

# Partial Joint Processing with Efficient Backhauling in Coordinated MultiPoint Networks

Tilak Rajesh Lakshmana\*, Carmen Botella†, Tommy Svensson\*

\*Signals and Systems, Chalmers University of Technology, Gothenburg, Sweden,

†Institute of Robotics and Information & Communication Technologies (IRTIC), Universitat de València, València, Spain.

{tilak, tommy.svensson}@chalmers.se, Carmen.Botella@uv.es

**Abstract**—Joint processing between base stations is a promising technique to improve the quality of service to users at the cell edge, but this technique poses tremendous requirements on the backhaul signaling capabilities. Partial joint processing is a technique aimed to reduce feedback load, in one approach the users feed back the channel state information of the best links based on a channel gain threshold mechanism. However, it has been shown in the literature that the reduction in the feedback load is not reflected in an equivalent backhaul reduction, unless additional scheduling or precoding techniques are applied. The reason is that reduced feedback from users yields sparse channel state information at the Central Coordination Node. Under these conditions, existing linear precoding techniques fail to remove the interference and reduce backhaul, simultaneously, unless constraints are imposed on scheduling. In this paper, a partial joint processing scheme with efficient backhauling is proposed, based on a stochastic optimization algorithm called particle swarm optimization. The use of particle swarm optimization in the design of the precoder promises efficient backhauling with improved sum rate.

**Index Terms**—Joint Processing, Zero Forcing, Backhaul load reduction, Particle Swarm Optimization, Stochastic Optimization

## I. INTRODUCTION

Future cellular communication systems tend to have a frequency reuse factor of one, causing intercell interference and reducing user experience close to the cell-edge. Joint Processing (JP) between Base Stations (BSs) is one of the techniques that falls in the framework of Coordinated MultiPoint (CoMP) transmission [1]. In downlink JP, the user receives its data from multiple coordinating BSs.

In a typical Centralized Joint Processing (CJP) approach, the cluster of BSs jointly coordinates and transmits the data to the intended user, without causing interference to other users. This poses users to feed back the Channel State Information (CSI) of all the BSs in the cluster to their serving BS. Then, the CSI needs to be forwarded over the backhaul towards the Central Coordination Node (CCN) to precancel the interference via BeamForming (BF) and power allocation. This non-casual availability of CSI at the CCN for interference avoidance can be treated as casual for a stationary channel, but needs regular

updates for non-stationary channels. Nevertheless, the need for the CSI being available at the CCN and for sending the precoding weights and user data from the CCN to the corresponding BSs puts tremendous requirements on backhauling. To alleviate them, clusters of BSs are usually arranged. The clustering techniques can be divided into network-centric or user-centric, depending on where the clustering decision is carried out. To this end, Partial Joint Processing (PJP) has been proposed to reduce the CSI feedback load [2].

PJP can be viewed as a general framework for feedback and backhaul reduction. In the particular approach considered in this paper, a CCN or the serving BS might instruct the User Equipments (UEs) to report the CSI of the links in the cluster of BSs whose channel gain fall within an active set threshold or window, relative to their best link (usually the serving BS). With this PJP scheme, feedback load is reduced, as CSI of only a subset of BSs is fed back per UE. This subset is also referred to as an active set. Note that this user-centric clustering technique results in the formation of overlapping active set of BSs for each user, and the user should preferably only receive its data from the set of BSs included in its active set. The CSI being available at the CCN are marked as active links and those that are not available (or not reported) are marked inactive. The CCN forms an aggregate channel matrix based on these active and inactive links for interference avoidance. As a result, the aggregated channel matrix is now sparse, due to the reduced CSI feedback giving rise to inactive links, which are modeled as zeros.

In JP, linear BF techniques such as Zero Forcing (ZF) can be used for interference avoidance, as long as the aggregated channel matrix is well conditioned for inversion. It has been shown in the literature that the reduction in the CSI feedback load is not necessarily reflected in an equivalent BF backhaul reduction, unless additional scheduling or precoding techniques are applied [3], or unless the aggregated channel matrix is diagonal or block-diagonal. In other words, when calculating the ZF BF based on the sparse aggregated channel matrix, one inactive link may be mapped into a non-zero BF weight for that link. This causes unnecessary backhauling, since the UE has reported that link as inactive and that BS is then outside the active set of that UE, i.e., the resources at this BS can be used to serve other UEs. A brute force approach would be that the CCN might resort to nulling the BF weights where the links are inactive, but this might lead to inefficient power allocation and increased interference, resulting in reduced sum rate of the system in the cluster area.

