHs may vary by orders of magnitude due to the physical distribution of users and RRHs.

After each RRH calculates and applies its dimension reduction filters, the CP jointly detects the user symbols using the dimension reduction filter. Numerical results presented in Section VIII show that the use of decentralised filter design often incurs only a small performance sacrifice.

VI. PRACTICAL ASPECTS

In practical mobile channels with finite channel coherence time the dimension reduction filters must be regularly updated, and hence the computational and signalling overheads associated with distributed dimension reduction must be considered.

A. Computational Complexity

The centralised approaches to filter design are iterative in nature, with each iteration requiring a $K \times K$ matrix inversion with complexity $O(K^3)$, an $M \times M$ eigendecomposition, complexity $O(M^3)$, and some matrix multiplications. Simulations indicate that the block coordinate procedure typically converges with a small number ($\sim 3$) of updates at each RRH.

In contrast, the decentralised approach has a closed form solution, requiring a single $M \times K$ singular value decomposition with complexity $O(M K^2)$ per RRH. This computation is performed locally at the RRHs, meaning the decentralised approach requires the RRHs to have enhanced computational capabilities compared to the centralised approach.

B. Fronthaul CSI Overheads

In the centralised approach, the CP requires full network CSI for calculating the dimension reduction filters, which must then be transferred back to the RRHs. Assuming CSI is initially obtained at the RRHs through uplink pilot symbols, additional data must be transferred both to and from the CP.

One benefit of the decentralised approach is the reduction in overheads, since only the reduced dimension MIMO channels, $H_l^{RD}$, need to be transferred from RRHs to CP (with only changes in large scale fading communicated back). The number of coefficients that must be transferred, per channel coherence block, for each scheme are shown in Table I.

Figure 3 shows the reduction in total fronthaul load (including CSI overheads) in terms of mean coefficients per channel use, for a system with $K = 8, M = 8$. When the coherence block is small, CSI overheads account for a larger proportion of the fronthaul load, and hence the benefit from using decentralised dimension reduction is greater.

C. Imperfect CSI

In practical systems, the CSI estimated using uplink pilots is imperfect. Assuming the use of MMSE channel estimation, this can be accounted for by modelling the estimated channel between user $k$ and RRH $l$, $\hat{h}_{l,k} \in \mathbb{C}^M$, as

$$\hat{h}_{l,k} = h_{l,k} - e_{l,k}$$  

(29)

where $h_{l,k}$ is the propagation channel and $e_{l,k}$ the random orthogonal channel error, with covariance $E\{e_{l,k} e_{l,k}^\dagger\} = C_{l,k}$.

Following the reasoning in [12], the uplink received signal at RRH $l$ can be modelled as

$$y_l = \sqrt{\rho} H_l x + \sqrt{\rho} E_l x + \eta_l = \sqrt{\rho} \hat{H}_l x + \nu_l$$  

(30)

where $\nu_l \in \mathbb{C}^M$ is a noise term that includes the effect of both receiver noise and orthogonal channel estimation error, with

$$E\{\nu_l \nu_l^\dagger\} = \Omega_l + \rho \sum_{k=1}^K C_{l,k}.$$  

(31)

We then apply a ‘whitening’ transform to produce an equivalent MIMO system

$$\tilde{y}_l = \Omega_l^{-1/2} y_l = \sqrt{\rho} \hat{H}_l x + \tilde{\eta}_l$$  

(32)

where $\tilde{\eta}_l \sim \mathcal{CN}(0, I_M)$ and $E\{\tilde{\eta}_l \tilde{\eta}_l^\dagger\} = I_M$. The proposed dimension reduction methods may now be applied to $\tilde{y}_l$, and, following the reasoning in [12], the uplink rates calculated using standard expressions and these equivalent channels represent expected achievable rates under imperfect CSI.

