

Research Article

Downlink Multicell Processing with Limited-Backhaul Capacity

O. Simeone,¹ O. Somekh,² H. V. Poor,² and S. Shamai (Shitz)³

¹ CWCSR, New Jersey Institute of Technology, Newark, NJ 07102, USA

² Department of Electrical Engineering, Princeton University, Princeton, NJ 08544, USA

³ Department of Electrical Engineering, Technion, Haifa, 32000, Israel

Correspondence should be addressed to O. Simeone, osvaldo.simeone@njit.edu

Received 16 November 2008; Revised 7 March 2009; Accepted 18 May 2009

Recommended by Robert W. Heath

Multicell processing in the form of joint encoding for the downlink of a cellular system is studied under the assumption that the base stations (BSs) are connected to a central processor (CP) via finite-capacity links (finite-capacity backhaul). To obtain analytical insight into the impact of finite-capacity backhaul on the downlink throughput, the investigation focuses on a simple linear cellular system (as for a highway or a long avenue) based on the Wyner model. Several transmission schemes are proposed that require varying degrees of knowledge regarding the system codebooks at the BSs. Achievable rates are derived in closed-form and compared with an upper bound. Performance is also evaluated in asymptotic regimes of interest (high backhaul capacity and extreme signal-to-noise ratio, SNR) and further corroborated by numerical results. The major finding of this work is that even in the presence of oblivious BSs (that is, BSs with no information about the codebooks) multicell processing is able to provide ideal performance with relatively small backhaul capacities, unless the application of interest requires high data rate (i.e., high SNR) and the backhaul capacity is not allowed to increase with the SNR. In these latter cases, some form of codebook information at the BSs becomes necessary.

Copyright © 2009 O. Simeone et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. Introduction

Multicell processing promises to dramatically increase the throughput of infrastructure (cellular) systems [1, 2]. The technology prescribes joint processing of different base stations' (BSs') signals for downlink or uplink, so as to mimic a multiantenna (MIMO) transmitter or receiver, respectively. It is enabled by the presence of backbone links connecting the BSs, either among themselves or with a central processor (CP).

Previous works on multicell processing have dealt with different cellular models that capture various tradeoffs between adherence to "real" systems and analytical tractability. Such scenarios range from the Wyner model [3], which accounts for the essence of cellular systems in terms of inter-cell interference and possibly fading (see [2] for an extensive literature overview), to more complex and realistic scenarios which also model other effects such as random geometric distribution of the users, multiple antennas transmitters/receivers, and so forth, (see, e.g., [4–6]). While the former

modelling provides basic analytical and theoretical insight into the network performance and optimal operation, the latter provides the necessary framework to assess the impact of relevant practical issues via analysis backed by numerical results. Reference [7] discusses the relationship between the two approaches, showing that the Wyner model predicts well results for more general and realistic frameworks.

Focusing on the downlink, which is of interest here, the current state-of-the-art on multicell processing encompasses investigations of the throughput achievable by a number of different joint transmission strategies, including Dirty Paper Coding (DPC) [1, 8, 9], joint beamforming (linear precoding) [4, 10, 11], and joint scheduling [5, 6]. As is well summarized in the conclusions of [1], a number of issues remain to be addressed before realistically considering multicell processing for future wireless systems, namely (in the order of [1]): the need for a high-speed backbone enabling information (data, control/synchronization, and channel state) exchange between the base stations, the

requirement of channel information availability for coherent methods, and timing/phase synchronization.

1.1. Main Contributions and Related Work. In this paper, we are concerned with investigating the first issue listed in [1], that is, the role of the backbone capacity in enabling multicell processing. More specifically, we are interested in assessing how the backhaul capacity influences the achievable rate or energy efficiency gains for the downlink, and in identifying corresponding effective processing techniques. It is noted that such an analysis is particularly critical in systems such as WiMax, where multicell cooperation is envisaged (see, e.g., [12]) but, unlike existing cellular systems, an infrastructure of high-capacity backhaul links is not readily available, thus raising the issue of designing the backhaul. Our focus is on obtaining analytical insight and, therefore, we concentrate on a simple Wyner-type model, following, for example, [9, 13, 14], and thus do not account for the other practical issues (Channel State Information (CSI) [15], synchronization) mentioned above.

Related work has also attempted to study the impact of limitations in the backhaul on the throughput of multicell processing. Specifically, while [8, 16, 17] limit the connectivity between the BSs and the CP (only a subset of BSs is connected to the same CP), references [17–19] (uplink) and [20, 21] (downlink) enforce a topological constraint in that only backhaul links between adjacent BSs are available (no CP) and message passing techniques are implemented. Finally, reference [22] focuses on the uplink and assumes that the links between all the BSs and a CP have finite-capacity (*finite-capacity backhaul*).

Here, we extend the analysis of [22] to the downlink with finite-capacity backhaul. Our theoretical contribution can be also seen in the more general context of decentralized processing (see [23] and references therein). We consider different transmission strategies which require different amounts of information at the BSs regarding the codebook used to communicate with the Mobile Stations (MSs). In particular, similarly to [22], we first consider *oblivious BSs* that do not possess any *Codebook Information (CI)*. (It should be remarked that, when employing techniques such as DPC, encoding is performed with a more sophisticated encoding strategy than simple look-up on a table of codewords on the basis of the transmitted message. The transmitted signal is in fact a function of the interference sequence to be cancelled. Therefore, a more appropriate term for what we refer to as CI would be encoding function information. We choose the first for simplicity but this distinction should be kept in mind.) This scenario is of specific interest for nomadic applications where information about roaming users is not available at the BSs. Having identified the limits of this technique in specific regimes of interest, we then investigate other solutions based on some degree of CI at the BSs (either *local CI* or *cluster CI*). Our results shed light, via analytical insights, into the relative performance of different transmission strategies in the presence of limitations on the backhaul links. In particular, by comparing the derived achievable rates to an upper bound, optimality of some of the proposed techniques

is established in specific regimes. It is finally noted that the results in this paper were partially presented in [24], and that in [25] some of the techniques proposed here are applied to a single source-destination distributed MIMO setting.