This work has been supported by The Swedish Agency for Innovation Systems (VINNOVA) and by the EU FP7 project INFSO-ICT-247223 ARTIST4G. C. Botella's work is supported by the Spanish MEC Grants CONSOLIDER-INGENIO 2010 CSD2008-00010 "COMONSENS" and COSIMA TEC2010-19545-C04-01. The authors would like to thank associate professor Ivar Gustafsson for the discussion on sparse matrix theory and Jingya Li for the power allocation discussion. The computations were performed on C<sup>3</sup>SE computing resources.

To the best of our knowledge, this problem has only been addressed in [3], where two solutions are proposed, one based on scheduling (MAC layer approach) and the other based on ZF precoding (PHY layer approach). In the scheduling solution of [3], a suitable set of UEs are selected for transmission such that the sparse aggregated channel matrix available at the CCN is block-diagonal and hence, invertible. These suitable sets are formed by arranging disjoint active sets in each time slot, i.e., each BS involved in serving a UE belongs to only one active set. The drawback of this approach is that, in each time slot, a given set of disjoint BS active sets is selected for transmission; if the UEs prefer services from the same set of BSs they need to wait for their turn to be served in a TDMA fashion. Fairness is guaranteed but at the cost of UEs needing to wait for a long time. In case of the ZF precoding solution in [3], no constraints are assumed on scheduling. To reduce the backhaul load, the zeros in the sparse aggregated channel matrix are mapped to the aggregated BF matrix. The interference is reduced by formulating this as a constrained optimization problem. The proposed solution needs a well constructed aggregated channel matrix and hence, it is heavily dependent on scheduling. On the other hand, there is no linear technique existing in the literature that can invert the aggregated channel matrix with zeros (inactive links) and preserve these zeros in the transposed version of the inverse, when the aggregated channel matrix is not diagonal or block-diagonal.

In this paper, Particle Swarm Optimization (PSO), a stochastic optimization method, is proposed as a tool to design a BF that can achieve a backhaul reduction in terms of zero BF weights equivalent to the CSI feedback reduction. Compared to the ZF precoding solution in [3], this technique works on sparse aggregated channel matrices, without any constraint on scheduling. PSO has already been shown to obtain the optimal multiuser MIMO linear precoding vector, where the objective function of the PSO was to maximize the system capacity [4]. Whereas, in our paper, the PSO is used in a multicell scenario performing PJP CoMP with perfect CSI with the main objective of minimizing interference.

The paper is organized as follows, in section II the system model is described. Section III introduces PSO and discusses how the BF weights are treated as particles. Simulation results are presented in section IV, and section V concludes the paper. The notation used in this paper is summarized in the footnote below.

## II. SYSTEM MODEL

Consider a cluster of  $K$  single antenna BSs involved in the downlink transmission to  $M$  single antenna UEs located at the cluster center, as shown in Figure 1. Any form of intercluster

Notation: Boldface upper-case letters denote matrices,  $\mathbf{X}$ , boldface lower-case letters denote vectors,  $\mathbf{x}$ , and italics denote scalars,  $x$ . The  $\mathbb{C}^{m \times n}$  is a complex valued matrix of size  $m \times n$ . The  $(\cdot)^H$  is the conjugate transpose of a matrix. The  $\|\cdot\|_F$  is the Frobenius norm,  $\text{OffDiag}(\mathbf{X})$  is an operation on the matrix  $\mathbf{X}$  that sets the elements in the main diagonal to zero.  $\mathbf{X}(i, j)$  is the  $(i, j)$ th element of matrix  $\mathbf{X}$ .  $\text{vec}(\mathbf{X})$  is the vector of stacked columns of matrix  $\mathbf{X}$  and  $\otimes$  denotes the Kronecker product.  $\Re\{\mathbf{X}(i, j)\}$  and  $\Im\{\mathbf{X}(i, j)\}$  are the real part and the imaginary parts of  $\mathbf{X}(i, j)$ .

