Foundations and Trends® in
Communications and Information Theory
Vol. 8, Nos. 1-2 (2011) 1–177
© 2012 O. Simeone, N. Levy, A. Sanderovich,
O. Somekh, B. M. Zaidel, H. Vincent Poor
and S. Shamai (Shitz)
DOI: 10.1561/0100000048

now
the essence of knowledge

Cooperative Wireless Cellular Systems: An Information-Theoretic View

By Osvaldo Simeone, Nathan Levy,
Amichai Sanderovich, Oren Somekh,
Benjamin M. Zaidel, H. Vincent Poor
and Shlomo Shamai (Shitz)

Contents

Foundations and Trends® in
Communications and Information Theory
Vol. 8, Nos. 1-2 (2011) 1–177
© 2012 O. Simeone, N. Levy, A. Sanderovich,
O. Somekh, B. M. Zaidel, H. Vincent Poor
and S. Shamai (Shitz)
DOI: 10.1561/0100000048

now
the essence of knowledge

Cooperative Wireless Cellular Systems: An Information-Theoretic View

Osvaldo Simeone[1], Nathan Levy[2], Amichai Sanderovich[3], Oren Somekh[4], Benjamin M. Zaidel[5], H. Vincent Poor[6] and Shlomo Shamai (Shitz)[7]

[1] *New Jersey Institute of Technology (NJIT), Newark, USA, osvaldo.simeone@njit.edu*
[2] *Technion, Haifa 32000, Israel, nlevy@techunix.technion.ac.il*
[3] *Technion, Haifa 32000, Israel, amichi@tx.technion.ac.il*
[4] *Yahoo! Haifa Labs, Matam Industrial Park, Haifa 31905, Israel, orensomekh@yahoo.com*
[5] *Technion, Haifa 32000, Israel, benjamin.zaidel@gmail.com*
[6] *Princeton University, Princeton, USA, poor@princeton.edu*
[7] *Technion, Haifa 32000, Israel, sshlomo@ee.technion.ac.il*

Abstract

In this monograph, the impact of cooperation on the performance of wireless cellular systems is studied from an information-theoretic standpoint, focusing on simple formulations typically referred to as Wyner-type models. Following ongoing research and standardization efforts, the text covers two main classes of cooperation strategies. The first class is cooperation at the base station (BS) level, which is also known

as Multi-Cell Processing (MCP), network Multiple-Input Multiple-Output (MIMO), or Coordinated Multi-Point transmission/reception (CoMP). With MCP, cooperative decoding, for the uplink, or encoding, for the downlink, is enabled at the BSs. MCP is made possible by the presence of an architecture of, typically wired, backhaul links connecting individual BSs to a central processor (CP) or to one another. The second class of cooperative strategies allows cooperation in the form of relaying for conveying data between Mobile Stations (MSs) and BSs in either the uplink or the downlink. Relaying can be enabled by two possible architectures. A first option is to deploy dedicated Relay Stations (RSs) that are tasked with forwarding uplink or downlink traffic. The second option is for the MSs to act as RSs for other MSs.

MCP is first studied under ideal conditions on the backhaul links, namely by assuming that all BSs are connected to a CP with unlimited-capacity links. Both Gaussian (nonfading) and flat-fading channels are analyzed, for the uplink and the downlink, and analytical insights are drawn into the performance advantages of MCP in different relevant operating regimes. Performance comparison is performed with standard Single-Cell Processing (SCP) techniques, whereby each BS decodes, in the uplink, or encodes, in the downlink, independently, as implemented with different spatial reuse factors. Then, practical constraints on the backhaul architecture enabling MCP are introduced. Specifically, three common settings are studied. In the first, all the BSs are connected to a CP via finite-capacity links. In the second, only BSs in adjacent cells are connected via (finite-capacity) backhaul links. In the third, only a subset of BSs is connected to a CP for joint encoding/decoding (clustered cooperation). Achievable rates for the three settings are studied and compared for both the uplink and the downlink.

The performance advantages of relaying are analyzed for cellular systems with dedicated RSs and with cooperative MSs. Different techniques are reviewed that require varying degrees of information about system parameters at the MSs, RSs, and BSs. Performance is investigated with both MCP and SCP, revealing a profound interplay between cooperation at the BS level and relaying. Finally, various open problems are pointed out.

1

Introduction

In this section, we provide a brief introduction to the subject of this monograph and discuss the goals and outline of the text.

1.1 Motivation

Cooperative communication refers to the coordinated transmission or reception by some nodes of a given communication network (see, e.g., [48]). In the context of wireless cellular systems, cooperation is being currently studied in both academic research activities and standardization efforts in two different forms.

- ***Multi-Cell Processing (MCP)***: The first type of cooperative strategies is at the base station (BS) level, and is known as Multi-Cell Processing (MCP), network Multiple-Input Multiple-Output (MIMO) [122] or Coordinated Multi-Point transmission/reception (CoMP), with the last term being used in the 4th Generation (4G)-Long Term Evolution (LTE)-Advanced cellular standard (see, e.g., [3]). With MCP, cooperative decoding, for the uplink, or encoding, for the downlink, is enabled at the BSs. MCP is made possible

by the presence of an architecture of, typically wired, backhaul links connecting individual BSs to a central processor (CP), or to one another.

- ***Relaying***: The second class allows cooperation in the form of relaying for conveying data between Mobile Stations (MSs) and BSs in either the uplink or downlink. Relaying can be enabled by two possible architectures. A first option is to deploy *dedicated Relay Stations (RSs)* that are tasked with forwarding uplink or downlink traffic. The second option is for the MSs to act as RSs for other MSs.

The two approaches have different merits and expected performance gains, as discussed below. Overall, they are seen as extremely promising strategies to overcome the current problem of "bandwidth crunch" affecting cellular systems due to the ever increasing capacity demands [19].

By enabling joint encoding and decoding across multiple cells, MCP has the capability to turn inter-cell interference from one of the main limitations on the system performance, as it is in conventional non-cooperative cellular systems, into an asset. For instance, focusing on the downlink, thanks to inter-cell "interference", MCP enables all the BSs to communicate to any MS in the system in a cooperative fashion. This allows the BSs to control, and potentially cancel, inter-cell interference, and thus to serve the MSs at a rate that is not limited by such interference. Similar considerations apply to the uplink as well.

Beside interference mitigation, through BS cooperation, MCP allows *beamforming, diversity*, and *multiuser diversity* gains to be harnessed. The first of these gains refers to the possibility of performing coherent decoding or encoding across multiple BSs so as to boost the effective signal-to-noise ratio (SNR) in the uplink or downlink. The second, diversity, refers to the possibility of leveraging different signal paths from transmitter to receiver in order to increase the probability the transmitted signal is received with a sufficiently large SNR. The third, multiuser diversity, accounts for the design degrees of freedom afforded by the ability to schedule different users depending on their channel conditions across the whole network.

Relaying, instead, is mostly seen as a way to extend the coverage of a given BS by enabling multihop transmission between an MS and a BS (see, e.g., [55, 110]). Equivalently, it can also be seen as a means to reduce the required transmission energy in the uplink or downlink for given SNR requirements. Relaying also enables beamforming and diversity gains to be harnessed via cooperation between MSs or between MSs and RSs (see, e.g., [51]).

1.2 Approach and Goals

The analysis of MCP was initiated in the early works [33, 132] for the uplink and in [96] for the downlink. The analysis in these works is based on the assumption that the BSs are connected via unrestricted backhaul links (*error-free and unlimited capacity*) to a CP and focuses on models that, in information-theoretic terms, can be seen as *symmetric Gaussian multiple access or broadcast interference channels*. In these models, typically referred to as *Wyner-type models*, a number of users per cell are served by a single-antenna BS, as in a multiple access or broadcast channel, and interference takes place only between adjacent cells, as in partially connected interference networks. Both the models where cells are arranged along a line or in a more conventional bidimensional geometry can be considered, where the first class may model systems deployed along a highway, railroad, or long corridor (see [122] for an implementation-based study), while the second applies to more general scenarios.

Wyner-type models are simple abstractions of cellular systems. They capture well one of the main aspects of such settings, namely the locality of inter-cell interference. The advantage of addressing the study of given transmission strategies on such models is the possibility to obtain analytical insights. These insights provide an invaluable stepping stone for the simulation-based studies that are necessary for a full performance assessment under more realistic operating conditions (see, e.g., [135] for further considerations on this point).

In this monograph, we aim at providing an information-theoretic view of the advantages of cooperation in wireless cellular systems in terms of MCP, relaying and their interplay. In order to enable analysis,

we will adopt Wyner-type models. The treatment reviews a number of results available in the literature in a unified fashion that reveals their connections and illuminates general conclusions. We will keep the treatment as self-contained as possible, but we will privilege intuition over analytical details and technicalities. In particular, we will not provide any detailed proof. In doing so, our goal here is to provide an understanding of the performance of the cooperative techniques at hand and of the analytical tools used for this purpose.

As research in the field is still ongoing, the treatment will be far from complete, and we will point to open problems along the way. Nevertheless, we feel that the available results are mature and complete enough to warrant the treatment given in this text. We will provide a clear pointer to the main references used in compiling this monograph. However, we will not attempt to provide a comprehensive bibliography on the subject of cooperation for cellular system. It should also be emphasized that the focus here is based purely on information-theoretic arguments and is limited to Wyner-type models. In particular, we will leave out discussion of other more complex models and of important issues such as the signal-processing aspects of optimal beamforming and power allocation. A more complete list of references in this regard can be found in [28]. Previous shorter tutorials can be found in [92, 93].

1.3 Outline

The text is organized as follows. In Section 2, the models that will be adopted for analysis throughout the monograph are presented. Then, in Section 3, results are discussed that provide the performance of MCP under ideal conditions on the backhaul links, namely by assuming that all BSs are connected to a CP with unlimited-capacity links, and without accounting for fading. Section 4 extends these results from the Gaussian model studied in the previous section to flat-fading channels for both uplink and downlink. Performance comparison is performed with standard Single-Cell Processing (SCP) techniques, whereby each BS decodes, in the uplink, or encodes, in the downlink, independently, as implemented with different spatial reuse factors. In Sections 5 and 6, practical constraints on the backhaul architecture enabling MCP are

introduced for the uplink and downlink, respectively. Specifically, three common settings are studied. In the first, all BSs are connected in a CP via finite-capacity links. In the second, only BSs in adjacent cells are connected via (finite-capacity) backhaul links. In the third, only a subset of nearby BSs is connected to a CP for joint encoding/decoding (clustered cooperation). The performance advantages of relaying are finally analyzed for cellular systems with dedicated RSs (Section 7) and with cooperative MSs (Section 8) over Gaussian channels. Different techniques are proposed that require varying degrees of information about system parameters at the MSs, RSs, and BSs. Performance is analyzed with both MCP and SCP, revealing a profound interplay between cooperation at the BSs and relaying.