2. System Model

We study the downlink of a cellular system modelled as in Figure 1, where M cells, with single-antennas BSs, are arranged in a linear geometry, as would be the case for a system deployed along a highway or a long avenue [26]. At any given time, one single-antenna terminal is served in each cell via, for example, intracell TDMA. The served users are located close to the border between successive cells in a “soft-handoff” condition. In this case, each active terminal, say the m th, receives signals from the local m th BS and the previous, $(m - 1)$ th, BS. This framework is a variation of the Wyner model [3] and has been studied in [9] and later [13, 14] in terms of sum-rate for the case where there are no restrictions on the backbone connecting the BSs. We refer to [19] (see Appendix A therein) where some further discussion on the validity of this model is presented. Deviating from the ideal condition of unlimited backhaul, here we assume that each BS is connected to a central processor via a finite-capacity link of capacity C (bits/channel use) as in [22]. (It is noted that the condition of finite-capacity backhaul limits the number of bits that can be conveyed over each backhaul link per transmission block (codeword). The normalization of this overall amount of bits over the number of downlink channel symbols considered here is rather arbitrary but convenient for the interpretation of the results.) The model is further characterized by a single parameter to account for intercell interference, namely, the power gain $0 \leq \alpha \leq 1$ (in [9] it was $\alpha = 1$). Accordingly, the signal received at the m th MS is given by

$$Y_m = X_m + \alpha X_{m-1} + Z_m, \quad (1)$$

where X_m is the symbol transmitted at a given discrete time by the m th BS with per-symbol power constraint $E[|X_m|^2] = P$ and the noise Z_m is a white proper complex Gaussian process with unit power. We remark that we will be interested in asymptotic results where the number M of cells is large, which, as discussed in [2], provides a reliable measure of the performance also with finite and small M . Moreover, we focus on Gaussian (nonfaded) channels for simplicity. It is noted that such an assumption clearly bypasses the critical issues of acquiring CSI at the BSs or at the CP in a more realistic fading channel. Finally, we assume that each MS has available CI of the local transmission only, thus ruling out sophisticated joint decoding techniques at the MSs (see, e.g., [27]).

Messages $\{W_m\}_{m=1}^M$ to be delivered to the M MSs are generated randomly and uniformly in the set $\{1, 2, \dots, 2^{nR}\}$ at the CP (see Figure 1), where R (bits/channel use) is the common rate of all the messages (*per-cell rate*). Using standard definitions, we will say that a per-cell rate R is achievable if there exists a sequence of codes (i.e., encoders and decoders) with codewords of length n such that the

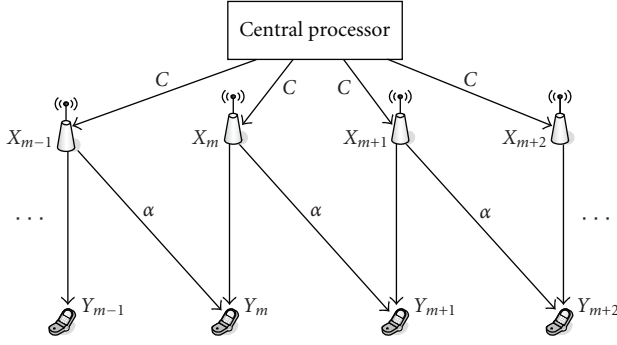


FIGURE 1: Linear cellular model of interest characterized by finite-capacity links between a central unit processor (that generates the messages to be delivered to each user) and the base stations.

probability of having at least one decoding error in the system vanishes as $n \rightarrow \infty$, that is, $\Pr[\cup_m \{\widehat{W}_m \neq W_m\}] \rightarrow 0$, where \widehat{W}_m is the estimated message at the m th MS.

3. Reference Results

In this section, we review an upper bound on the per-cell rate that can be easily derived from a result presented in [9] for $\alpha = 1$, and later extended by [13] to any $\alpha \leq 1$. (Notice that this result was not given in this form in [13] but can be easily derived from Lemma 3.5 therein.)

Proposition 1 (Upper bound). *The per-cell capacity of the system is upper bounded by*

$$R_{ub} = \min \left\{ C, \log_2 \left(\frac{1 + (1 + \alpha^2)P + \sqrt{1 + 2(1 + \alpha^2)P + (1 - \alpha^2)^2 P^2}}{2} \right) \right\}. \quad (2)$$

Proof. This result follows by considering a cut-set bound for two cuts, one dividing the central processor from the BSs and one the BSs from the MSs. For the second cut, it is noted that the system is equivalent to the infinite-capacity backbone case for which the per-cell capacity has been derived in [8, 13]. \square

It is relevant to notice that upper bound (2) remains valid even if we allow multiple MSs to be simultaneously active in each cell (and P is the per-cell power constraint), as follows easily from [3] and duality arguments [9]. Therefore, *whenever achievable rates will be shown in the following to attain (2) in specific regimes, optimality should be intended not only under the restriction of intracell TDMA strategies but also for the general case where more MSs can be scheduled at the same time* (with a total per-cell power constraint).

For future reference, two further observations on the upper bound (2) are in order. First, it is interesting to

study the *low-SNR behavior*, in the sense of [28], which is significant for systems with sufficiently large bandwidth available. Accordingly, the minimum energy per bit for reliable communication $E_b/N_{0,\min}$, and the corresponding slope of the spectral efficiency [28] are easily shown to be given by

$$\frac{E_b}{N_{0,\min,\text{ub}}} = \frac{\log_e 2}{1 + \alpha^2}, \quad S_{0,\text{ub}} = \frac{2(1 + \alpha^2)^2}{1 + 4\alpha^2 + \alpha^4}. \quad (3)$$

This result shows that the power gain with respect to a single-link (interference-free) Gaussian channel (for which $E_b/N_{0,\min} = \log_e 2$) due to multicell processing can be quantified in the low-SNR regime by the factor $(1 + \alpha^2) \geq 1$ (and the slope $S_{0,\text{ub}}$ is a decreasing function of α^2). A second observation concerns the following question regarding the *high-SNR behavior*: how fast need the backhaul capacity C grow with increasing P in order to guarantee the optimal multiplexing gain of a system with unlimited backhaul capacity? Recalling that the maximum multiplexing gain of a multiantenna broadcast channel with channel state information at the transmitter equals the number of transmit antennas [15], it easily follows that the optimal multiplexing gain of the per-cell rate (2) is 1 and that, in order to achieve it, the capacity C needs to grow as $C \sim \log_2 P$. In the following, this requirement in terms of capacity C will be compared with that of practical transmission schemes.