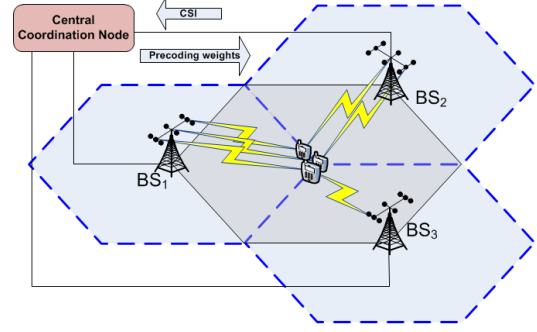


Figure 1. The cluster layout, the hexagon in the middle denotes the cluster area under consideration.

interference affecting the demodulation of signals at the UE is assumed to be negligible and is thus neglected. With a frequency reuse factor of one in this layout, the transmission to a UE will cause interference to other UEs. Assuming CJP, the discrete time signal received at  $M$  UEs,  $\mathbf{y} \in \mathbb{C}^{M \times 1}$  is

$$\mathbf{y} = \mathbf{H}\mathbf{W}\mathbf{x} + \mathbf{n}, \quad (1)$$

where  $\mathbf{H} \in \mathbb{C}^{M \times K}$  is the aggregated channel matrix of the form  $[\mathbf{h}_1^T \mathbf{h}_2^T \dots \mathbf{h}_M^T]^T$ ,  $\mathbf{h}_m \in \mathbb{C}^{1 \times K}$  is the channel from all the BSs in the cluster to the  $m$ th UE,  $\mathbf{W} \in \mathbb{C}^{K \times M}$  is the aggregated BF matrix of the form  $\mathbf{W} = [\mathbf{w}_1 \mathbf{w}_2 \dots \mathbf{w}_M]$ ,  $\mathbf{w}_m \in \mathbb{C}^{K \times 1}$  is the BF for the  $m$ th UE,  $\mathbf{x} \in \mathbb{C}^{M \times 1}$  is the transmitted symbols to the  $M$  UEs, and  $\mathbf{n}$  is the spatially and temporally white receiver noise with variance  $\sigma^2$ , and it is uncorrelated with the transmitted symbols.

When CJP is used, the CCN has a full channel matrix  $\mathbf{H}$ . In the literature, a ZF BF matrix  $\mathbf{W}$  is obtained by taking the right inverse of  $\mathbf{H}$ ,

$$\mathbf{W} = \mathbf{H}^H (\mathbf{H}\mathbf{H}^H)^{-1}. \quad (2)$$

CJP can be seen as a particular case of PJP when the threshold is high, such that all links are active for a given UE. For convenience, at the CCN, the active and inactive links can be represented as an active set, a binary matrix of size  $[M \times K]$ , whose  $(m, k)$ th element represents the  $(m, k)$ th link between the  $m$ th user and the  $k$ th BS. These elements take the value '1' and '0' representing links whose CSI is available (active) and not available (inactive), respectively [2]. Few links are active in some scenarios, e.g., small values of the active set threshold result in a sparse aggregated channel matrix  $\tilde{\mathbf{H}}$  at the CCN. If  $\tilde{\mathbf{H}}$  is invertible, the  $\tilde{\mathbf{W}}$  thus formed may not have zeros at places where needed, i.e., a UE will receive its data from BSs outside its active set, corresponding to its inactive BSs. For example, say  $UE_1$  reported CSI for  $BS_1$ ,  $BS_3$  and not for  $BS_2$  i.e.,  $BS_2$  falls outside the active set of  $UE_1$ , hence,  $\tilde{\mathbf{H}}(1, 2) = 0$ . The CCN having the CSI of all the UEs in that cluster tries to invert the aggregated channel matrix to obtain the BF weights. These weights are only needed at  $BS_1$  and  $BS_3$  for  $UE_1$ , but they might show up at  $BS_2$  for  $UE_1$ , due to the behavior of the pseudo-inverse involved with ZF, i.e.,  $\tilde{\mathbf{W}}(2, 1) \neq 0$ . This is highly undesirable as it results in extra and unnecessary backhaul load on the