2

Cellular Models

In this section, we review the basic system models that will be considered throughout this monograph. We focus our attention on *linear* Wyner-type models, as done in the original works [33, 132]. Note that extension of the given results to planar models is possible, though not always straightforward, and we refer to [113] for further discussion on this point. In the following, we first introduce standard Wyner-type models for the uplink and downlink, and then present extensions that encompass practical limitations on the backhaul links and for relaying.

2.1 Uplink

A linear Wyner-type model, sketched in Figure 2.1, is characterized by M cells arranged along a line (as for a highway, railway, or corridor), each with a single-antenna BS and K single-antenna MSs. In this class of models, inter-cell interference at a given BS is limited to L_ℓ BSs to the left and L_r to the right of the BS at hand (left and right are taken with respect to the reader's viewpoint). Considering the uplink, the baseband received signal at the m-th BS, $m \in [1, M]$, at a given time instant $t \in [1, n]$ (n is the size of the transmitted block) can then be

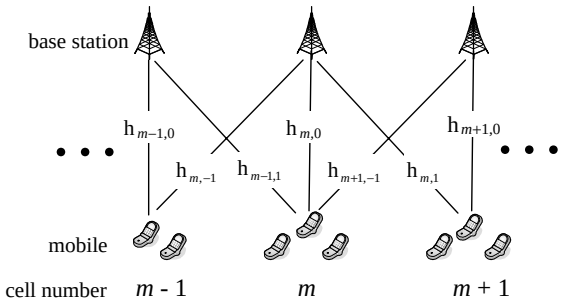

Fig. 2.1 A linear Wyner-type model with $L_r = L_\ell = 1$ and $K = 3$.

written as

$$y_m(t) = \sum_{l=-L_\ell}^{L_r} \mathbf{h}_{m,l}^T(t)\mathbf{x}_{m+l}(t) + z_m(t), \tag{2.1}$$

where $\mathbf{x}_m(t)$ is the $K \times 1$ (complex) vector of signals transmitted by the K MSs in the m-th cell, the $K \times 1$ vector $\mathbf{h}_{m,k}(t)$ contains the channel gains toward the m-th BS from mobiles placed k cells apart (see Figure 2.1 for an illustration) and $z_m(t)$ is complex Gaussian noise with unit power, assumed to be uncorrelated over BS index m and time t.[1] Denoting as $[\mathbf{x}]_k$ the k-th element of vector $\mathbf{x}$, we assume equal per-user *per-block* power constraints

$$\frac{1}{n}\sum_{t=1}^{n} \mathbb{E}[|[\mathbf{x}_m(t)]_k|^2] \le \frac{P}{K}, \tag{2.2}$$

for all $m \in [1, M]$ and $k \in [1, K]$, so that the *per-cell per-block power constraint* is given by P. Notice that model (Equation (2.1)) assumes full frame and symbol-level synchronization among all MSs, even though extensions of the available results may be possible in the asynchronous case following, e.g., [123]. Also, note that in some scenarios, it may be more relevant to impose a *per-symbol*, rather than per-block, power constraint, as in $\mathbb{E}[|[\mathbf{x}_m(t)]_k|^2] \le \frac{P}{K}$ for all $t \in [1, n]$. The latter will be considered only when explicitly mentioned.

[1] Spatially correlated noise could arise, for instance, due to interference from other wireless systems. However, if such correlation is known at the receivers, it can be compensated via whitening leading to model (Equation (2.1)).

The model (Equation (2.1)) discussed above reduces to the following special cases that will be referred to throughout the monograph:

- ***Gaussian Wyner model***: This corresponds to a static scenario with symmetric intercell interference and cell-homogeneous channel gains, i.e., we have $L_\ell = L_r = L$ and $\mathbf{h}_{m,k}(t) = \alpha_k \mathbf{1}$ $(-L \leq k \leq L)$ with $\alpha_k = \alpha_{-k}$ and $\alpha_0 = 1$. We emphasize that these gains are assumed to be real for simplicity of presentation, but that most results are easily extended to complex-valued inter-cell gains. By cell-homogeneous, we mean that the channel gains do not depend on the cell index m, but only on the distance between cells. Note that this implies that all users in the same cell share the same path loss toward all BSs, as it is approximately the case if all users are located at the center of the cell. This model is clearly a crude approximation of reality, but it captures, at least to a first order, the locality of intercell interference. The parameter L is referred to as the *intercell interference span*, as it measures how many cells on the left or right of any cell affect reception at the corresponding BS. We will often refer to the original Wyner model in [33, 132] which corresponds to the special case $L = 1$, for which we have, denoting α_1 as α for simplicity, the baseband received signal as

$$y_m(t) = \mathbf{1}^T \mathbf{x}_m(t) + \alpha \mathbf{1}^T \mathbf{x}_{m+1}(t) + \alpha \mathbf{1}^T \mathbf{x}_{m-1}(t) + z_m(t). \tag{2.3}$$

  Note that $\mathbf{1}^T \mathbf{x}$ is simply the sum of all elements in the vector $\mathbf{x}$. This model is sketched in Figure 2.2 for reference.
- ***Gaussian soft-handoff model***: This corresponds to a static cell-homogeneous system like the Gaussian Wyner model, in which, however, there is no symmetry in the intercell channel gains. Specifically, we have intercell interference only from the cells to the left as $L_\ell = L$, $L_r = 0$ and $\mathbf{h}_{m,k}(t) = \alpha_k \mathbf{1}$, $(0 \leq k \leq L)$ with $\alpha_0 = 1$, where, as above, $\alpha_k \geq 0$ are deterministic (real) quantities. A standard scenario is the special case $L = 1$, studied, e.g., in [43, 118], for which the

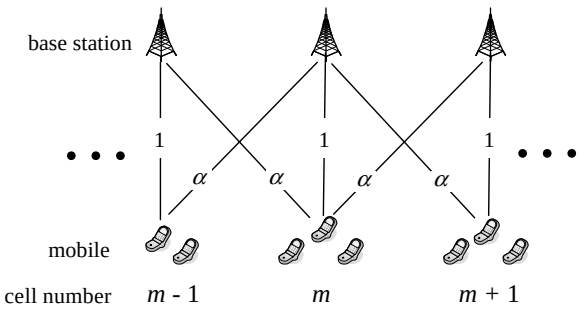

Fig. 2.2 The Gaussian linear Wyner model with $L = 1$ and $K = 1$.

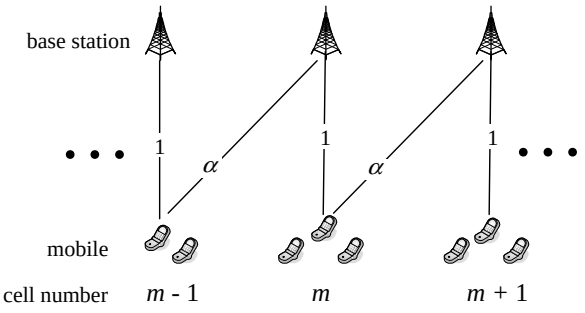

Fig. 2.3 The Gaussian linear soft-handoff model with $L = 1$ and $K = 1$.

baseband received signal is given by

$$y_m(t) = \mathbf{1}^T \mathbf{x}_m(t) + \alpha \mathbf{1}^T \mathbf{x}_{m-1}(t) + z_m(t), \qquad (2.4)$$

where we have again denoted α_1 as α for simplicity. This special case clarifies the name "soft-handoff". In fact, model (Equation (2.4)) accounts for a simple scenario where the MSs are placed at the border between two cells, so that the signal transmitted by each MS is received with nonnegligible power only by two BSs, the local BS and the BS on the right (see Figure 2.3);

- ***Fading Wyner model***: This model incorporates fading, accounted for by random channel gains $\mathbf{h}_{m,k}(t)$, in the Gaussian Wyner model. In particular, we have $L_\ell = L_r = L$ and $\mathbf{h}_{m,k}(t) = \alpha_k \tilde{\mathbf{h}}_{m,k}(t)$ $(-L \le k \le L)$, where vectors $\tilde{\mathbf{h}}_{m,k}(t)$, $t \in [1, n]$, are independent over m and k and distributed

according to a joint distribution π_k with power of each entry of $\tilde{\mathbf{h}}_{m,k}(t)$ normalized to one. For simplicity, similar to the Gaussian Wyner model, *statistical* symmetric intercell interference is assumed, i.e., $\alpha_k = \alpha_{-k}$ (and $\alpha_0 = 1$) and $\pi_k = \pi_{-k}$. As for temporal variations, two scenarios are typical: (*i*) *Quasistatic fading*: Channels $\tilde{\mathbf{h}}_{m,k}(t)$ are constant over the transmission of a given codeword (i.e., for $t \in [1, n]$); and (*ii*) *Ergodic fading*: Channels $\tilde{\mathbf{h}}_{m,k}(t)$ vary in an ergodic fashion along the symbols of the codeword. The ergodic model was studied in [111] with $L = 1$; and

- ***Fading soft-handoff model***: This model is the fading counterpart of the Gaussian soft-handoff model, and has $L_\ell = L$, $L_r = 0$, and $\mathbf{h}_{m,k}(t) = \alpha_k \tilde{\mathbf{h}}_{m,k}(t)$ $(0 \le k \le L)$ where $\tilde{\mathbf{h}}_{m,k}(t)$ are independent and modeled as for the fading Wyner model. This scenario was considered in [61] (under more general conditions on the joint distribution of vectors $\tilde{\mathbf{h}}_{m,k}(t)$).

Where not stated otherwise, for both fading Wyner and fading soft-handoff models, we assume that the fading gains are modelled according to Rayleigh fading, i.e., each entry of $\tilde{\mathbf{h}}_{m,k}(t)$ is complex Gaussian with zero mean and unitary power. In the rest of the monograph, we will concentrate the analysis on Gaussian and fading Wyner and soft-handoff models.