4. Oblivious BSs (No CI)

We start by considering rate achievable with oblivious BSs (no CI), which was investigated in [22] for the uplink of the channel at hand. Specifically, with oblivious BSs, the BSs are not aware of *any* codebook in the system so that encoding can take place only at the CP. In such conditions, we propose the following scheme. The CP performs joint DPC, which would be optimal for the case $C \rightarrow \infty$ (as for any MIMO broadcast system, see, e.g., [9]), producing the sequences of n symbols $\{\tilde{\mathbf{X}}_m\}_{m=1}^M$, one per BS. It is noted that, in a fading channel, such a scheme would require full CSI at the CP. To convey such sequences to the BSs over the finite-capacity links, each $\tilde{\mathbf{X}}_m$ is quantized using a Gaussian quantization codebook with 2^{nC} codewords, producing the compression codeword of n symbol $\hat{\mathbf{X}}_m$. The index of such a codeword is sent to the corresponding BS and simply forwarded by the latter towards the MSs (i.e., $\mathbf{X}_m = \hat{\mathbf{X}}_m$). In designing DPC at the CP, one should pay attention to the fact that the BSs forward inevitably also quantization noise, so that

- (i) in order to meet the power constraint $E[|\mathbf{X}_m|^2] = P$ one should ensure that (see the proof for details)

$$E\left[|\tilde{\mathbf{X}}_m|^2\right] = \frac{P}{1 + 1/(2^C - 1)}; \quad (4)$$

- (ii) the equivalent SNR at the MSs is decreased from P to

$$\tilde{P} = \frac{P}{(1 + (1 + \alpha^2)P)/(2^C - 1) + 1}. \quad (5)$$

From these considerations, the following result can be proved.

Proposition 2 (Oblivious BSs). *Assuming that the BSs are oblivious (no CI), the following rate is achievable with central encoding:*

$$R_{obl} = \log_2 \left(\frac{1 + (1 + \alpha^2)\tilde{P} + \sqrt{1 + 2(1 + \alpha^2)\tilde{P} + (1 - \alpha^2)^2\tilde{P}^2}}{2} \right). \quad (6)$$

Proof. See Appendix A. \square

Remark 1. It is emphasized that the considered scheme does not exploit in any way the specific structure of the considered channel, but the latter enables a simple closed-form expression to be obtained for the achievable rate (6). This expression permits further insight to be gained into the performance of downlink transmission schemes based on oblivious BSs and *compress-and-forward*-type communication schemes, which are expected to be qualitatively valid also for more complicated models (see, e.g., the discussion in [7]).

4.1. Performance in Asymptotic Regimes. It is clear that the rate R_{obl} (6) does not achieve the upper bound (2) for all values of the system parameters. However, we will show that the proposed technique is optimal or nearly optimal under a number of conditions.

High backhaul capacity. At first, we notice that, in the absence of constraints on the backhaul, that is, $C \rightarrow \infty$ the scheme proposed above is optimal, $R_{obl} \rightarrow R_{ub}$, since $\tilde{P} \rightarrow P$. This conclusion is not surprising since for $C \rightarrow \infty$ the system is free from the impairment due to compression noise and thus DPC achieves capacity.

High SNR. More interesting is the “dual” regime $P \rightarrow \infty$ where compression noise on the backhaul links plays a major role. In this case, we have

$$\lim_{P \rightarrow \infty} R_{obl} = C - 1 + \log_2 \left(1 + \sqrt{1 - \frac{4\alpha^2}{(1 + \alpha^2)^2} (1 - 2^{-C})^2} \right), \quad (7)$$

thus falling short of achieving the upper bound $R_{ub} = C$ (for $P \rightarrow \infty$) by at most one bit (unless $\alpha = 0$, in which case we clearly have optimality). Another measure of interest in the high SNR regime is obtained by letting P grow and allowing the backhaul capacity C to scale with P , in order to assess under which condition can the optimal *multiplexing gain* be retained (see discussion in the previous section). Substituting $C = r \log_2 P$ in (6), it can be seen that the multiplexing gain with this choice is given by $\min(r, 1)$, so that the optimal multiplexing gain of 1 can be achieved by having $C \sim \log_2 P$, which is optimal according to our discussion on the upper bound.

Low SNR. Finally, the low-SNR (wideband) characterization is given by

$$\frac{E_b}{N_{0 \min}} = \frac{E_b}{N_{0 \min, ub}} \cdot \frac{1}{(1 - 2^{-C})} \quad (8)$$

$$S_0 = S_{0, ub} \cdot \frac{1}{1 + S_{0, ub}(2^{-C})/(1 - 2^{-C})}.$$

This result shows that *the energy efficiency loss due to finite-capacity backhaul can be quantified in the low-SNR regime by $(1 - 2^{-C})$* . This loss, accordingly to the discussion above, tends to zero for $C \rightarrow \infty$. It is remarked that, interestingly, the low-SNR performance (8) of the scheme at hand coincides with the uplink transmission strategy of [22], suggesting a limited duality between uplink and downlink with finite-capacity backhaul.

5. Cluster CI

In the previous section, it was shown that, while oblivious BSs are able to achieve capacity if the backhaul capacity C is large enough, in other regimes of interest (such as with large power P) a performance loss is incurred. In this section, we thus propose two techniques that can overcome these limitations by exploiting CI at the BS (i.e., nonoblivious BSs). Specifically, each BS is assumed to know its own encoding function and the encoding functions of a number of other nearby BSs (*cluster CI*). Moreover, unlike the previous section, encoding is carried out at each BS and no encoding is performed at the CP. We define the two techniques as sequential DPC and joint DPC.

5.1. Sequential DPC. Sequential DPC exploits the locality of the interference (recall Figure 1) and is inspired by the approach in [29], where a similar approach is used in a cognitive radio context. It is noted that the results of [29] were limited to the multiplexing gain and cannot be directly applied here given the different setting. It is also worth pointing out that the deployment of this scheme for more general channel model would require some approximation, for example, treating some of the interference contributions from other cells as noise, thus achieving only partial interference cancellation.

To elaborate, every m th BS knows its encoding function and the encoding functions of the J BSs preceding it (i.e., BSs $m - i$ with $i = 1, \dots, J$). At the beginning of the transmission block, each BS receives from the CP $J + 1$ messages $\{W_{m-i}\}_{i=0}^J$, that is, the local message and the messages of the J preceding BSs. The basic idea is now that, based on these J additional messages and the knowledge of the corresponding encoding functions, the m th BS can perform DPC over these messages and cancel the intercell interference achieving the single-user (interference-free) rate $\log_2(1 + P)$. It is noted that, in a fading channel, each BS requires the CSI corresponding to the J preceding BSs. As pointed out in [29], in order to implement the sequential DPC scheme correctly, we need to “turn off” every $(J + 2)$ th BS (e.g., BSs $J + 2, 2(J + 2), \dots$) and consider the clusters of

$J + 1$ BSs in between silent BSs. To elaborate, let us focus on any cluster and index the participating BSs as $m' = 0, 1, \dots, J + 1$, where $m' = 0$ corresponds to a silent BS, $m' = 1$ is the leftmost BS in a cluster and so on. Due to the intercell interference structure in the model at hand (recall Figure 1), the BS with $m' = 1$ in each cluster can achieve single-user rate without performing any DPC. The second BS ($m' = 2$) instead needs to perform DPC on the signal transmitted by $m' = 1$, which can be done, since the second BS can reconstruct this signal knowing the encoding function and the corresponding message of the BS with $m' = 1$. Proceeding this way, it is clear that clustering is necessary because the signal transmitted by the m' th BS in each cluster depends in fact on the signals transmitted by all the preceding $m' - 1$ BSs within the cluster due to the successive DPC encodings.