cluster. It should be pointed out that  $BS_2$  being inactive is not involved in JP for serving  $UE_1$  but  $BS_2$  can be involved in serving other UEs. A BS that is not involved in serving any UE need not be considered in this setup at all. Due to the overlapping clusters formed with PJP, the subset of BSs reported by the UEs differ for a given frequency/time resource. Hence, a BS serving only one UE at the cluster center, should be included in the precoding design as this UE is sharing the same frequency/time resource and the interference thus generated needs to be accounted.

To realize the gains of the active set based PJP scheme, the problem at the CCN under a ZF assumption, is two fold: firstly, invert a sparse matrix and secondly, obtain null BF weights in the correct places. Hence, BF is an important challenge to be realized, especially without the need to have any special constraints on scheduling.

Every BS involved in JP has a maximum per-BS power constraint of  $P_{\max}$ . The precoding matrix  $\bar{\mathbf{W}}$  is realized such that at least one of the BSs can transmit at maximum power as defined in

$$\bar{\mathbf{W}} = \left( \sqrt{P_{\max} / \left( \max_{k=1, \dots, K} \|\tilde{\mathbf{W}}^k\|_F^2 \right)} \right) \cdot \tilde{\mathbf{W}}, \quad (3)$$

where  $\tilde{\mathbf{W}}^k$  stands for the BF weights of the  $k$ th BS towards the  $M$  users.

The Signal to Interference plus Noise Ratio (SINR) for the  $m$ th UE is given as

$$\text{SINR}_m = \frac{\|\mathbf{h}_m \bar{\mathbf{w}}_m\|^2}{\sum_{j=1, j \neq m}^M \|\mathbf{h}_m \bar{\mathbf{w}}_j\|^2 + \sigma^2}, \quad (4)$$

and the sum rate per cell at the cluster center is given as

$$R_{\text{tot}} = \frac{1}{K} \sum_{m=1}^M \log_2(1 + \text{SINR}_m) [\text{bps}/\text{Hz}/\text{cell}]. \quad (5)$$

### III. PARTICLE SWARM OPTIMIZATION

Particle Swarm Optimization (PSO) is a stochastic optimization algorithm inspired from the movement of a flock of birds, a shoal of fish, etc [5]. The birds are modeled as particles traveling in the search space to find food or the feasible solution of a given objective function. Their social behavior is modeled as a swarm. In [5], the algorithm simulating the social behavior was simplified and was observed to be performing optimization. PSO being metaheuristic does not guarantee a global optimum, but when implemented as a stochastic optimization, randomness is injected into the algorithm to move away from the local solution when searching for a global optimum. In this paper, a basic PSO capable of finding an equilibrium solution with *inertia weight* and without the *craziness* operator [5], is presented. As per [6], a basic PSO does not satisfy the convergence condition for global search. PSO is chosen over classical optimization methods, as the overlapping clusters (dynamic changes in the aggregated channel matrix in every frequency/time resource) formed with

PJP make the linear ZF BF technique presented in [3] difficult to realize.

The ZF precoding solution proposed in [3] for single antenna systems is simplified into a classic linear algebra problem  $\mathbf{A}\mathbf{x} = \mathbf{b}$ , where  $\mathbf{A}$  is formed by block diagonalizing the aggregated channel matrix;  $\mathbf{x}$  and  $\mathbf{b}$  are the vectorized BF and identity matrices, respectively. The zeros in the  $\mathbf{x}$  eliminate the columns of  $\mathbf{A}$  and the solution reduces to a classic right inverse as in eqn (2). This can be written compactly as  $(\mathbf{I}_M \otimes \tilde{\mathbf{H}}) \cdot \text{vec}(\tilde{\mathbf{W}}_e) = \text{vec}(\mathbf{I}_M)$ , where the inactive links or zeros in  $\tilde{\mathbf{W}}_e$  have eliminated the columns of the Kronecker product. This elimination can give rise to an overdetermined system and the right inverse does not exist. An overdetermined system is encountered  $\sim 83\%$  of the time, when this approach is applied for PJP with active set threshold of 10 dB in our evaluation setup, see section IV. Hence, we propose PSO to overcome the limitations in the state of the art ZF BF solution in [3], and we later show that PSO performs better.