It is remarked that the purpose of the Wyner model is not that of proving a means for an exact quantitative prediction of the actual performance of a real-life cellular system. Rather, its adoption is aimed at providing insights and general guidelines for the design of cellular systems that rely on cooperative processing. Moreover, the assumption of a limited intercell interference span made in the Wyner models is in practice a reasonable approximation of reality up to a certain SNR level. In fact, only if the interference is sufficiently below the noise level, can it be potentially neglected without affecting the analysis (see, e.g., [68]). Therefore, the results obtained in the rest of the monograph have always to be considered as valid within the SNR range in which the Wyner model is a reasonable approximation.

2.1.1 Edge Effects

In order to obtain compact and insightful analytical results, it is often convenient to neglect edge effects. This can be done in two ways. A first way is to focus on the regime of large number of cells, i.e., $M \to \infty$. This way, virtually all cells see exactly the same intercell interference scenario, possibly in a statistical sense, as discussed above.

An alternative approach, considered, e.g., in [33, 118], is to consider a system in which cells are placed on a circle. Specifically, the system is as described above, but the cell on the right of the first one ($m = 1$) is considered to be the last cell ($m = M$), and the cell on the right of the last one ($m = M$) is considered to be the first one ($m = 1$). This way, each cell is affected by statistically homogeneous intercell interference for any finite M. It is noted that the two models coincide in the limit of large M and, in practice, results for the two models are very close for relatively small values of M [113].

2.1.2 Matrix Formulation

We now rewrite the model (Equation (2.1)) in a more compact matrix form. We drop dependence on time t for simplicity. To proceed, construct an $M \times MK$ channel matrix $\mathbf{H}$ such that the m-th row collects all channel gains to the m-th BS, i.e., $[\mathbf{h}_{m,m-1}^T, \mathbf{h}_{m,m-2}^T, \ldots, \mathbf{h}_{m,0}^T, \mathbf{h}_{m,-1}^T, \ldots, \mathbf{h}_{m,-(M-m-1)}^T]$, where $\mathbf{h}_{m,k}^T$ with $k \notin [-L_r, L_\ell]$ are to be considered as zero. We can then write the $M \times 1$ vector of received signals $\mathbf{y} = [y_1, \ldots, y_M]^T$ as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{z}, \tag{2.5}$$

where $\mathbf{x} = [\mathbf{x}_1^T \cdots \mathbf{x}_M^T]^T$ is the vector of transmitted signals and $\mathbf{z}$ the uncorrelated vector of unit-power complex Gaussian noises. From the definition above, it is clear that, in general, $\mathbf{H}$ is a finite-band matrix (in the sense that only a finite number of diagonals have nonzero entries). Moreover, it is not difficult to see that for the Gaussian Wyner and Gaussian soft-handoff models, the matrix $\mathbf{H}$ has a block-Toeplitz structure, which will be useful in the following.

2.1.3 Channel State Information (CSI)

For the uplink, we assume that full CSI is available at the BSs, and also at the CP when MCP is enabled. On the other hand, the MSs are only informed about the transmission strategy, namely the transmission codebook, and in particular the transmission rate and power to be used for transmission. In other words, the MSs cannot modify their operation based on the current channel conditions. Therefore, for the uplink, due to practical constraints, we do not consider the possibility for sophisticated channel-aware scheduling strategies.

We emphasize that assessing the impact of imperfect CSI is of critical importance in comparing the performance of different system designs. To this end, an analysis in terms of performance lower bounds for the uplink can be performed following the standard approach of [35] of treating the CSI error as an additional source of Gaussian noise. This way, the analysis presented in this monograph can be adapted to account also for imperfect CSI. However, we leave a full investigation of this important point to future work and to related studies such as [38, 68].

2.2 Downlink

The model for the downlink can be easily obtained from the discussion above. We will reuse some symbols, since their meaning will be clear from the context. Define as $\mathbf{y}_m$ the $K \times 1$ vector of signals received by the K MSs in the m-th cell, $\mathbf{y} = [\mathbf{y}_1^T \cdots \mathbf{y}_M^T]^T$, and $\mathbf{x}$ as the $M \times 1$ transmitted signal by the BSs. We then have

$$\mathbf{y} = \mathbf{H}^\dagger \mathbf{x} + \mathbf{z}, \tag{2.6}$$

where $\mathbf{z}$ is the vector of unit-power uncorrelated complex Gaussian noise and the channel matrix $\mathbf{H}$ is defined as above. We assume a *per-BS (and thus per-cell) per-block power constraint* $\frac{1}{n}\sum_{t=1}^{n} \mathrm{E}[|[\mathbf{x}(t)]_m|^2] \leq P$ for all $m \in [1, M]$. When explicitly mentioned, a per-BS per-symbol power constraint $\mathrm{E}[|[\mathbf{x}(t)]_m|^2] \leq P$, for all $t \in [1, n]$, will also be considered. Model Equation (2.6) is the dual of Equation (2.5) in the sense discussed below and in Appendix B.

For reference, the signal received by the k-th MS in the m-th cell in the downlink Gaussian Wyner model with $L = 1$ is given by

$$[\mathbf{y}_m(t)]_k = x_m(t) + \alpha x_{m+1}(t) + \alpha x_{m-1}(t) + z_m(t), \tag{2.7}$$

while for the Gaussian soft-handoff model with $L = 1$ we have

$$[\mathbf{y}_m(t)]_k = x_m(t) + \alpha x_{m+1}(t) + z_m(t), \tag{2.8}$$

where we recall that $x_m(t)$ is the signal transmitted by the m-th BS.

2.2.1 Channel State Information (CSI)

We assume, unless stated otherwise, that the BSs and the CP have full CSI, while the MSs are informed about the channels affecting their respective received signals and about the transmission strategy employed for transmission by the BSs, and by the CP if MCP is enabled. We note that assessing the impact of imperfect CSI on the downlink is even more critical than for the uplink. However, this analysis turns out to be more complex than for the uplink and has been mostly been dealt with in simulation-based studies (see, e.g., [74, 84]). We remark that, even in the case of ideal BS cooperation and focusing only on the multiplexing gain, the problem of assessing the impact of imperfect CSI at the transmitter is open [54]. However, results for simplified settings suggest that, with a CSI quality that scales sufficiently fast with the SNR, the performance of the system is not significantly degraded as compared to the case of ideal CSI in the high-SNR regime [11]. Here we leave this important issue to future work.

2.3 Backhaul Architectures for Multi-Cell Processing

In order for MCP to be possible, the BSs must be connected via backhaul links. For both the uplink and downlink, we will consider the three following models for the backhaul connections, as illustrated in Figure 2.4.

- ***Central processor (CP) with finite-capacity backhaul*** (Figure 2.4(a)): In this case, all the BSs are connected to a CP for joint decoding (for the uplink) or encoding (for the

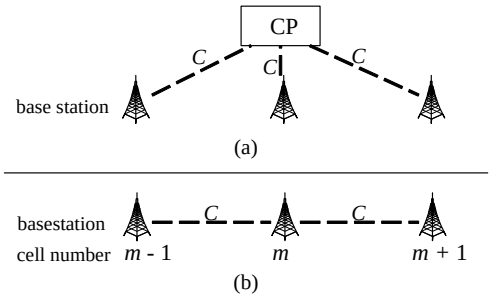

Fig. 2.4 Backhaul architectures that enable Multi-Cell Processing.

downlink) via finite-capacity backhaul links of capacity C (bits/s/Hz), where bandwidth normalization is done with respect to the spectrum available for the uplink or downlink transmission. Recall that the original works [33, 96] assume unlimited backhaul capacity, i.e., $C \to \infty$, whereas here we will consider also the effect of a finite backhaul capacity C;

- ***Local finite-capacity backhaul between adjacent BSs*** (Figure 2.4(b)): Here the BSs are connected only to their neighboring BSs, via finite-capacity links of capacity C (bits/s/Hz), which may be uni- or bidirectional;
- ***Clustered Cooperation***: Multiple CPs are available, one for each cluster of cells. Clusters of BSs are connected to a CP for joint encoding/decoding. This scenario is similar to the first, with the difference that here not all interfering cells can be jointly processed at the CP, although no capacity limitation is imposed on the individual links between BSs and their CP. This is done so as to focus on the impact of finite-size BS clusters on the performance of MCP (see [91] for a model that considers both clustered cooperation and finite-capacity backhaul links).

It is noted that the first two models coincide in the case of unlimited backhaul capacity $C \to \infty$.

A remark on the choice made here to constrain the capacity C of the backhaul links is in order. It might seem that, at least for conventional cellular systems, the backhaul capacity required for macro BSs is not of

primary concern, and can be considered infinite for practical purposes, due to the availability of fiber optic cables. A first consideration in this regard is that, as of today, backhaul links are still for a sizable fraction copper or wireless [64]. Moreover, copper and wireless links are overwhelmingly used in heterogeneous cellular to connect the femto- and pico-BSs to the network (see, e.g., [73]).

A second observation has to do with the capacity needs for transmission of compressed baseband information from the BS to the network. This approach is central to many proposals for MCP [64]. For instance, for a single base station supporting three carriers of Wideband Code Division Multiple Access (W-CDMA) plus 20 MHz LTE, one needs about 7.7 Gb/s even using advanced compression strategies [64]. This fills nearly the capacity of an entire 10 Gigabit Passive Optical Network, thus having very serious implications on the transmission layer of the system. In fact, this type of capacity requirement is orders of magnitude larger compared to current systems in which local decoding is performed. Therefore, even with fiber optic cables the issue of backhaul capacity limitations is of critical importance. We finally remark that an alternative solution to this problem is based on radio over fiber (see, e.g., [89]).

2.4 Linear Wyner Models with Relaying

In modern cellular systems, cooperation is advocated not only for joint encoding/decoding at the BS level, but also in terms of relaying. There are two main scenarios where relaying is enabled:

- ***Dedicated relay stations***: Dedicated RSs are deployed in fixed locations within each cell, and are tasked with relaying traffic from the MSs to the BSs in the uplink, or from the BSs to the MSs in the downlink.
- ***Mobile station relaying***: Relaying in the uplink or downlink is performed by MSs for traffic originating at or intended for other MSs.