Proposition 3 (Sequential DPC). *Assuming that every m th BS knows its own encoding function and the encoding function of the J BSs preceding it (cluster CI), the following rate is achievable with sequential DPC:*

$$R_{\text{seq}} = \min \left\{ \frac{2C}{J+2}, \left(1 - \frac{1}{J+2}\right) \log_2(1+P) \right\}. \quad (9)$$

Proof. We consider equal-time time-sharing among $J + 2$ cluster configurations so that in the j th configuration ($j = 1, \dots, J+2$), we silence cells $(J+2)+j-1, 2(J+2)+j-1, \dots$. This way, each BS occupies all the $J+2$ positions $m' = 0, 1, \dots, J+1$ in a cluster, one for each configuration. Rate splitting is then performed so that a given message W_m is split into $J + 1$ messages with equal rate R' ($R = R'(J + 1)$) to be transmitted during the $J + 1$ configurations where the m th BS is not silent. It is easy to see that, since each BS occupies all the $J + 2$ positions in a cluster and that m' messages need to be delivered by the central processor when the BS occupies position m' (see discussion above), the backhaul links to all the BSs are equally utilized and the constraint on the backhaul capacity becomes $C \geq R' \sum_{m=0}^{J+1} m = (R'/2)(J + 1)(J + 2) = (R/2)(J + 2)$. Moreover, from the fact that each BS is active in $J + 1$ out of the overall $J + 2$ configurations, we have the following further constraint on the rate: $R \leq (J + 1)/(J + 2) \log_2(1 + P)$. From the two constraints above, rate (9) easily follows. \square

Remark 2. An alternative scheme could be devised that exploits the transmission power of the “silent” cell ($m' = 0$) in each cluster. This could be done by sending the message of the first BS ($m' = 1$) to the “silent” BS ($m' = 0$) on the corresponding finite-capacity link in order to allow the latter to cooperate via coherent power combining with the first BS. Following the same steps as in the proof of Proposition 3, the rate achievable by this scheme is easily derived to be

$$\min \left\{ \frac{2C(J+1)}{(J^2+3J+4)}, \frac{J}{(J+2)} \log_2(1+P) + \frac{1}{(J+2)} \log_2(1+(1+\alpha)^2P) \right\} \quad (10)$$

Performance comparison of this rate with (9) depends on the operating regime of interest, and will not be further

considered here since it would not alter meaningfully the main conclusions.

5.1.1. Performance in Asymptotic Regimes. High Backhaul Capacity. In the limit of a large backhaul capacity $C \rightarrow \infty$, for fixed cluster size $J + 1$, sequential DPC achieves rate $R_{\text{seq}} \rightarrow (1 - 1/(J + 2)) \log_2(1 + P)$ and is therefore limited by the loss in multiplexing gain (see also below) that follows from the need to silence a fraction $1/(J + 2)$ of the BSs [29].

High SNR. Consider now the regime of large power $P \rightarrow \infty$. In this case, the performance is limited by the backhaul capacity and we have $R_{\text{seq}} \rightarrow 2C/(J + 2)$, which, if we allow optimization of the cluster size, becomes $R_{\text{seq}} \rightarrow R_{\text{ub}} = C$ (for $J = 0$, that is, each cluster consists of only one active cell. This corresponds to the InterCell-Time-Sharing (ICTS) strategy [2], see also discussion in the next section.) Letting C increase with power P , for any finite J , the maximal multiplexing gain is $1 - 1/(J + 2) < 1$, and, from (9), achieving this rate scaling requires the backhaul capacity C to grow as $C \sim (J + 1)/2 \cdot \log_2 P$. Thus, sequential DPC, unlike oblivious BSs, entails a loss in terms of multiplexing gain, that can be made arbitrarily small by increasing the cluster size J but only at the expense of a proportionally faster increase of the backhaul capacity C .

Low-SNR. The low-SNR characterization for R_{seq} is given by

$$\frac{E_b}{N_{0 \min}} = \frac{\log_e 2}{1 - (1/(2 + J))}, S_0 = 2 \left(1 - \frac{1}{2 + J}\right), \quad (11)$$

which shows that sequential DPC falls short of achieving the performance of the upper bound, being designed only to cancel intercell interference. Moreover, by selecting a sufficiently large J it is clear that the single-user performance $E_b/N_{0 \min} = \log_e 2$, and $S_0 = 2$, can be achieved.

5.2. Joint DPC. A second scheme based on cluster CI can be devised that, unlike the sequential scheme described above, is able to achieve the upper bound R_{ub} in the regime of unlimited backhaul capacity ($C \rightarrow \infty$), as with oblivious BSs. The idea is to cluster the BSs as in the previously discussed scheme by silencing one every $(J + 2)$ th BS, then send all the messages to be delivered within the cluster to all the participating BSs, and finally perform joint DPC for the messages in the cluster at each BS. Notice that this scheme requires that every BS within a cluster needs to be informed about the encoding functions of all the $J + 1$ BSs within the same cluster (instead of the preceding BSs). This implies, once time-sharing is taken into account as explained above, that knowledge of $2J + 1$ encoding function is required at each node (instead of $J + 1$ as in the case of sequential DPC). As can be easily inferred from the results in [9], the rate achievable by this scheme is

$$R_{\text{joint}} = \min \left\{ \frac{C}{J+1}, \frac{1}{J+2} \min_{\text{tr}(\mathbf{Y}) \leq 1/P} \max_{\text{tr}(\mathbf{P}) \leq 1} \log_2 \frac{|\mathbf{Y} + \mathbf{H}\mathbf{P}\mathbf{H}^H|}{|\mathbf{Y}|} \right\}, \quad (12)$$

with the $(J + 1) \times (J + 1)$ channel matrix defined as a Toeplitz matrix defined by the first column $[1 \ \alpha \ \mathbf{0}^T]^T$, and

$\mathbf{P} = \text{diag}([P_1 \cdots P_{J+1}]) \succcurlyeq \mathbf{0}$ and $\mathbf{Y} = \text{diag}([\gamma_1 \cdots \gamma_{J+1}]) \succ \mathbf{0}$ being diagonal matrices collecting signal and noise powers. As it can be concluded from (12) and the results in [9], for $C \rightarrow \infty, J \rightarrow \infty$ and $C/J \rightarrow \infty$, we have $R_{\text{joint}} \rightarrow R_{\text{ub}}$. However, it will be shown in Section 7, that for relatively small values of C , rate (12) is generally smaller than (9). Finally, it is easy to see that this scheme has the same limitations in terms of multiplexing gains as sequential DPC and that its requirement in terms of scaling of capacity C is more demanding ($C \sim (J+1)^2/(J+2) \cdot \log_2 P$).