| Method                | RRH to CP             | CP to RRH                  |
|-----------------------|-----------------------|----------------------------|
| Standard Distributed MIMO | $H_l, MK$ coeff.      | –                          |
| Centralised Dimension Red. | $H_l, MK$ coeff.      | $A_l, MN$ coeff.            |
| Decentralised Dimension Red. | $H_l^{RD}, NK$ coeff. | –                          |
\[
\mathbb{E}\left\{ \log_2 \det \left( \mathbf{I}_N + \rho \mathbf{A}_l^H \mathbf{H}_l \mathbf{A}_l \right) \right\} \geq \mathbb{E}\left\{ \log_2 \det \left( \mathbf{I}_N + \rho \mathbf{A}_l^H \mathbf{H}_l \left( \mathbf{I}_K + \rho \sum_{j \neq l} \mathbf{H}_j^H \mathbf{H}_j \right)^{-1} \mathbf{H}_l^H \mathbf{A}_l \right) \right\},
\]

\[
\geq \log_2 \det \left( \mathbf{I}_N + \rho \mathbf{A}_l^H \mathbf{H}_l \left( \mathbf{I}_K + \rho \sum_{j \neq l} \mathbb{E}\{ \mathbf{H}_j^H \mathbf{H}_j \} \right)^{-1} \mathbf{H}_l^H \mathbf{A}_l \right). \tag{26}
\]

The second, outer, linear precoding stage is then applied to the \(N\)-dimension signal at each RRH to produce an \(M\)-dimension transmit signal,

\[
s_l = \mathbf{B}_l z_l. \tag{34}
\]

The vector of received signals at the \(K\) users is then

\[
y = \sum_{l=1}^L \mathbf{H}_l^H s_l + \eta = \sum_{l=1}^L \mathbf{H}_l^H \mathbf{B}_l \mathbf{F}_l x + \eta. \tag{35}
\]

In principle \(\mathbf{F}_l\) and \(\mathbf{B}_l\) can be directly optimised according to some downlink performance criteria. We instead consider a simpler approach where the uplink dimension reduction filters are re-used in the outer precoding stage, i.e

\[
\mathbf{B}_l = \mathbf{A}_l. \tag{36}
\]

This is attractive from a practical perspective for TDD networks serving the same users on the uplink and downlink, as it removes the computation and signalling overheads associated with generating additional outer precoding matrices.

For semi-orthogonal \(\mathbf{B}_l\), the per-RRH transmit power depends only on the inner precoders,

\[
\mathbb{E}\{s_l^H s_l\} = \text{tr} \left( \mathbf{F}_l^H \mathbf{F}_l \right) \leq P, \tag{37}
\]

These inner precoders can be thought of as acting on the set of reciprocal reduced dimension downlink channels,

\[
\mathbf{H}_l^H \mathbf{B}_l = \mathbf{P}^{-1/2} \tilde{\mathbf{H}}_l^{DL}, \tag{38}
\]

and can be designed using standard downlink precoding methods, e.g. Moore-Penrose zero-forcing (ZF),

\[
\mathbf{F}_l = \mathbf{B}_l^H \left( \sum_{j=1}^L \mathbf{B}_j^H \mathbf{H}_j \right)^{-1} \Gamma^{1/2}, \tag{39}
\]

where \(\Gamma = \text{diag}(\rho_k)\). Under perfect CSI, the downlink rate of user \(k\) is then

\[
\mathcal{R}_k^{DL} = \log_2 (1 + \rho_k). \tag{40}
\]

The \(\rho_k\) can then be allocated to maximise some objective, such as max-min user rate, subject to the power constraints in (37).

VIII. NUMERICAL RESULTS

We now provide a selection of numerical results demonstrating the benefits of the proposed dimension reduction approach. These are obtained via Monte Carlo simulation for a dense urban environment, where the \(K\) users and \(L\) RRHs are distributed randomly within a \(200 \text{ m} \times 200 \text{ m}\) coverage area.

The user channels follow complex normal independent fading

\[
h_{l,k} \sim \mathcal{CN}(0, \beta_{l,k} \mathbf{I}_M), \tag{41}
\]

and can be designed using standard downlink precoding methods, e.g. Moore-Penrose zero-forcing (ZF),
with a log-distance shadow fading path loss model,
\[ \beta_{l,k} = 147.5 - 20 \log_{10} f_c - 10\gamma \log_{10} d_{l,k} + \psi \]  
(42)

where \( f_c \) is the carrier frequency in GHz, \( \gamma \) the path loss exponent, \( d_{l,k} \) the distance between RRH \( l \) and user \( k \) in metres, and \( \psi \sim CN(0,\sigma_{\psi}^2) \) the log-normal shadow fading term in dB. Simulation parameters are summarised in Table II, chosen based on the results given in [13].