In the following we describe the models that will be considered in this monograph for the assessment of the performance gains of the two

approaches at hand. In order to keep the discussion focused, we will concentrate solely on the uplink.

2.4.1 Linear Wyner Model with Dedicated Relay Stations

Here we describe the model that will be studied to assess the performance advantages of the deployment of dedicated RSs. We focus on the uplink and concentrate on the performance of users that are at the edge of each cell. This is done since the main benefits of relay deployment are expected to occur for MSs that happen to be far from the BS, and thus need to transmit at high power in order to be received at sufficient SNR. In fact, dedicated RSs, if conveniently located between an MS and a BS provide the possibility for a two-hop transmission from the MS to the BS, which can reduce the necessary transmitted power and hence increase coverage.

To model transmission from edge-cell MSs in a way that is amenable to analysis, we consider the model in Figure 2.5, where the direct link from MS to BS is assumed to be so small that it can be, as a first approximation, neglected. Moreover, we assume that a single RS is present in each cell and is placed between the edge-cell MSs and the BS. The model thus reduces to a cascade of two Wyner-type models, one from the MSs to the RSs and one from the RSs to the BSs (see Figure 2.5). We will refer to these two Wyner models as *first and second*

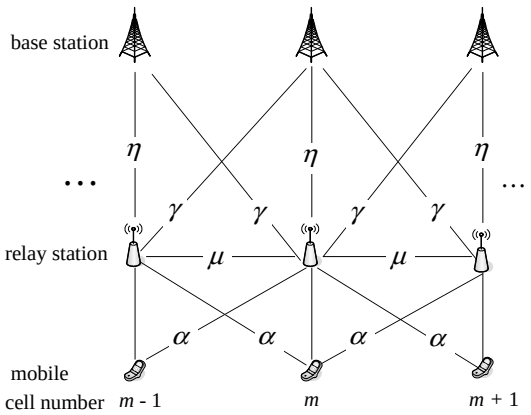

Fig. 2.5 Gaussian Linear Wyner model with dedicated relay stations ($L = 1, K = 1$).

hop, respectively. We also assume, where not stated otherwise, that the relays operate in full-duplex mode. Recent results on the feasibility of full-duplex transmission can be found in [21].

To simplify the discussion, we present the main system equations for the case where both hops are Gaussian Wyner models with $L = 1$ and there is only one user active in each cell ($K = 1$). Based on the discussion above, the extension to more general models for the two hops is immediate. For the first hop, the baseband signal received by the RS in the m-th cell at time t is given by (cf. Equation (2.3))

$$\begin{aligned} y_m^{(1)}(t) = x_m^{(1)}(t) + \alpha x_{m-1}^{(1)}(t) + \alpha x_{m+1}^{(1)}(t) + \mu x_{m+1}^{(2)}(t) \\ + \mu x_{m-1}^{(2)}(t) + z_m^{(1)}(t), \end{aligned} \tag{2.9}$$

where the superscript identifies the hop, so that $x_m^{(1)}(t)$ is the signal transmitted by the m-th MS (in the first hop), $x_m^{(2)}(t)$ is the signal transmitted by the m-th RS (in the second hop) and $z_m^{(1)}(t)$ is complex Gaussian noise affecting reception at the m-th RS, assumed to have power σ_1^2. Note that from Equation (2.9) reception at the RS in the m-th cell is affected by the inter-cell interference from MSs *and* RSs in adjacent cells (see Figure 2.5).[2] Similarly the baseband signal received at the m-th BS is given by

$$y_m^{(2)}(t) = \eta x_m^{(2)}(t) + \gamma x_{m-1}^{(2)}(t) + \gamma x_{m+1}^{(2)}(t) + z_m^{(2)}(t), \tag{2.10}$$

where definitions follow as above and $z_m^{(2)}(t)$ is complex Gaussian with power σ_2^2.

Finally, power constraints are imposed on the MSs as (cf. Equation (2.12))

$$\frac{1}{n}\sum_{t=1}^{n} \mathbb{E}[|x_m^{(1)}(t)|^2] \leq P, \tag{2.11}$$

[2] The fact that the path loss is the same toward the adjacent BSs is not at odds with the assumption that the MS is at the cell edge if one considers cells that have some extension on the plane. In fact, the MS can be facing its BS in a direction orthogonal to the deployment of cells and thus still be in a symmetric position with respect to the neighboring cells.

and on the RSs as

$$\frac{1}{n}\sum_{t=1}^{n}\mathbb{E}[|x_m^{(2)}(t)|^2] \leq Q, \tag{2.12}$$

Full CSI is assumed at both transmitters and receivers.

2.4.2 Linear Wyner Model with Mobile Cooperation

In systems where dedicated RSs are not available, it is still possible for MSs to obtain gains from relaying as long as the MSs are enabled to forward traffic on behalf of other MSs. In order for MSs to perform relaying, it is necessary for the relaying MS to obtain some information about the signal or the message of the MS whose signal should be forwarded to the BS. This can take place in two main ways:

- ***Out-of-band Cooperation***: Each MS is able to communicate with nearby MSs using a different radio interface than the one used for communication to the BS. For instance, inter-MS communication may take place over a Wi-Fi link, while transmission to the BS may be over a 3G or 4G cellular link.
- ***In-band Cooperation***: Each MS is able to receive in the same bandwidth used for communication to the BSs. This way, each MS can overhear the signals sent by neighboring MSs. Note that here we assume full-duplex transmission so that each MS is able to transmit and receive at the same time.

MSs can communicate, and hence cooperate, with other MSs in the same cell and/or in different cells. Note that the latter case is practically relevant for MSs close to the edge of the cells. The performance advantages achievable by cooperating with same-cell MSs are akin to the relaying gains attainable with dedicated RSs. Indeed, in both cases, performance gains arise due to the potential power savings due to multi-hop communication toward the same-cell BS. Instead, cooperation with MSs in other cells enables a different type of gains when implemented along with MCP. In fact, the signal of any m-th MS, when forwarded

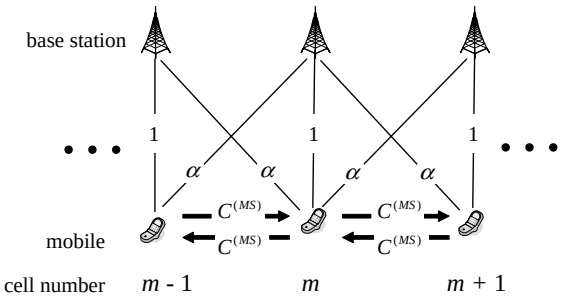

Fig. 2.6 A linear Gaussian Wyner model with out-of-band MS cooperation.

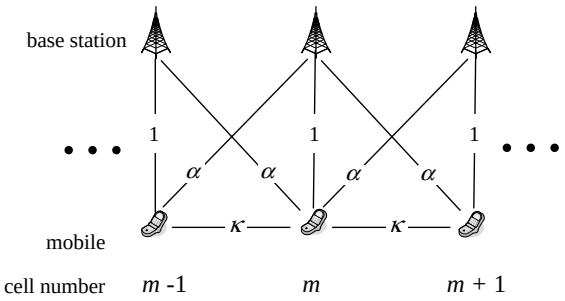

Fig. 2.7 A linear Gaussian Wyner model with in-band MS cooperation.

by an MS in a different cell, say the $(m+1)$-th, will be received with large power by the $(m+1)$-th BS. With MCP, this transmission creates an alternative path for the signal of the m-th MS to reach the CP. Moreover, such path can combine constructively with the direct one in order to provide coherent (beamforming) gains.

We consider two models for inter-cell cooperation. The first, modelling *out-of-band cooperation,* is shown in Figure 2.6. Here MSs in adjacent cells are connected with orthogonal links of capacity $C^{(MS)}$ (bits/s/Hz) in either direction. In the second, modelling *in-band cooperation*, the signal received by the m-th MS is given by

$$y_m^{(MS)}(t) = \kappa x_{m+1}(t) + \kappa x_{m-1}(t) + z_m^{(MS)}(t), \tag{2.13}$$

where $\kappa \geq 0$ is a channel gain modelling the quality of the inter-MS links and $z_m^{(MS)}(t)$ is the unit-power complex Gaussian noise affecting reception at the m-th MS. This is shown in Figure 2.7. Full CSI is assumed at both transmitters and receivers.

3

Multi-Cell Processing in Gaussian Channels

In this section, we elaborate on the *per-cell sum-rate* achievable for the uplink of Wyner-type models without relaying. We will focus on the *Gaussian Wyner model,* while fading models are discussed in Section 4. Note that, as discussed in the previous section, the works [33, 132] consider only the case $L = 1$ (in Equation (2.1)) so the analysis presented here is more general.

The per-cell sum-rate provides an important performance metric, since it measures the overall amount of information carried by the network per cell. Moreover, in symmetric channels such as the Wyner and soft-handoff models studied here, achievability of a per-cell sum-rate R implies also achievability of a per-MS rate equal to R/M. It is noted that, in some analyses, one might be interested in a more refined understanding of the system performance in terms of achievable rate regions for all users in the system or for given groups of users in the system (e.g., users at the border of the cell versus users in the interior of the cell). This analysis is not considered in this monograph and is mostly a problem that remains open for research.

3.1 Uplink

In this section, we analyze the uplink of a Gaussian Wyner model.

3.1.1 Single-Cell Processing and Spatial Reuse

Consider at first a baseline scheme, where SCP is performed, so that each BS decodes individually its own K users by treating users in other cells as Gaussian noise. A standard technique to cope with intercell interference is *spatial reuse*, which consists of activating at a given time (or equivalently in a given subband) only one cell every $F \geq 1$ cells. The parameter F is referred to as the *spatial reuse factor*. SCP with spatial reuse is easily seen to achieve the following per-cell sum-rate (in bits/s/Hz):

$$R_{SCP}(P,F) = \begin{cases} \frac{1}{F}\log_2\left(1+\frac{FP}{1+2FP\sum_{k=1}^{\lfloor L/F \rfloor}\alpha_{kF}^2}\right) & \text{if } F \leq L \\ \frac{1}{F}\log_2(1+FP) & \text{if } F > L \end{cases}, \tag{3.1}$$

where L is the intercell interference span.