6. Local CI

Finally, here we consider a technique that also aims at alleviating the limitations of oblivious BSs but requires each BS to know *only* the local codebook (*local CI*). This does not entail any further control signalling among BSs to exchange codebook information. Moreover, the scheme proposed in this section avoids the large computational complexity associated with the cluster CI-based scheme discussed in the previous section, where multiple DPC encodings were to be carried out at each BS. Toward this goal, here the burden of encoding is shared by the BSs and the CP.

It should be mentioned right away that rate

$$R_{\text{ICTS}} = \min\left\{C, 1/2 \log_2(1+P)\right\} \quad (13)$$

can be straightforwardly achieved under the assumption of local CI by turning off one of every two BSs and using single-user codes for the active BSs (which now see interference-free channels). Notice that this corresponds to the scheme presented in the previous section with $J = 0$, and that it follows the so called Inter Cell-Time-Sharing (ICTS) approach (see [2]).

In order to improve on R_{ICTS} , we consider the following transmission scheme, which is based on the interference structure of the network and could be approximately implemented in other scenarios by treating the remaining interference terms as noise. As far as the first BS is concerned, the CP simply sends message W_1 and the BS uses a regular Gaussian codebook transmitting the sequences of n symbols \mathbf{X}_1 . The CP then quantizes \mathbf{X}_1 using a proper Gaussian quantization codebook with 2^{nR_q} codewords, producing the sequence of n symbols $\hat{\mathbf{X}}_1$. This is delivered, along with the local message W_2 , on the limited-capacity link toward the second BS. The latter transmits its message W_2 by performing DPC over the quantized signal $\hat{\mathbf{X}}_1$. The procedure is repeated in the same way for the successive BSs (notice that the CP must reproduce the transmitted signal \mathbf{X}_m , which is possible given that the CP knows the messages, encoding functions and quantization codebooks). In order to satisfy the capacity constraint on the backhaul links, the quantization rate must satisfy $R_q + R \leq C$. Finally, we remark that, in a fading channel, the CP would require full CSI, while each BS would require only local CSI regarding the useful and interfering channels.

Proposition 4 (Local CI). *Assuming that every m th BS knows only its own encoding function (local CI), the following rate is achievable:*

$$R_{\text{local}} = \begin{cases} C & \text{if } C \leq \log_2\left(1 + \frac{P}{1 + \alpha^2 P}\right) \\ R'_{\text{local}} & \text{otherwise,} \end{cases} \quad (14)$$

where

$$R'_{\text{local}} = \log_2\left(1 - \frac{2^C}{\alpha^2 P} + \sqrt{1 + \frac{2^{C+1}}{\alpha^2} \left(2 + \frac{1}{P}\right) + \frac{2^{2C}}{\alpha^4 P^2}}\right) - 1 \quad (15)$$

for $\alpha > 0$ and $\log_2(1+P)$ for $\alpha = 0$.

Proof. See Appendix B. \square

It is noted that the condition $C \leq \log_2(1 + P/(1 + \alpha^2 P))$ in (14) corresponds to the case where a rate C , which upper bounds the performance as per (2), can be achieved by simple single-user encoding and decoding in each cell, whereby intercell signals are treated as interference. Also notice that it can be easily proved that the rate R'_{local} (15) is a continuous function of α for $\alpha \geq 0$.

6.1. Performance in Asymptotic Regimes. High Backhaul Capacity. For $C \rightarrow \infty$, we have $R_{\text{local}} \rightarrow \log_2(1+P) < R_{\text{ub}}$ (as for R_{seq} and R_{joint}), which corresponds to perfect interference pre-cancellation via DPC.

High SNR. For $P \rightarrow \infty$, we have

$$\lim_{P \rightarrow \infty} R_{\text{local}} = \min\left(C, \log_2\left(1 + \sqrt{1 + \frac{2^{C+2}}{\alpha^2}}\right) - 1\right), \quad (16)$$

which is a nonincreasing function of α and reduces to C when $\alpha = 0$. It is noted that the second term of (16) is dominant for $\alpha^2 \geq 1/(2^C - 1)$, in which case R_{local} , unlike R_{seq} , is asymptotically (with P) smaller than the upper bound C . In particular, with $\alpha^2 = 1$ and increasing C , the rate $R_{\text{local}} \rightarrow C/2$ for $P \rightarrow \infty$. We now turn to the analysis of the *multiplexing gain*: setting $C = r \log_2 P$ in (14), the multiplexing gain is found to be $\min(r/2, 1)$ so that the optimal multiplexing gain of 1 can be achieved by having $C \sim 2 \log_2 P$. This contrasts with the case of local BS processing studied in the previous section where the optimal multiplexing gain was not achievable.

Low SNR. Finally, the low-SNR characterization is given by

$$\frac{E_b}{N_{0 \min}} = \log_e 2, \quad S_0 = \frac{2}{1 + 2\alpha^2 2^{-C}}, \quad (17)$$

where we see that single-user performance in terms of $E_b/N_{0 \min}$ is achieved, similarly to the case treated in the previous section, whereas the same can be said for the slope only as $C \rightarrow \infty$ (see also the discussion above).

7. Numerical Results and Discussion

In this section, we further investigate the performance of the proposed techniques in the regime of finite-capacity C and power P via numerical results. It is remarked that, while the considered techniques pose different computational requirements on the BSs and the central unit, as discussed throughout the paper, the amount of information exchanged between the central unit and the BSs is the same and limited by the backhaul capacity C .