The BF weight matrix  $\tilde{\mathbf{W}}$  is stochastically initialized, and zeros are inserted according to the active set matrix. The non-zero BF weights are mapped to the  $i$ th particle as  $\mathbf{X}(i, j) \leftarrow \Re\{\tilde{\mathbf{W}}(m, n)\}$  and  $\mathbf{X}(i, j+1) \leftarrow \Im\{\tilde{\mathbf{W}}(m, n)\}$ . The search space of the particles is initially limited to  $[x_{\min}, x_{\max}]$ , where  $x_{\max} = 1/\max\{|\tilde{\mathbf{H}}(i, j)|\}$ . This value is chosen as the starting limit of the particles in the search space for faster convergence. PSO stochastically changes these limits in every iteration. The aim of the PSO is to optimize the position of the particles of the swarm based on the objective function. The resulting best particle is chosen as the best BF weights. In this work, two different cases are considered:

$$\text{case a)} \quad \arg \min_{\tilde{\mathbf{W}}_a} \left\{ \|\tilde{\mathbf{H}}\tilde{\mathbf{W}}_a - \mathbf{I}\|_F \right\} \quad (6)$$

$$\text{case b)} \quad \arg \min_{\tilde{\mathbf{W}}_b} \left\{ \|\text{OffDiag}(\tilde{\mathbf{H}}\tilde{\mathbf{W}}_b)\|_F \right\} \quad (7)$$

Both objective functions for the PSO are subject to  $\forall i, j : \tilde{\mathbf{H}}(i, j) = 0$  maps to  $\tilde{\mathbf{W}}_x(j, i) = 0$ , where  $x$  represents case a) or case b). In case a), the identity matrix is subtracted from the product of the sparse aggregated channel matrix and the aggregated BF matrix. The identity matrix tries to ensure that all the users are fairly served, i.e., it aims for equal receive power to all the users. This is based on the ZF philosophy that  $\mathbf{H}\mathbf{W} = \mathbf{I}$ . The interference is minimized considering fairness between users. In case b), only the off diagonal elements of the product of the sparse aggregated channel matrix and the aggregated BF matrix is considered. The non-zero off diagonal elements represent the presence of interference in the system involved in JP. The aim of this objective function is to minimize the interference alone.

To evaluate the objective function, the positions of the particles are demapped to form the complex BF weight, i.e.,  $\tilde{\mathbf{W}}(m, n) \leftarrow \{\mathbf{X}(i, j)\} + i \cdot \{\mathbf{X}(i, j+1)\}$ , where 'i' is the imaginary unit. The core of the PSO algorithm is described in eqns (8)-(11). These equations are evaluated in every iteration, i.e.,  $\forall i = 1, \dots, N; j = 1, \dots, n$ , where the time step length is  $\Delta t = 1$ ,  $N$  is the number of particles and  $n$  is the

number of variables. Eqn (8) updates the velocity of the particle, where the term involving  $c_1$  is called the *cognitive* component (weighs the self confidence of that  $i$ th particle) and the term involving  $c_2$  is called the *social* component (weighs the reliability of other particles for that  $i$ th particle).  $\mathbf{X}^{pb}(i, :) \leftarrow \mathbf{X}(i, :)$  is the best position attained by the  $i$ th particle itself carrying the best BF weights it could find,  $\mathbf{x}^{sb}$  is the best position attained by any particle carrying the best BF weights in the entire swarm.  $p$  and  $q$  are uniform random numbers in  $[0, 1]$ . Eqn (9) restricts the maximum velocity of the particle to  $v_{\max}$ , such that the particles do not diverge and eqn (10) updates the position of the particle. An *inertia weight*,  $w$ , is used to bias the current velocity based on its previous value in eqn (8), such that when the inertia weight is initially 1.4, being greater than 1, the particles are biased to explore the search space. When  $w$  decays to 0.4, due to a constant decay factor  $\beta$  in eqn (11), the cognitive/social components are given more attention [7].