Rate (Equation (3.1)) is obtained by either one of the following techniques:

- *Wide-Band (WB) transmission*: All users in a given cell transmit at the same time with power FP/K. We emphasize that, since we focus on frequency-flat channels, a WB transmission does not imply frequency selectivity here, but only transmission across all the available spectral resources.
- *Intra-cell Time Division Multiple Access (TDMA)*: Each user in a cell transmits with power FP for a fraction of time $1/K$.

That both techniques attain rate (Equation (3.1)) follows from the well-known results on the sum-capacity of a Gaussian multiple access channel [17]. Moreover, from such results, it can be seen that, with both WB and intra-cell TDMA, users can be allocated an equal share of the overall per-cell sum-rate. Moreover, WB allows for more flexibility

in the rate allocation, while still obtaining Equation (3.1), whereas intra-cell TDMA is not able to support different users' rates without performance loss.

It is also noted that achievability of Equation (3.1) hinges on the per-block power constraint. In fact, the power allocations above satisfy the per-block power constraint (Equation (2.12)), due to the fact that each cell transmits for a fraction $1/F$ of the time and, for intra-cell TDMA, users are active for a fraction $1/FK$ of the time. With a *per-symbol* power constraint of P/K per user, instead, the per-cell sum-rate of SCP would be limited to

$$\begin{cases} \frac{1}{F}\log_2\left(1+\frac{aP/K}{1+2P\sum_{k=1}^{\lfloor L/F\rfloor}\alpha_{kF}^2}\right) & \text{if } F\le L \\ \frac{1}{F}\log_2(1+aP/K) & \text{if } F>L \end{cases}, \tag{3.2}$$

where $a=K$ for WB and $a=1$ for intra-cell TDMA. Therefore, with per-symbol power constraints, WB is advantageous over intra-cell TDMA.

Finally, we remark that, while still considering SCP processing, performance could be in principle improved by informing each BS of the codebooks of the interfering MSs belonging to other cells. This would allow each BS to perform joint decoding of the useful and interfering signals, rather than necessarily treating interference as noise. We will discuss related techniques in Section 4 and we refer to references [149] and [67] for application to the Gaussian Wyner and soft-handoff models, respectively.

3.1.1.1 Extreme-SNR Analysis

In order to get further insight into the performance of SCP, we resort here, as will be often done in this monograph, to the analysis of extreme SNR conditions, namely high- and low-SNR regimes.

By analyzing the performance in the *high-SNR regime*, one is able to focus on the effect of interference in the system, by rendering the impact of noise virtually negligible as compared to that of interference. In other words, to fix the ideas, in this regime, the performance is ruled by the signal-to-interference ratio, rather than the SNR. Instead, the

low-SNR regime provides a complementary viewpoint, by concentrating on an operating regime where performance is mostly determined by the SNR, and thus by the amount of (signal) energy available, rather than by the signal-to-interference ratio. So, broadly speaking, the high-SNR analysis provides insights into how well interference is handled by the system, and the low-SNR analysis brings information about how efficiently energy is used by the system. We now recall some relevant definitions and apply them to the analysis of the performance of SCP.

High-SNR Analysis In the high-SNR regime, the main performance criterion of interest is the *per-cell multiplexing gain* $\mathcal{S}_\infty$. For a given transmission scheme achieving rate $R(P)$ (bits/s/Hz), this quantity is defined as

$$\mathcal{S}_\infty = \lim_{P\to\infty} \frac{R(P)}{\log_2 P}, \tag{3.3}$$

which is measured in (bits/s/Hz/(3 dB)). The per-cell multiplexing gain provides the slope of the achievable rate $R(P)$ versus the SNR measure P on a plot where P is shown in 3 dB units, that is, as $10\log_{10}(P)/(10\log_{10}2)$. Based on the per-cell multiplexing gain, a system is said to be *interference-limited* if the multiplexing gain is zero and *noninterference-limited* otherwise. Given the interpretation of the multiplexing gain, this implies that a system is interference-limited if its plot of the rate $R(P)$ versus the SNR P attains a rate floor for large P.

Applying definition (Equation (3.3)) to the performance (Equation (3.1)) of SCP, we see that, if the reuse factor F is smaller than or equal to the intercell interference span L, the system operates in the interference-limited regime [33, 132]. Instead, if the reuse factor F is larger than the inter-cell interference span L, SCP with spatial reuse completely eliminates inter-cell interference and provides a noninterference-limited behavior with per-cell multiplexing gain equal to $\mathcal{S}_\infty = 1/F$. Since the multiplexing gain of an interference-free system (i.e., with all intercell power gains equal to zero) is unity, from the discussion above, we conclude that the presence of inter-cell interference,

if handled via SCP, leads to a rate degradation with respect to an interference-free system at high SNR given by a factor equal to $F > L$.[1]

Before moving to the low-SNR regime, we note that a more refined analysis of the high-SNR regime is possible following [69, 94]. In particular, rather than evaluating only the slope, one can also evaluate the offset of the curve $R(P)$ versus $P[\mathrm{dB}]/(3\,\mathrm{dB})$ discussed above. In other words, the rate $R(P)$ can be expanded as

$$R \simeq \mathcal{S}_\infty \left(\frac{P[\mathrm{d}B]}{3\mathrm{d}B} - \mathcal{L}_\infty \right), \tag{3.4}$$

where $\mathcal{L}_\infty$ is the so-called high-SNR offset. In this monograph, we will mostly concentrate on the multiplexing gain $\mathcal{S}_\infty$. For a thorough discussion of the high-SNR offset in the context of multicell systems, we refer the reader to [118].

Low-SNR Analysis In the low-SNR regime, where noise dominates interference, we can use the formalism of [125] and analyze the system performance in terms of the minimum (transmit) energy-per-bit necessary for reliable communications $E_b/N_0|_{\min}$ and of the so-called low-SNR (also known as "wideband") slope S_0 (bits/s/Hz/(3 dB)). Specifically, reference [125] proposes to expand an achievable rate $R(P)$ as a function of the energy-per-bit

$$E_b = \frac{P}{R(P)}, \tag{3.5}$$

as

$$R \simeq \frac{\mathcal{S}_0}{3\mathrm{d}B} \left(\frac{E_b}{N_0}[\mathrm{d}B] - \frac{E_b}{N_{0\,\min}}[\mathrm{d}B] \right), \tag{3.6}$$

where N_0 is the noise spectral density (normalized to 1 here). Moreover, reference [125] showed that this quantities can be calculated as $\frac{E_b}{N_{0\,\min}} = \frac{1}{\dot{R}(0)}$ and $\mathcal{S}_0 = (2\ln 2)\frac{(\dot{R}(0))^2}{(-\ddot{R}(0))}$, where $\dot{R}(P)$ and $\ddot{R}(P)$ represent the first and second derivative, respectively, of the rate $R(P)$.

[1] This discussion should be tempered by the observation that the high-SNR regime should always be defined within the limits of approximation of the Wyner model, as discussed in the previous section. The results obtained from the high-SNR analysis should thus be considered to be valid only within the SNR range that makes the Wyner model approximation acceptable.

Using the expression for the SCP rate (Equation (3.1)), it can be seen that intercell interference does not cause any increase in the minimum (transmit) energy-per-bit necessary for reliable communications $E_b/N_{0\min}$, which equals $\ln 2 = -1.59\,\text{dB}$, as for interference-free channels. However, if one observes also the slope of the spectral efficiency S_0 (bits/s/Hz/(3 dB)), which accounts for a higher-order expansion of the spectral efficiency as the SNR $P \to 0$, the loss due to intercell interference is seen also in the low-SNR regime. In fact, we have the following result for the rate (Equation (3.1)):

$$\mathcal{S}_0 = \begin{cases} \dfrac{2}{F\left(1 + 4\sum_{k=1}^{\lfloor L/F \rfloor} \alpha_{kF}^2\right)} & \text{if } F \leq L \\ \dfrac{2}{F} & \text{if } F > L \end{cases}, \tag{3.7}$$

where we recall that interference-free channels have $\mathcal{S}_0 = 2$.

As shown below, MCP allows the network to overcome the limitations of SCP and spatial reuse.

3.1.2 MCP with Unlimited Backhaul

We now consider the performance of MCP assuming unlimited-capacity backhaul links to a CP, i.e., $C \to \infty$. The per-cell sum-capacity $R_{MCP}(P)$ with MCP in this scenario (for any M) is given by [132]

$$R_{MCP}(P) = \frac{1}{M} \log_2 \det\left(\mathbf{I} + \frac{P}{K}\mathbf{H}\mathbf{H}^\dagger\right) \tag{3.8}$$

$$= \frac{1}{M} \sum_{m=1}^{M} \log_2\left(1 + \frac{P}{K}\lambda_m(\mathbf{H}\mathbf{H}^\dagger)\right) \tag{3.9}$$

$$= \int_0^\infty \log_2\left(1 + \frac{P}{K}x\right) dF_{\mathbf{H}\mathbf{H}^\dagger}(x), \tag{3.10}$$

where the $\lambda_m(\mathbf{H}\mathbf{H}^\dagger)$ are the eigenvalues of the argument matrix and $F_{\mathbf{H}\mathbf{H}^\dagger}(x)$ is the empirical distribution of such eigenvalues:

$$F_{\mathbf{H}\mathbf{H}^\dagger}(x) = \frac{1}{M} \sum_{m=1}^{M} 1(\lambda_m(\mathbf{H}\mathbf{H}^\dagger) \leq x). \tag{3.11}$$

The per-cell capacity (Equation (3.8)) is achieved by performing ideal multiuser detection at the CP (which can in practice be realized by following approaches such as in [1]). Moreover, transmission across all cells takes place with a reuse factor of $F = 1$, so that all cells are active at all times. As far as scheduling within each cell is concerned, both intra-cell TDMA, where users transmit with power P for a fraction of time $1/K$, and the WB scheme, whereby all users transmit with full power P/K at all times, lead to the same sum-rate.

It is noted again that the optimality of TDMA is strictly dependent on the per-block power constraint (Equation (2.12)), and would not hold in general under more restrictive conditions, such as per-symbol power constraints. More general conditions under which TDMA is optimal, under per-block power constraints, can be found in [33]. For instance, from [33], it is found that TDMA would generally not be optimal in scenarios where users had different intra- and intercell channel gains, such as in fading scenarios (see Section 3).