Dependence on the Backhaul Capacity C . Figure 2 shows the rates achievable by oblivious BSs R_{obl} , sequential and joint DPC R_{seq} and R_{joint} (with optimized J), ICTS R_{ICTS} , and local CI R_{local} versus the backhaul capacity C for $P = 10$ dB and $\alpha = 1$. (What we report is actually an upper bound on R_{joint} obtained by setting $\mathbf{Y} = 1/(P(J+1))\mathbf{I}$ in (12): $R_{joint} \leq \min\{C/(J+1), 1/(J+1) \max_{\text{tr}(\mathbf{P}) \leq 1} \log |\mathbf{I} + P(J+1)\mathbf{H}\mathbf{P}\mathbf{H}^H|\}$ which can be easily solved by numerical tools for convex optimization. This choice has no consequences in our discussion since it is enough to give evidence to the negative conclusion about the performance of R_{joint} discussed in the text.) The optimal cluster-size J for R_{seq} and R_{joint} is, as expected from the discussion in Section 5.1.1, increasing with the capacity C (not shown). It is also seen that if C is large enough, and for relatively small to moderate values of P (see next figure), the proposed scheme with oblivious BSs is to be preferred. Moreover, if central processing is not feasible for limitations at the CP, it is seen that for sufficiently small values of C ($C < 30$), sequential DPC is generally advantageous over joint DPC, even though the latter is asymptotically ($C \rightarrow \infty$) optimal. For this reason in the following we will not consider R_{joint} . Also notice that while the schemes based on local CI and oblivious BSs attain the respective asymptotic values for $C \approx 10$, convergence is much slower for schemes based on local BS processing (that is, sequential and joint DPC).

Dependence on the Power P . Figure 3 shows the same achievable rate discussed above versus the power P for $C = 6$ and $\alpha = 1$. Here, the optimal cluster-size J for R_{seq} is, as expected from the discussion in Section 5.1.1, decreasing with the power P . For small-to-moderate power P , as discussed in the previous example, the preferred scheme is that based on oblivious BSs for its capability of performing joint DPC via central processing. However, as the power increases, we know from the asymptotic analysis that CI, either local (as in ICTS) or cluster (as in sequential DPC), plays a critical role. This is confirmed by Figure 3, where it is clearly shown that R_{seq} and R_{ICTS} become advantageous over R_{obl} for $P > 30$ dB.

Dependence on the Intercell Power Gain α . The impact of the intercell power gain α^2 is shown in Figure 4. Sequential DPC is designed to cancel the intercell interference and thus its performance does not depend on α . (See also discussion in Section 5.1.1 on the low-SNR regime.) Moreover, while the local CI-based scheme suffers from increasing α^2 due to the

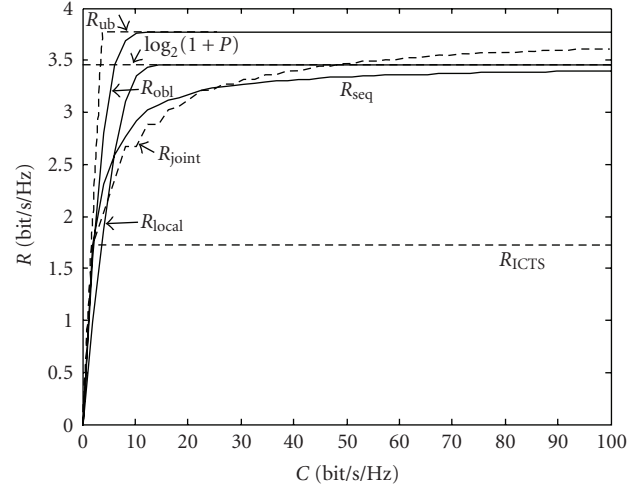


FIGURE 2: Rates achievable with oblivious BSs R_{obl} , sequential and joint DPC R_{seq} and R_{joint} (with optimized J), ICTS R_{ICTS} , and local CI R_{local} versus the backhaul capacity versus C for $P = 10$ dB and $\alpha = 1$.

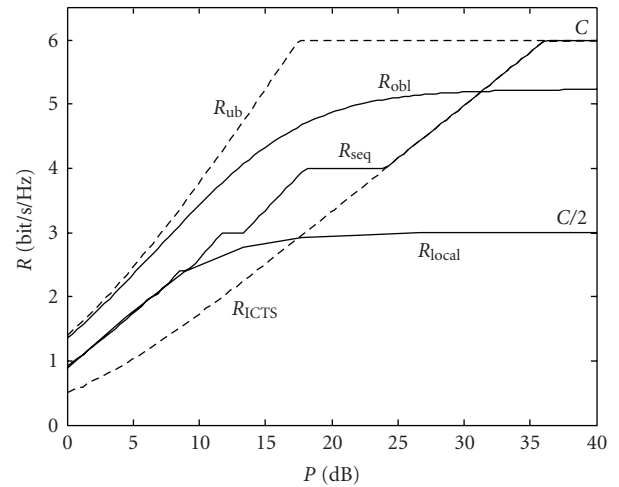


FIGURE 3: Rates achievable with oblivious BSs R_{obl} , sequential DPC R_{seq} (with optimized J), ICTS R_{ICTS} , and local CI R_{local} versus P for $C = 6$ and $\alpha = 1$.

enhanced noise level caused by quantization of the adjacent-cell transmission signal, oblivious BSs, similarly to the upper bound R_{ub} , are able to exploit the extrasignal path due to a larger α^2 .

7.1. Discussion. Our analysis and the numerical results above have shown that

- (i) the upper bound R_{ub} can be easily achieved in the regime of sufficiently large backhaul capacity even with oblivious BSs;
- (ii) in the regime of large power with fixed capacity C , achieving the upper bound is only possible if some form of CI is available at the BSs, as, for instance, by ICTS;

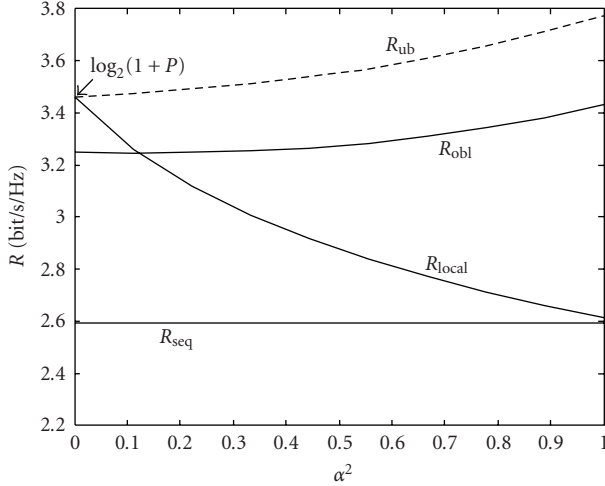


FIGURE 4: Rates achievable with oblivious BSs R_{obl} , sequential DPC R_{seq} (with optimized J), ICTS R_{ICTS} , and local CI R_{local} versus α^2 for $C = 6$ and $P = 10$ dB.

- (iii) allowing the capacity C to increase with power P , the optimal multiplexing gain of 1 with minimum scaling $C \sim \log_2 P$ can be achieved with oblivious BSs;
- (iv) for finite C and P , low-SNR analysis and numerical results have shown that all the considered schemes fall short of achieving the upper bound.