On the uplink, user power control is applied based on slow fading coefficients, such that each user has unity overall average power adjusted channel gain,
\[ p_k \sum_{l=1}^{L} \frac{\beta_{l,k}}{L} = 1. \]  
(43)

All rates are averaged over both user & RRH locations and random channel realisation. As a reference, the proposed schemes are compared to dimension reduction using joint antenna selection (adapted from [14]), and simple antenna reduction (i.e. where each RRH has only \( N \) antennas).

A. Uplink Sum Rate under MMSE-SIC Detection

Figure 5 shows the sum rates that can be achieved for varying reduced signal dimensions, \( N \), in a system with \( L = 4 \) RRHs, \( K = 8 \) users, \( M = 8 \) antennas and MMSE-SIC detection. At all reduced signal dimensions (\( N < 8 \)), the proposed centralised & decentralised dimension reduction methods achieve significantly higher sum rates than can be achieved by either antenna selection or antenna reduction.

For example, under either of the proposed schemes, reducing the signal dimension at each RRH from \( N = M = 8 \) to \( N = 3 \) results in only \( \sim 5\% \) loss in cell sum rate, whilst reducing signal fronthaul data by 63%. This compares favourably to simply reducing the number of antennas at each RRH to \( M = 3 \), which incurs a \( \sim 30\% \) loss in sum rate.

Figure 6 shows the case where the signal dimension at each RRH, \( N \), is kept constant but the number of antennas deployed at each RRH, \( M \), is increased, with centralised (CKLT) dimension reduction. Despite the fixed dimension signal, the cell sum rate is always increased by increasing the number of antennas – demonstrating the benefit of operating in the 'massive' regime even when fronthaul capacity is limited.

B. Uplink User Rates under Linear MMSE Detection

Figure 7 shows the mean user rates under linear MMSE detection for the two proposed dimension reduction techniques, \( K = 8, L = 4, M = 8 \). Note that for \( N = 1 \) the overall equivalent reduced dimension MIMO system is degraded (rank deficient), and linear detection cannot effectively isolate the user streams, resulting in poor performance. For \( N > 2 \) the reduced dimension a high proportion of the full user rate is achieved.

C. Uplink User Outage Performance

Figure 8 shows the user outage performance under linear MMSE detection for varying numbers of antennas (top), and fixed number of antennas (\( M = 8 \)) with dimension reduction.
applied. The benefits of deploying excess antennas and then performing dimension reduction to reduce fronthaul load are again clear for both centralised and decentralised cases. For \( N = 5 \), the outage performance is close to that of the full dimension system. A theoretical analysis of the outage performance remains as an interesting topic for future work.

**D. Scalability**

Figure 9 shows that as the network becomes denser – by adding more users and RRHs – dimension reduction continues to capture a very high proportion of the full dimension MIMO system throughput. We note that whilst there is some performance penalty compared to the centralised approach, the fully decentralised approach (D-CKLT) has a much reduced computational load that is distributed between the RRHs, and therefore represents an attractive scalable solution for use in large C-RAN networks.

**E. Imperfect CSI**

We now consider the case of imperfect (noisy) CSI at the RRHs. Figure 10 shows the achievable sum rate, based on the model discussed in Section VII assuming the MIMO channels are estimated from orthogonal pilot signals with SNR \( \rho_{\text{CSI}} \). The performance of the reduced dimension system is close to that of the full dimension system irrespective of the CSI quality.

**F. Downlink Two-Stage Precoding**

As discussed in Section VII, the dimension reduction approach to fronthaul data reduction readily extends to the downlink in the form of a two-stage precoding scheme. Figure 11 illustrates that similar performance benefits are afforded as on the uplink – establishing dimension reduction as an effective bidirectional approach to fronthaul data reduction for massive MIMO systems.