For the Gaussian Wyner model, assuming intra-cell TDMA, the matrix $\mathbf{H}$ is easily seen to be a symmetric Toeplitz matrix with first column given by $[\alpha_1 \alpha_2 \cdots \alpha_L \mathbf{0}_{M-L}]$. This implies that, using Szegö's theorem [30], for $M \to \infty$, rate (3.10) can be written in a simple integral form as [10, 132]

$$R_{MCP}(P) = \int_0^1 \log_2 \left(1 + P \left(1 + 2 \sum_{k=1}^{L} \alpha_k \cos(2\pi k \theta) \right)^2 \right) d\theta. \quad (3.12)$$

The following interpretation of expression (Equation (3.12)) is useful. Consider the case $K = 1$. This is done without loss of generality, given the optimality of intra-cell TDMA. We can now identify the signal received at the CP as the output, for each time instant, of an Linear Time Invariant (LTI) filter, whose input is given by the signals transmitted by the MSs and whose impulse response is $h_m = \delta_m + \sum_{k=1}^{L} \alpha_k \delta_{m-k} + \sum_{k=1}^{L} \alpha_k \delta_{m+k}$ (δ_m is the Kronecker delta function). This is illustrated in Figure 3.1 and can be explained as follows.

The maximum rate for communications over an LTI channel with frequency response $H(\theta) = \sum_m h_m e^{-j2\pi m\theta}$ is well-known to be given

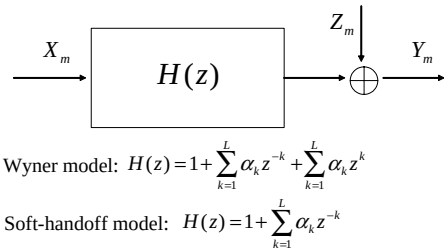

Fig. 3.1 LTI interpretation of the signal received over a Gaussian Wyner-type models.

by the maximum over $S(\theta)$ of the rate

$$R(S(\theta), H(\theta)) = \int_0^1 \log_2\left(1 + S(\theta)|H(\theta)|^2\right) d\theta, \tag{3.13}$$

where $S(\theta)$ is the *power spectral density* of the input signal (see, e.g., [90]). In the Wyner models, the spectrum $S(\theta)$ models the spatial correlation of the transmitted signals across different MSs, since the input signal X_m corresponds to the signal transmitted by the m-th MS. Given that the MSs cannot cooperate in the scenario at hand, their signals are necessarily independent. This is because the MSs cannot coordinate their transmissions since they have no common information. Therefore, we need to choose $S(\theta)$ as a constant in Equation (3.13), which leads to Equation (3.12). As a side remark, as will be discussed in Section 8, in case MSs can cooperate, then the input spectrum $S(\theta)$ need not be constant, leading to potential performance gains.

For future reference, we define

$$R_{\mathrm{w}}(a, b, \rho) = \int_0^1 \log_2\left(1 + \rho H(\theta)^2\right) d\theta \ , \tag{3.14}$$

where the channel frequency response is $H(\theta) = b + 2a\cos(2\pi\theta)$, as the maximum rate over a Gaussian Wyner model with direct (same-cell) gain b ($b = 1$ in Figure 2.2), intercell interference gain a ($a = \alpha$ in Figure 2.2) and power constraint ρ ($\rho = P$ in Figure 2.2). We will use this definition in Section 8.

For the Gaussian soft-handoff model with $L = 1$, the LTI filter modelling the effect of intercell interference can be seen to be $h_m = \delta_m + \alpha\delta_{m-1}$ and the integral (Equation (3.10)) can be evaluated in

closed form. Specifically, for the model at hand, we obtain [43, 118]

$$R_{MCP}(P) = \log_2\left(\frac{1+(1+\alpha^2)P+\sqrt{1+2(1+\alpha^2)P+(1-\alpha^2)^2P^2}}{2}\right). \tag{3.15}$$

3.1.2.1 Extreme-SNR Analysis

Following Section 3.1.1.1, we consider here the performance of MCP with unlimited-capacity backhaul in the extreme-SNR regimes. First, the multiplexing gain $\mathcal{S}_\infty$ of the MCP capacity (Equations (3.12)–(3.15)) equals one, as for an interference-free scenario. This implies that MCP is able to overcome the rate loss due to inter-cell interference of SCP identified above.

For the low-SNR regime, let us start with the Wyner model. The minimum energy-per-bit is given by

$$\frac{E_b}{N_0}_{\min} = \frac{\ln 2}{(1+2\sum_{k=1}^{L}\alpha_k^2)}, \tag{3.16}$$

showing an energy gain due to MCP with respect to SCP and to an interference-free system given by $(1+2\sum_{k=1}^{L}\alpha_k^2)$. This energy gain can be interpreted as the *array gain* due to the fact that the signal transmitted by each MS is received by the CP with a combined power of $(1+2\sum_{k=1}^{L}\alpha_k^2)P$ across all BSs. The parameter $\mathcal{S}_0$ for the special case $L=1$ is instead given by

$$\mathcal{S}_0 = \frac{2(1+2\alpha^2)^2}{1+12\alpha^2+6\alpha^4}. \tag{3.17}$$

Note that the second-order comparison between MCP and SCP provided by the low-SNR slope is here less relevant given the gain of MCP in terms of $E_b/N_0|_{\min}$.

For the soft-handoff model, we get

$$\frac{E_b}{N_0}_{\min} = \frac{\log_e 2}{1+\sum_{k=1}^{L}\alpha_k^2}, \tag{3.18}$$

showing the expected array gain of $(1 + \sum_{k=1}^{L} \alpha_k^2)P$ and the low-SNR slope

$$\mathcal{S}_0 = \frac{2(1+\alpha^2)^2}{1+4\alpha^2+\alpha^4}. \tag{3.19}$$

3.1.2.2 Numerical Example

We conclude this section with numerical evaluations of the rates discussed above for the uplink. Figure 3.2 shows the per-cell sum-rates achievable by MCP and SCP for the Gaussian Wyner model and the Gaussian soft-handoff model with $L = 1$ versus the inter-cell gain α^2 ($P = 1,\ K = 1$). SCP with reuse factor $F = 1$ has interference-limited performance and the achievable rate decreases for increasing interference power α^2. Since $L = 1$, a reuse factor $F = 2$ is instead able to fully mitigate the inter-cell interference, but the achievable rate improves on $F = 1$ only for sufficiently large α^2.

The performance gains achievable with MCP are remarkable if the inter-cell interference α^2 is large enough. This shows that with MCP,

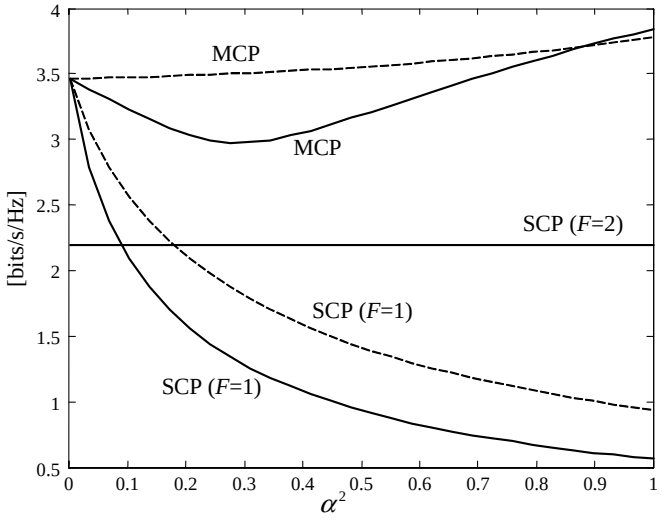

Fig. 3.2 Per-cell sum-rates for the Gaussian Wyner model (solid lines) and Gaussian soft-handoff models (dashed lines) versus the inter-cell gain α^2 ($L = 1,\ P = 10$).

interference should not be generally considered as a hindrance, since the CP can use the signal received across all BSs. Specifically, in the Wyner model, it can be seen that if α^2 is small enough, the effect of a larger inter-cell gain does decrease the achievable rate. In this regime, the joint decoding capabilities of the CP are not sufficient to overcome the effect of inter-cell interference. However, when α^2 is large enough, this is not the case and MCP provides remarkable gains. Note also that in the soft-handoff model, the achievable rate with MCP always increases with α^2.

3.2 Downlink

In this section, we briefly turn our attention to the downlink of a Gaussian Wyner model.

3.2.1 SCP and Spatial Reuse

Achievable rates using SCP and frequency reuse are easily seen to be given as for the uplink and we refer to Section 3.1.1 for discussion.

3.2.2 MCP with Unlimited Backhaul

As throughout this section, assume that MCP is enabled by unlimited-capacity backhaul to a CP. The per-cell sum-capacity is derived in [118] using the uplink–downlink duality results of [142] as (see Appendix B)

$$R_{MCP}(P) = \frac{1}{M} \min_{\mathbf{\Lambda}} \max_{\mathbf{D}} \log_2 \frac{\det\left(\mathbf{\Lambda} + \frac{P}{K}\mathbf{H}\mathbf{D}\mathbf{H}^{\dagger}\right)}{\det\left(\mathbf{\Lambda}\right)}, \tag{3.20}$$

with $\mathbf{\Lambda}$ and $\mathbf{D}$ being diagonal $MK \times MK$ matrices with the constraints $tr(\mathbf{\Lambda}) \leq M$ and $tr(\mathbf{D}) \leq M$. This rate can be interpreted as the per-cell sum-rate of a "dual" uplink system with channel gains defined as in the uplink of the system at hand, but where the m-th MS transmits with power $[\mathbf{D}]_{mm}$(i.e., the m-th diagonal element of matrix $\mathbf{D}$) and the noise covariance matrix across the BSs is Λ. In this formulation, the effects of the per-BS power constraints are reflected by the need to minimize over $\mathbf{\Lambda}$ (see Appendix B).