8. Conclusions

While multicell processing is by now regarded as a key candidate technology for future wireless communication standards, a number of issues remain to be investigated to fully assess its potentiality, most notably the impact of finite-capacity backhaul, imperfect synchronization, and availability of accurate CSI. In this paper, we have taken a first step towards addressing one of these issues by focusing on the impact of a finite-capacity backhaul on the downlink of a simple cellular system abstracted according to a Wyner-type model. A number of transmission techniques have been proposed that present different tradeoffs regarding the amount of codebook information (CI) and processing burden required at each BS. The main conclusion of this work is that multicell processing presents a graceful degradation in the presence of decreasing backhaul capacity, even with BSs oblivious to all the codebooks used in the system (i.e., no CI). It has also been shown that for high SNR and fixed backhaul capacity, a system with oblivious BSs is limited by the quantization noise, and knowledge of the codebooks at the BSs is necessary to avoid performance bottlenecks.

The conclusions of this work have been obtained via analytical insights enabled by the considered framework. Important issues left for future research pertain to both theoretical and more practical aspects. In the first category, of primary importance is devising strategies for the setting at hand that are optimal for all values of the system parameters. Another interesting open question is whether it is possible to assess possible duality results between uplink and downlink

channels with limited-capacity backhaul. (A low-SNR result in this sense has been identified in this paper.) In the latter, implementation of the proposed techniques in more realistic and general channel models, and possibly in the presence of synchronization errors and imperfect CSI, is of interest.

Appendices

A. Proof of Proposition 2

Quantization is performed at the CP using the forward test channel $\hat{X}_m = \tilde{X}_m + Z_{q,m}$, where \tilde{X}_m and $Z_{q,m}$ are independent complex Gaussian random variables with zero means and variances $P/(1 + 1/(2^C - 1))$ (due to the power constraint (4)) and σ_q^2 , respectively. It is noted that $Z_{q,m}$ models the quantization error. In order to send the quantized sequence represented by \hat{X}_m to the m th BS, the following condition must be satisfied from standard rate-distortion theory results:

$$C \geq I(\hat{X}_m; \tilde{X}_m) = \log_2 \left(1 + \frac{P}{\sigma_q^2(1 + 1/(2^C - 1))} \right), \quad (\text{A.1})$$

so that, taking (A.1) with equality, we have $\sigma_q^2 = P/2^C$. The signal transmitted by each BS is the quantized sequence $X_m = \hat{X}_m$, which satisfies the power constraint $E[|X_m|^2] = P/(1 + 1/(2^C - 1)) + P/2^C = P$ by construction. The signal received at each MS is then given by $Y_m = \tilde{X}_m + \alpha \tilde{X}_{m-1} + \tilde{Z}_m$, with $\tilde{Z}_m = Z_m + Z_{q,m} + \alpha Z_{q,m-1} \sim \mathcal{CN}(0, 1 + (1 + \alpha^2)\sigma_q^2)$, independent of \tilde{X}_m and \tilde{X}_{m-1} . From the previous equation we see that the system can be seen as a modified Wyner model in the sense of (1) with enhanced noise due to quantization. The corresponding SNR is

$$\tilde{P} = \frac{E[|\tilde{X}_m|^2]}{1 + (1 + \alpha^2)\sigma_q^2} = \frac{P}{(1 + (1 + \alpha^2)P)/(2^C - 1) + 1}. \quad (\text{A.2})$$

It is noted that the noise correlation between the noise samples \tilde{Z}_m does not affect the achievable rates, which depend only on the marginal distributions of the received signals. The result then follows from application of the upper bound (2).

B. Proof of Proposition 4

Following the discussion in Section 6, the signal sequence \mathbf{X}_{m-1} is quantized using the test channel $X_{m-1} = \hat{X}_{m-1} + Z_{q,m-1}$, where $Z_{q,m-1}$ is a complex Gaussian random variable with zero mean and variance σ_q^2 , independent of \hat{X}_{m-1} , which models the quantization error. In order to send the quantized signal \hat{X}_{m-1} to the m th BS, the following condition must be satisfied from standard rate-distortion theory results (the subscript "2" is dropped from the rate for simplicity of notation): $R_q = C - R \geq I(\hat{X}_{m-1}; X_{m-1}) = \log_2(P/\sigma_q^2)$, where $C - R \geq 0$ is the excess capacity on the m th link (recall that message W_m must be transmitted as well). From the previous equation, we can conclude that the variance of the

quantization error is $\sigma_q^2 = P/2^{C-R}$. The m th BS performs DPC over the quantized codeword represented by \hat{X}_{m-1} . In order to derive the rate achieved by DPC, we can write the received signal at the m th MS (1) as $Y_m = X_m + \alpha X_{m-1} + Z_m = X_m + \alpha \hat{X}_{m-1} + \tilde{Z}_m$, where \tilde{Z}_m is complex Gaussian with power $1 + \alpha^2 \sigma_q^2$ and is independent of X_m and \hat{X}_{m-1} . Therefore, recalling the property of DPC, we have that the achievable rate with the scheme at hand satisfies

$$R \leq \log_2 \left(1 + \frac{P}{1 + \alpha^2 \sigma_q^2} \right) = \log_2 \left(1 + \frac{P}{1 + (\alpha^2 P)/2^{C-R}} \right). \quad (\text{B.3})$$

From (B.3), if $C \leq \log_2(1 + P/(1 + \alpha^2 P))$, then the rate $R = C$, which corresponds to the upper bound (2), is clearly achievable. Otherwise, we can consider (B.3) with equality and solve the corresponding fixed-point equation. This leads to (14) and (15).

Acknowledgments

The research was supported in part by a Marie Curie Outgoing International Fellowship and the NEWCOM++ network of excellence both within the 7th European Community Framework Programme, by the U.S. National Science Foundation under Grants CNS-06-25637, CNS-06-26611, and ANI-03-38807, and also by the REMON Consortium.