**IX. Conclusion**

This paper has provided an investigation into the use of dimension reduction for reducing fronthaul data in massive MIMO C-RAN systems. Centralised and decentralised methods for designing appropriate dimension reduction filters have been proposed, and a combination of analysis and numerical results used to show that dimension reduction, applied locally at each RRH, can significantly reduce fronthaul traffic without incurring significant loss in cell throughput. Furthermore, we...
have shown that this dimension reduction is readily extended to the downlink of MIMO systems, establishing it as an effective and practical bidirectional approach to fronthaul data reduction that is particularly well suited to TDD networks.

Overall, our work shows that it is possible to use dimension reduction to produce an equivalent reduced dimension distributed MIMO system that preserves the performance benefits of operating in the ‘massive’ regime with a large overall excess of BS antennas. Future work should build on these contributions to further analyse the behaviour and performance of reduced dimension MIMO systems, and investigate & implement dimension reduction schemes in practical systems.

ACKNOWLEDGMENT
The authors would like to thank Toshiba Europe Ltd. and EPSRC for financially supporting this work.

REFERENCES
[1] M. Jaber, M. A. Imran, R. Tafazolli, and A. Tukmanov, “5G Backhaul Challenges and Emerging Research Directions: A Survey,” IEEE Access, vol. 4, pp. 1743–1766, 2016.
[2] M. Gastpar, P. L. Dragotti, and M. Vetterli, “The Distributed Karhunen–Loève Transform,” IEEE Transactions on Information Theory, vol. 52, no. 12, pp. 5177–5196, 2006.
[3] V. Goyal, “Theoretical foundations of transform coding,” IEEE Signal Processing Magazine, vol. 18, no. 5, pp. 9–21, 2001.
[4] I. D. Schizas, G. B. Giannakis, and Z.-Q. Luo, “Distributed Estimation Using Reduced-Dimensionality Sensor Observations,” IEEE Transactions on Signal Processing, vol. 55, no. 8, pp. 4284–4299, 2007.
[5] H. Ma, B. Wang, and K. J. R. Liu, “Distributed Signal Compressive Quantization and Parallel Interference Cancellation for Cloud Radio Access Network,” IEEE Transactions on Communications, vol. 66, no. 9, pp. 4186–4198, 2018.
[6] F. Wiffen, M. Z. Bocus, A. Doufexi, and W. H. Chin, “MF-based Dimension Reduction Signal Compression for Fronthaul-Constrained Distributed MIMO C-RAN,” in 2020 IEEE Wireless Communications and Networking Conference (WCNC), 2020, pp. 1–8.
[7] B. Dai and W. Yu, “Sparse Beamforming and User-Centric Clustering for Downlink Cloud Radio Access Network,” IEEE Access, vol. 2, pp. 1326–1339, 2014.
[8] C. Pan, H. Zhu, N. J. Gomes, and J. Wang, “Joint Precoding and RRH Selection for User-Centric Green MIMO C-RAN,” IEEE Transactions on Wireless Communications, vol. 16, no. 5, pp. 2891–2906, 2017.
[9] S. Sanayei and A. Nosratinia, “Antenna selection in MIMO systems,” IEEE Communications Magazine, vol. 42, no. 10, pp. 68–73, 2004.
[10] R. Bellman, Introduction to matrix analysis. SIAM, 1997.
[11] K. K. Kim, “A Note on the Convexity of \( \log \det(1 + KX^{-1}) \) and its Constrained Optimization Representation,” CoRR, vol. abs/1509.00777, 2015. [Online]. Available: http://arxiv.org/abs/1509.00777
[12] B. Hassibi and B. Hochwald, “How much training is needed in multiple-antenna wireless links?” IEEE Transactions on Information Theory, vol. 49, no. 4, pp. 951–963, 2003.
[13] S. Sun, T. S. Rappaport, T. A. Thomas, A. Ghosh, H. C. Nguyen, I. Z. Kovács, I. Rodriguez, O. Koymen, and A. Partyka, “Investigation of prediction accuracy, sensitivity, and parameter stability of large-scale propagation path loss models for 5g wireless communications,” IEEE Transactions on Vehicular Technology, vol. 65, no. 5, pp. 2843–2860, 2016.
[14] M. Gharavi-Alkhansari and A. Gershman, “Fast antenna subset selection in MIMO systems,” IEEE Transactions on Signal Processing, vol. 52, no. 2, pp. 339–347, 2004.