The per-cell sum-rate (Equation (3.20)) is known to be achieved by Dirty Paper Coding (DPC) [15] at the CP [118]. DPC is discussed

in more technical terms in Appendix A, but, in short, it operates as follows. The KM users in all cells are sorted in an optimized order, and the signal intended for each user is encoded in the selected order. When encoding for a given user, the CP "matches" the signal intended for the user at hand to the signals of the already encoded users, so that the effect of the latter on the user's reception is mitigated. This "matching" can be done in a number of ways (see, e.g., [24]), but the main idea is that the CP constructs a large codebook in which each message is associated with multiple possible code words. Among all the code words for the message to be encoded, the CP chooses the one that is "aligned" with the already encoded signals. This allows the CP to leverage, at least to some extent, rather than necessarily suffer from, interference from other signals.

For the Gaussian Wyner model, it can be shown that the per-cell sum-capacity (Equation (3.20)) is exactly equal to the corresponding capacity for the uplink (Equation (3.12)) for $M \to \infty$. From this fact, one can conclude that, as for the uplink, intra-cell TDMA, where only one user is served per-cell, is optimal with Gaussian (unfaded) channels.

3.3 Summary

In this section, we have discussed the performancee advantages of ideal MCP as implemented via unlimited capacity links connecting BSs and the CP. We have focused on Gaussian channels, leaving the study of the impact of fading to the next section. As compared to SCP and spatial reuse, MCP was found to offer significant gains in both the high-SNR regime, in which interference is the major performance bottleneck, and in the low-SNR regime, in which the performance depends mostly on an efficient use of the available power. In fact, in the high-SNR regime, thanks to its enhanced interference mitigation capabilities, MCP was shown to offer multiplexing gains, thus increasing the number of users that can be served at full rate by the system. In fact, with SCP and spatial reuse, the multiplexing gain is limited by the number of interfering cells; it becomes zero if the BSs do not have the capability of decoding also the messages transmitted by interfering cells and spatial reuse is not allowed. Moreover, in the low-SNR regime, MCP was shown

to provide beamforming gains, being able to coherently combine the signals received by different BSs. Finally, by using the concept of duality, it was discussed that, under the given ideal backhaul assumptions, MCP has the same performance, in terms of per-cell sum-rate, in both the uplink and the downlink.

4

Multi-Cell Processing in Fading Channels

In the previous section, we elaborated on the potential gains attainable via MCP in a scenario with no fading and ideal backhaul links. In this section, we extend the analysis from Gaussian to fading models. We consider first the uplink and then the downlink, and deal mostly with ergodic fading. In order to focus on the impact of fading, we will assume throughout this section unlimited-capacity backhaul connections to the CP, as in the previous section. The effect of limitations on the backhaul links will be discussed in the following section.

4.1 Uplink

Here we study the performance of MCP for the uplink of fading Wyner and soft-handoff models (see Section 2.1), assuming unlimited-capacity backhaul links to the CP. We focus on ergodic fading, and leave some discussion on quasistatic fading to Section 4.1.3.

We start by observing that, with ergodic fading, the per-cell sum-capacity is given by the expectation of Equation (3.10) with respect to the distribution of $\mathbf{H}_M$, i.e.,

$$R_{MCP}^{erg}(P) = \mathbb{E}\left(R_{MCP}(P)\right). \tag{4.1}$$

This conclusion follows from standard information-theoretic considerations (see, e.g., [22]). It accounts for the fact that, if the fading is ergodic, encoding over a sufficiently long period enables averaging of the effect of fading. Note that we have added a subscript to the channel matrix $\mathbf{H}_M$ defined in the previous section in order to stress its dependence on M.

Following arguments similar to the ones put forth in the previous section, this rate can be seen to be achieved by the WB scheme, so that all users transmit with full power at all times, along with joint decoding at the CP. However, unlike the Gaussian (unfaded) models discussed in the previous section, the per-cell sum-rate achieved by intra-cell TDMA is generally smaller. This is in line with standard results on multiple access channels (see, e.g., [17, 95]) and can be interpreted as follows. With fading and intra-cell TDMA, if the currently scheduled user has an unfavorable channel state, then the current transmission time is, in some sense, wasted. In contrast, if all users are concurrently scheduled as with WB, then the unfavorable channel conditions of some users are potentially compensated for by the good channel conditions of other users, thus leading to possible performance gains. This can be seen as an instance of the gains attainable via *multiuser diversity*. Evaluating Equation (4.1) analytically turns out to be a generally complex problem, and it will be studied below for both the Wyner and the soft-handoff models.

Before discussing the evaluation of Equation (4.1) further, we note that the performance of SCP with spatial reuse can be obtained by taking the expectation of the rates achievable in the corresponding Gaussian models with respect to the channel gains (these rates can be calculated by adapting Equation (3.1)). By doing so, it can be seen that, with SCP, intra-cell TDMA may be advantageous over WB for $F \leq L$ if the intercell interference is large enough [95]. This is because intra-cell TDMA reduces the overall power of the intercell interference. Some further performance comparisons between intra-cell TDMA and spatial reuse in the presence of MCP for the soft-handoff model with $L = 1$ and Rayleigh fading can be found in [63].

4.1.1 Wyner Model

Evaluating Equation (3.1) analytically can be potentially attempted as the number of cells M, and thus the $M \times MK$ matrix $\mathbf{H}_M$, grows large using (large) random matrix analytical tools. This is an attractive option since related results are often proved to be valid also for relatively small values of M (see, e.g., [113] and discussion below). Many recent studies have indeed successfully analyzed the asymptotic rates of various vector channels using results from the theory of random matrices. In Appendix D, we provide an introduction to concepts of random matrix analysis that are useful for the discussion.

In most cases where random matrix analysis bears fruit, the number of random variables present in the random matrix involved grows quadratically in the size of the matrix $\mathbf{H}_M$, so that "self-averaging" is strong enough to ensure convergence of the empirical measure Equation (3.11) of the eigenvalues as M grows large, and to derive equations for the corresponding limiting spectrum (or its Stieltjes transform; see Appendix D). In particular, this is possible if the normalized continuous power profile of $\mathbf{H}_M$, which is defined with $r,\ t \in [0,1]$ as

$$\mathcal{P}_M(r,t) \triangleq \mathbb{E}(|[\mathbf{H}_M]_{i,j}|^2) \quad \frac{i-1}{M} \leq r < \frac{i}{M}, \quad \frac{j-1}{MK} \leq t < \frac{j}{MK}, \quad (4.2)$$

converges uniformly to a bounded, piecewise continuous function as $M \to \infty$, see e.g., [121] (see also [2] for fluctuation results).

In the fading Wyner model (and also in the soft-handoff model to be discussed in Section 4.1.2), the matrix $\mathbf{H}_M\mathbf{H}_M^\dagger$, which appears in the rate (Equation (3.1)) we are interested in evaluating, is a finite-band matrix [116]. Therefore, the number of random variables grows linearly in the size of the matrix. For this class of matrices, it is easy to verify that for K fixed, $\mathcal{P}_M(r,t)$ does *not* converge uniformly, so that standard techniques are not applicable.

While standard random matrix tools are not useful for our purposes, reference [62] proved that under light assumptions on the fading coefficients, $R_{MCP}(P)$ converges almost surely as M goes to infinity, and the limit is expressed as a function of the *Lyapunov exponent* of sequence of fixed-size random matrices. This result is demonstrated

by using a version of the *Thouless formula for the strip* (see [62] for further details). It is also shown that $R_{MCP}^{erg}(P)$ converges to the same quantity as well, thereby proving that the ergodic sum-rate converges as the number of cells goes to infinity. The work [62] provides also a useful relationship between the value of $R_{MCP}^{erg}(P)$ for $M \to \infty$, say $R_{MCP}^{erg,\infty}(P)$, and the corresponding per-cell ergodic rates for finite M, say $R_{MCP}^{erg,M}(P)$. In particular, it is proved that

$$\frac{M}{M+L_\ell+L_r}\mathbb{E}\left(\tilde{R}_{MCP}^{M}\left(\frac{M+L_\ell+L_r}{M}P\right)\right) \leq R_{MCP}^{erg,\infty}(P) \leq \mathbb{E}(\tilde{R}_{MCP}^{M}(P)),$$

where $\tilde{R}_{MCP}^{M}(P)$ is the per-cell sum-rate of a slightly different model where one adds L_ℓ cells without MSs on the left and L_r on the right at the boundaries of the system.

While [62] proves convergence of the ergodic rate as M grows large, this result does not provide per se much insight into the performance of MCP. To this end, we will instead use different bounding or approximation techniques below, and we will study performance in asymptotic SNR regimes.

4.1.1.1 Wide-Band Transmission (Large Number K of Users Per Cell)

Consider the case in which the number of users per cell K is large while the total power per cell P is kept constant. Recall that in the fading Wyner model the channel gains of all MSs have the same statistics. Applying the strong law of large numbers, the entries of $\frac{1}{K}\mathbf{H}_M\mathbf{H}_M^\dagger$ thus consolidate almost surely (a.s.) to their mean values, so that matrix $\frac{1}{K}\mathbf{H}_M\mathbf{H}_M^\dagger$ becomes a Toeplitz matrix (with deterministic entries) [111]. Then, as done in the previous section, by applying Szegö's theorem [30], for $M \to \infty$, the per-cell sum-rate capacity for the linear fading Wyner model is given by $R_{MCP}^{\infty}(P) = R_{MCP}^{erg,\infty}(P)$, where

$$R_{MCP}^{\infty}(P) = \int_0^1 \log_2 \left(\begin{array}{l} 1 + P\sigma^2\left(1 + 2\sum_{k=1}^{L}\alpha_k^2\right) \\ \quad + P(1-\sigma^2)\left(1 + 2\sum_{k=1}^{L}\alpha_k \cos(2\pi k\theta)\right)^2 \end{array} \right) d\theta, \tag{4.3}$$

in which σ^2 is the variance of each fading coefficient. We remark that the rate (Equation (4.3)) does not depend on the actual (stationary and ergodic) distribution of an individual fading distribution but only on its first two moments. Note also that this result recovers Equation (3.12) for $\sigma^2 = 0$, which corresponds to an unfaded scenario.

For $\sigma^2 = 1$, as assumed throughout this monograph when considering fading, we get

$$R_{MCP}^{\infty}(P) = \log_2 \left(1 + P \left(1 + 2 \sum_{k=1}^{L} \alpha_k^2 \right) \right). \tag{4.4}$$

This latter simple expression shows that MCP, when the number of users is large, is able not only to fully mitigate the intercell interference, as can be seen from the unitary multiplexing gain, but also to leverage the beamforming gain of $1 + 2\sum_{k=1}^{L} \alpha_k^2$ afforded by the reception of multiple BSs.