References

- [1] M. K. Karakayali, G. J. Foschini, and R. A. Valenzuela, "Network coordination for spectrally efficient communications in cellular systems," *IEEE Wireless Communications*, vol. 13, no. 4, pp. 56–61, 2006.
- [2] O. Somekh, O. Simeone, Y. Bar-Ness, A. Haimovich, U. Spagnolini, and S. Shamai (Shitz), "An information theoretic view of distributed antenna processing in cellular systems," in *Distributed Antenna Systems: Open Architecture for Future Wireless Communications*, Auerbach, CRC Press, Boca Raton, Fla, USA, 2007.
- [3] A. D. Wyner, "Shannon-theoretic approach to a Gaussian cellular multiple-access channel," *IEEE Transactions on Information Theory*, vol. 40, no. 6, pp. 1713–1727, 1994.
- [4] H. Zhang and H. Dai, "Cochannel interference mitigation and cooperative processing in downlink multicell multiuser MIMO networks," *EURASIP Journal on Wireless Communications and Networking*, vol. 2004, no. 2, pp. 222–235, 2004.
- [5] W. Choi and J. G. Andrews, "The capacity gain from intercell scheduling in multi-antenna systems," *IEEE Transactions on Wireless Communications*, vol. 7, no. 2, pp. 714–725, 2008.
- [6] S. G. Kiani and D. Gesbert, "Optimal and distributed scheduling for multicell capacity maximization," *IEEE Transactions on Wireless Communications*, vol. 7, no. 1, pp. 288–297, 2008.
- [7] D. Park and G. Caire, "Hard fairness versus proportional fairness in wireless communications: the multiple-cell case," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT '08)*, pp. 2036–2040, Toronto, Canada, July 2008.
- [8] S. Shamai (Shitz) and B. M. Zaidel, "Enhancing the cellular downlink capacity via co-processing at the transmitting end," in *Proceedings of the IEEE Vehicular Technology Conference (VTC '01)*, vol. 3, pp. 1745–1749, Rhodes, Greece, May 2001.
- [9] O. Somekh, B. M. Zaidel, and S. Shamai (Shitz), "Sum rate characterization of joint multiple cell-site processing," *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4473–4497, 2007.
- [10] O. Somekh, O. Simeone, Y. Bar-Ness, and A. M. Haimovich, "Distributed multi-cell zero-forcing beamforming in cellular downlink channels," in *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM '06)*, pp. 1–6, San Francisco, Calif, USA, November 2006.
- [11] H. Dahrouj and W. Yu, "Coordinated beamforming for the multi-cell multi-antenna wireless system," in *Proceedings of the 42nd Annual Conference on Information Sciences and Systems (CISS '08)*, pp. 429–434, Princeton, NJ, USA, March 2008.
- [12] IEEE P802.16e/D11, September 2005, www.ieee802.org/16/tgm/contrib/C80216m-09_0970.doc.
- [13] S. Jing, D. N. C. Tse, J. B. Soriaga, J. Hou, J. E. Smee, and R. Padovani, "Downlink macro-diversity in cellular networks," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT '07)*, pp. 1–5, Nice, France, June 2007.
- [14] Y. Liang, T. Yoo, and A. Goldsmith, "Coverage spectral efficiency of cellular systems with cooperative base stations," in *Proceedings of the 40th Asilomar Conference on Signals, Systems and Computers (ACSSC '06)*, pp. 349–353, Pacific Grove, Calif, USA, October 2006.
- [15] N. Jindal, "MIMO broadcast channels with finite-rate feedback," *IEEE Transactions on Information Theory*, vol. 52, no. 11, pp. 5045–5060, 2006.
- [16] P. Marsch and G. Fettweis, "A framework for optimizing the uplink performance of distributed antenna systems under a constrained backhaul," in *Proceedings of the IEEE International Conference on Communications (ICC '07)*, pp. 975–979, Glasgow, Scotland, June 2007.
- [17] E. Aktas, J. Evans, and S. Hanly, "Distributed decoding in a cellular multiple-access channel," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT '04)*, p. 484, Chicago, Ill, USA, June-July 2004.
- [18] O. Sental, A. J. Weiss, N. Sental, and Y. Weiss, "Generalized belief propagation receiver for near-optimal detection of two-dimensional channels with memory," in *Proceedings of the IEEE Information Theory Workshop (ITW '04)*, pp. 225–229, San Antonio, Tex, USA, October 2004.
- [19] O. Simeone, O. Somekh, H. V. Poor, and S. Shamai (Shitz), "Local base station cooperation via finite-capacity links for the uplink of linear cellular networks," *IEEE Transactions on Information Theory*, vol. 55, no. 1, pp. 190–204, 2009.
- [20] B. L. Ng, J. Evans, and S. Hanly, "Distributed downlink beamforming in cellular networks," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT '07)*, pp. 6–10, Nice, France, June 2007.
- [21] F. Boccardi, H. Huang, and A. Alexiou, "Network MIMO with reduced backhaul requirements by MAC coordination," in *Proceedings of the Asilomar Conference on Signals, Systems and Computers (ACSSC '08)*, Pacific Grove, Calif, USA, October 2008.
- [22] A. Sanderovich, O. Somekh, and S. Shamai (Shitz), "Uplink macro diversity with limited backhaul capacity," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT '07)*, pp. 11–15, Nice, France, June 2007.
- [23] A. Sanderovich, S. Shamai (Shitz), Y. Steinberg, and G. Kramer, "Communication via decentralized processing," *IEEE*

- Transactions on Information Theory*, vol. 54, no. 7, pp. 3008–3023, 2008.
- [24] S. Shamai (Shitz), O. Simeone, O. Somekh, and H. V. Poor, “Joint multi-cell processing for downlink channels with limited-capacity backhaul,” in *Proceedings of the Information Theory and Applications Workshop (ITA ’08)*, pp. 345–349, San Diego, Calif, USA, January 2008.
- [25] O. Simeone, O. Somekh, H. V. Poor, and S. Shamai (Shitz), “Distributed MIMO systems with oblivious antennas,” in *Proceedings of the IEEE International Symposium on Information Theory (ISIT ’08)*, pp. 910–914, Toronto, Canada, July 2008.
- [26] S. Venkatesan, A. Lozano, and R. Valenzuela, “Network MIMO: Overcoming intercell interference in indoor wireless systems,” in *Proceedings of the Asilomar Conference on Signals, Systems and Computers (ACSSC ’07)*, pp. 83–87, Pacific Grove, Calif, USA, October 2007.
- [27] H. Dai, A. F. Molisch, and H. V. Poor, “Downlink capacity of interference-limited MIMO systems with joint detection,” *IEEE Transactions on Wireless Communications*, vol. 3, no. 2, pp. 442–453, 2004.
- [28] S. Verdú, “Spectral efficiency in the wideband regime,” *IEEE Transactions on Information Theory*, vol. 48, no. 6, pp. 1319–1343, 2002.
- [29] A. Lapidath, S. Shamai (Shitz), and M. A. Wigger, “A linear interference network with local side-information,” in *Proceedings of the IEEE International Symposium on Information Theory (ISIT ’07)*, pp. 2201–2205, Nice, France, June 2007.