$$\mathbf{V}(i, j) \leftarrow w \cdot \mathbf{V}(i, j) + c_1 \cdot p \cdot (\mathbf{X}^{pb}(i, j) - \mathbf{X}(i, j)) / \Delta t + c_2 \cdot q \cdot (\mathbf{x}^{sb}(j) - \mathbf{X}(i, j)) / \Delta t \quad (8)$$

$$|\mathbf{V}(i, j)| < v_{\max} \quad (9)$$

$$\mathbf{X}(i, j) \leftarrow \mathbf{X}(i, j) + \mathbf{V}(i, j) \cdot \Delta t \quad (10)$$

$$w \leftarrow w \cdot \beta \quad (11)$$

The *cognitive* factor,  $c_1$ , and the *social* factor,  $c_2$ , are equal to 2 as highlighted in [5], [6]. These references also indicate that the choice of the nonlinear decreasing in the inertia weight with an initial value of 1.4 ensures that the particles cover a large search space and then the particles focus on refining the solutions. The choice of the decay factor as 0.99 indicates slow decaying in the inertia weight. The number of particles is also fixed throughout the simulation. The choice of the number of particles,  $N = 30$ , was found to be a reasonable compromise with the computational complexity and the ability of the PSO to find an equilibrium solution, see [6] and the references therein. The values of the PSO parameters used in this initial work are the typical values and are summarized in Table I. The influence of the PSO parameters chosen for the design of the precoding matrix needs to be investigated further as part of our future work, and also with the channel data from field measurements. PSO is preferred over other genetic algorithms, as it has the least number of parameters to configure and promises better computational efficiency.

#### IV. SIMULATION RESULTS

Consider the scenario with  $K = M = 3$  single antenna BSs/UEs involved in JP. Each BS covers a hexagonal cell of radius,  $R = 0.5$  kms. The UEs are placed at the cluster center along an ellipse with semi-major and semi-minor axis of length  $\frac{R}{16}$  and  $\frac{h/2}{16}$ , respectively, where  $h$  is the height of the hexagon, as illustrated in Figure 1. The pathloss model in [8] is used,

$$\gamma_{\text{PL}}(\text{dB}) = 128.1 + 37.6 \log_{10} R.$$

The channel is realized as follows,  $\mathbf{H} = \Gamma \sqrt{G \cdot \gamma_{\text{PL}} \cdot \gamma_{\text{SF}}}$ , where  $\Gamma \sim \mathcal{CN}(0, 1)$  are i.i.d complex Gaussian fading coef-

Table I  
PARTICLE SWARM PARAMETERS

Parameters	Values
Number of Particles, $N$	30
$x_{\max} = -x_{\min}$	$1 / \max \{  \mathbf{H}(i, j)  \}$
Max. velocity, $v_{\max}$	$(x_{\max} - x_{\min}) / \Delta t$
Cognitive factor, $c_1$	2
Social factor, $c_2$	2
Inertia Weight, $w$	$1.4 \rightarrow 0.4$
Constant decay factor, $\beta$	0.99

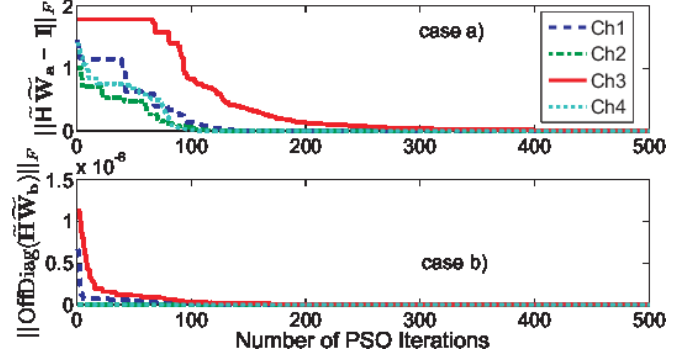


Figure 2. Convergence behavior of PSO for 4 different channel realizations

ficients,  $\gamma_{\text{SF}} \sim \mathcal{N}(0, 8 \text{ dB})$  is the shadow fading component, and  $G$  is the transmit antenna gain of 9 dBi. The system Signal to Noise Ratio (SNR) or the reference value of one UE located at the cell-edge, is fixed at 15 dB, giving rise to a maximum BS transmit power of 0.0603 W. It has been shown that it is difficult to estimate channels with pilot overhead for PJP with active set threshold greater than 15 dB at the cell-edge [9]. Hence, in this setup, a PJP threshold of 10 dB is considered.

The PSO with two different objective functions with case a) and b) can be compared with case c) based on [3]. The case c) is a typical ZF BF as in eqn (2) obtained from limited feedback where backhaul reduction is achieved with explicit nulling of the BF coefficients, i.e., zeros or nulls are placed in the BF matrix where needed. Similarly, case d) achieves backhaul reduction with explicit nulling but this is a genie aided case where full feedback is allowed i.e., complete CSI is available at the CCN. CJP has the full feedback and full backhauling, without any reduction in feedback or backhaul.

The convergence of the PSO algorithm is shown in Figure 2 for various aggregated channel matrices at CCN. The PSO in case b) converges the fastest within 100 iterations. Once the BFs are obtained from various algorithms, a simple power allocation per BS as in eqn (3) is performed, where there is at least one BS transmitting at maximum power. PSO being an iterative procedure, the per-BS power constraint can be applied in every iteration or after convergence. It should be noted that in case a), as defined in eqn (6), the power constraint is applied after the PSO algorithm has converged and in case b), as defined in eqn (7), the power constraint is applied after evaluating the objective function in every iteration. Other combinations were not considered, as the convergence was poor when the power constraint was applied in every iteration in case a). Applying the power constraint after convergence in case b) had more residual interference with slower conver-

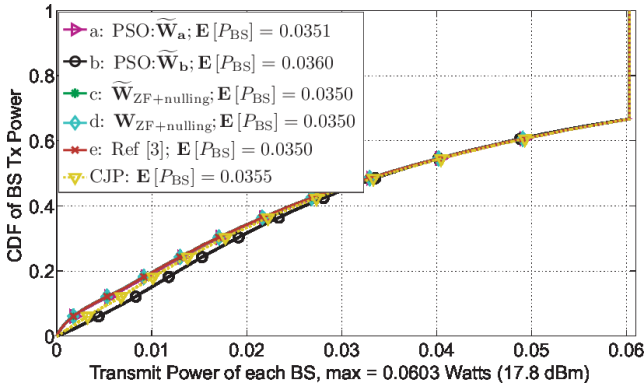


Figure 3. CDF of the BS power ( $P_{BS}$ ) transmitted with various precoding algorithms. The  $k$ th BS power is calculated as  $\|\tilde{\mathbf{W}}^k\|_F^2$ .

gence. Hence, the best combinations were considered. Figure 3 shows the Cumulative Distribution Function (CDF) of the BS transmit power. The values in the legend show the mean value of a given CDF. It can be observed that the PSO in case b) uses the BS power constraint,  $P_{\max}$  more effectively compared to others, as it is the bottom-most curve and it uses 2.85% relatively more power than case e), in average. All CDFs exhibit maximum power for 33.3% of the time, this is due to eqn (3), as at least one of the 3 BSs is transmitting at maximum power.

The residual interference power in the system was calculated based on  $\|\text{OffDiag}(\mathbf{H}\tilde{\mathbf{W}}_x)\|_F^2$ , where  $\mathbf{H}$  is the actual channel and  $x$  can be any one of the cases. It was observed that case a) trying to fairly serve all the users with equal power leaves 0.69% relatively more interference in the system when compared to case e). While, case b) reduces the interference by 3.94% relatively compared to case e). The CDF of the residual interference is not shown here, due to lack of space.

The CDF of the sum rate is shown in Figure 4. Case a) and b) perform the best compared to the other cases, with a PJP-10 dB threshold on CSI feedback. Case b) has a relative improvement in the average sum rate by 2.45% compared to case e), state of the art [3]. It should be noted that in this simulation setup, for fair comparison, only those cases are considered where the active set does not produce an overdetermined system, so that the state of the art solution in case e) works. This also means that PSO can be applied in all the scenarios without any restriction regarding scheduling or the need to have a well conditioned aggregated channel matrix. Hence, PSO not only achieves the reduction in backhaul, there is also a greater gain compared to the state of the art, as it always finds a solution. The case b) with PSO performs 2.98% relatively better than the genie aided full feedback case d) with reduced backhaul. The complexity analysis of the PSO algorithm with the constraint of backhaul load being equivalent to feedback load is treated as part of our future work.

## V. CONCLUSION

Efficient backhauling techniques are needed to realize the gains of joint processing with reduced feedback. Existing techniques of achieving efficient backhauling in this context

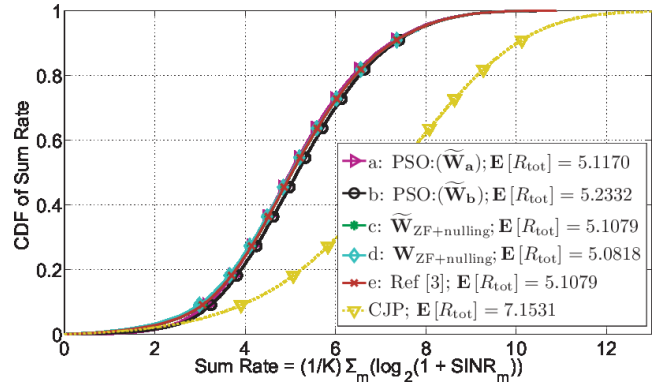


Figure 4. Sum rate per cell,  $R_{\text{tot}}$ , of each algorithm and  $\mathbb{E}[R_{\text{tot}}]$  is the mean of  $R_{\text{tot}}$ . CJP can be viewed as the upper bound for ZF.

have constraints on scheduling or need full channel state information to be fed back. The particle swarm optimization used in this paper, is able to perform without any such constraints.

When the state of the art technique can find a solution, the average sum rate of the stochastic optimization algorithm performs 2.45% better than the state of the art solution, without any restriction on scheduling, such that the backhauling load is equivalent to that of the feedback load. The proposed algorithm is also capable of performing even when the state of the art solution fails due to an overdetermined system. The best algorithm proposed in this paper converges in merely 100 iterations, with effective usage of the base station power and lowering interference. With parallel computing and with more improved flavors of swarm algorithms, this complexity is feasible. Robustness of the proposed particle swarm algorithm with optimized parameters and imperfect channel state information will be studied as part of our future work.

## REFERENCES

- [1] 3GPP TR 36.814-900, "3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Further Advancements for E-UTRA Physical Layer Aspects (Release 9)," Mar. 2010.
- [2] C. Botella, T. Svensson, X. Xu, H. Zhang, "On the performance of joint processing schemes over the cluster area," in *Proc. IEEE Vehicular Tech. Conf.*, May 2010.
- [3] A. Papadogiannis, H.J. Bang, D. Gesbert, E. Hardouin, "Efficient selective feedback design for multicell cooperative networks," *IEEE Trans. on Vehicular Tech.*, vol. 60, no. 1, pp. 196-205, Jan. 2011.
- [4] Fang Shu, Wu Gang, Li Shao-Qian, "Optimal multiuser MIMO linear precoding with LMMSE receiver," *EURASIP Journal on Wireless Comm and Networking*, vol. 2009, Article ID 197682, Mar. 2009.
- [5] J. Kennedy, R.C. Eberhart, "Particle swarm optimization," in *Proc. IEEE Int. Conf. on Neural Networks*, pp. 1942-1948, 1995.
- [6] AP Engelbrecht. *Fundamentals of computational swarm intelligence*, (John Wiley & Sons, 2005), pp. 171-172
- [7] Y.H. Shi, R.C. Eberhart, "Parameter Selection in Particle swarm optimization," *The 7th Annual Conf. on Evolutionary Programming*, San Diego, USA, 1998.
- [8] 3GPP TR 36.942-a20, "3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); Radio Frequency (RF) system scenarios (Release 10)," Jan. 2011.
- [9] ARTIST4G D1.2, "Innovative advanced signal processing algorithms for interference avoidance", *ARTIST4G technical deliverable*, 2010. [Online] Available: <https://ict-artist4g.eu/projet/deliverables>