A few more remarks are in order. It can be proved, similar to [111], that Equation (4.3) is increasing in σ^2 and is maximized for $\sigma^2 = 1$ as for Rayleigh fading. This implies that fading is beneficial with respect to a Gaussian channel with gains equal to the average of the fading coefficients, in the limit when the number of users per-cell increases without bound. It is remarked that the performance advantages due to fading may not hold for a finite number of users K. This can be easily seen by noticing that for $K = 1$ and no inter-cell interference ($\alpha = 0$), one obtains a point-to-point link for which fading is known not to increase the rate [111, 118]. The advantages of fading for a sufficiently large number of MSs can be interpreted, as argued above, in terms of *multiuser diversity*. In fact, if the number of users is large enough, it is likely that a sufficient number of users will experience better than average channel conditions, thus compensating for the bad channel conditions of other users.

Applying Jensen's inequality to the rate expression (Equation (4.1)), it can be seen that Equation (4.4) is also an upper bound on the ergodic per-cell capacity for any number of users K [111].

Figures 4.1 and 4.2, plotted for the linear Wyner model with $L = 1$ and Rayleigh fading, demonstrate the tightness of the asymptotic expression (and upper bound) (Equation (4.4)) already for a moderate

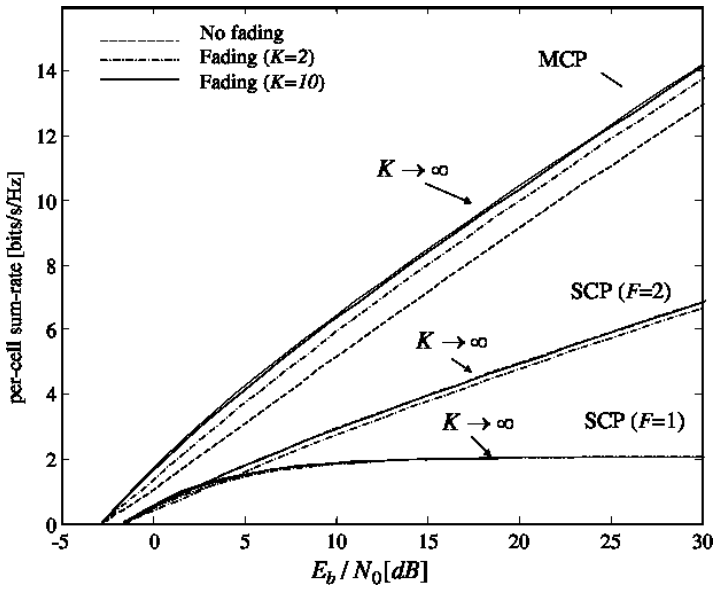

Fig. 4.1 Per-cell sum-rates versus E_b/N_0 for the uplink of the fading Wyner model (Rayleigh fading, $L = 1$, $\alpha = 0.4$).

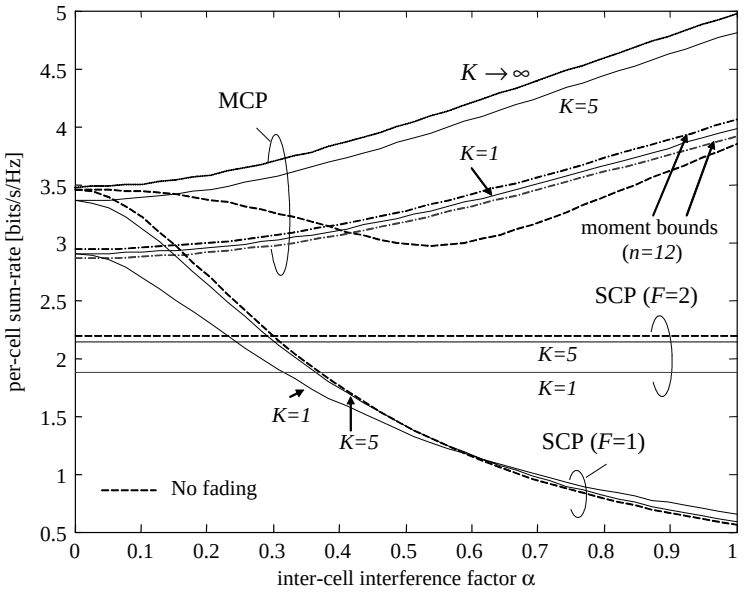

Fig. 4.2 Per-cell sum-rates versus inter-cell interference factor α in the fading Wyner model (Rayleigh fading, $P = 10\,\mathrm{dB}$).

number of users K. The fact that the presence of fading is beneficial with large K for all values of P and α is also demonstrated in the figures.

4.1.1.2 Moment Bounds for Intra-Cell TDMA

An alternative approach to approximating the ergodic per-cell sum-rate $R_{\text{MCP}}^{erg,\infty}$ is proposed in [111] for a fading Wyner model with $L = 1$ and fading coefficient with uniformly distributed phase (e.g., Rayleigh fading). Focusing on the intra-cell TDMA protocol, or equivalently $K = 1$, bounds on $R_{\text{MCP}}^{erg,\infty}$ were found that depend on the moments of the eigenvalue distribution (Equation (3.11)). Specifically, it was first shown that the average unordered eigenvalue distribution $\mathbb{E}(F_{\mathbf{H}_M\mathbf{H}_M^\dagger})$ converges weakly to a unique distribution $\mathcal{F}$ as the number of cells increases without bound, $M \to \infty$.[1] In addition, using a standard weighted paths summation over a restricted grid, the limiting values of the first several moments of this distribution were calculated. For example, listed below are the first two limiting moments:

$$\mathcal{M}_1 = m_2 + 2m_2\alpha^2 \tag{4.5}$$

and

$$\mathcal{M}_2 = m_4 + 8m_2^2\alpha^2 + (4m_2^2 + 2m_4)\alpha^4, \tag{4.6}$$

where m_i is the i-th power moment of the amplitude of an individual fading coefficient. Following [49] the upper and lower principle representations of $\mathcal{F}$ can be used to produce analytical lower and upper bounds on the per-cell sum-rate capacity. For instance, listed below are the lower and upper bounds of order $n = 2$ derived by this method are as follows:

$$\frac{(\mathcal{M}_1)^2}{\mathcal{M}_2}\log\left(1 + P\frac{\mathcal{M}_2}{\mathcal{M}_1}\right) \leq R_{\text{MCP}}^{erg,\infty} \leq \log(1 + P\mathcal{M}_1), \tag{4.7}$$

where $\mathcal{M}_1$ and $\mathcal{M}_2$ are the first and second limit moments of $\mathcal{F}$ given above. It is noted that this procedure can be extended in principle,

[1] It is conjectured that using similar methods the spectrum can be proved to converge a.s. to a unique limit as well.

although in a tedious manner, for any finite K, or also for Wyner like models where inter-cell interference stems from farther cells.

Examining the moment bounds (orders $n = 8,\ 10$) calculated for Rayleigh fading and presented in Figure 4.2, reveals that the bounds get tighter as the order n increases. Additional calculations [111] (not presented here) show that the bounds are tighter when P decreases. In addition, these calculations reveal that the rate of MCP with Rayleigh fading can be larger than for the corresponding Gaussian channel even for $K = 1$ but only for certain values of α and P. Following up the discussion above, this can be explained by the fact that with MCP inter-cell interference may not be deleterious, as explained in the discussion around Figure 3.2, and that fading allows the network to potentially leverage the good channel gains from any MS to any other BS it affects.

4.1.1.3 Free-Probability Approximation

In [56], the limiting spectrum of $\frac{1}{1+2\alpha^2}\mathbf{H}_M\mathbf{H}_M^\dagger$ for the fading Wyner model with $L_r = L_\ell = 1$ has been loosely shown by *free probability* tools to be approximated by the *Marŏenko–Pastur distribution* [121] (see Appendix D) with parameter K (number of users per-cell). The approximation is shown to match fairly well Monte-Carlo simulations for relatively large values of α and for Rayleigh fading. To demonstrate the approximation inaccuracy in the low α regime, we note that in the extreme case of $\alpha = 0$, the eigenvalues are evidently exponentially distributed, with no finite support, in contrast to the Marŏenko–Pastur distribution.

4.1.1.4 Extreme-SNR Analysis

Here we adopt the extreme-SNR analysis presented in Section 3.1.1.1 in order to get further insight into the performance of MCP over fading channels.

Recall from Section 3.1.1.1, the low-SNR regime is characterized through the minimum transmit E_b/N_0 that enables reliable communications, i.e., $E_b/N_{0_{\min}}$, and the low-SNR spectral efficiency slope $\mathcal{S}_0$. Using the results in [125] as in Section 3.1.1.1, the low-SNR parameters for the per-cell sum-rate capacity of the fading Wyner model with

$L = 1$ are given for Rayleigh fading by[2] [117]

$$\frac{E_b}{N_0}_{\min} = \frac{\ln 2}{1 + 2\alpha^2} \quad \text{and} \quad \mathcal{S}_0 = \frac{2K}{1 + K}. \tag{4.8}$$

Comparing these results with the corresponding parameters in the Gaussian (unfaded) Wyner model, namely Equations (3.16) and (3.17), reveals that fading does not entail any loss or gain in performance in terms of $\frac{E_b}{N_0}_{\min}$. In particular, the beamforming gain of $1 + 2\alpha^2$ is still achievable. Instead, the second-order analysis afforded by considerations of S_0 shows that, with low SNR, fading is beneficial when the number of users per cell exceeds a certain threshold which is a decreasing function of α [113].

Turning to the high-SNR regime, it is proved in [62] that the multiplexing gain is unity, just as for the corresponding Gaussian model (an expression for the so-called power offset is also given as a function of a Lyapunov exponent).

4.1.2 Soft-Handoff Model

We now discuss evaluation of the MCP per-cell sum-rate (Equation (4.1)) for the soft-handoff model with $L = 1$. Unlike the Wyner model, here the matrix $\mathbf{H}_M \mathbf{H}_M^\dagger$ has a Jacobi structure, that is, it is tridiagonal. This fact, as discussed in [118] facilitates analytical treatment beyond what is known to be possible for the Wyner model. This was already seen for the Gaussian scenario (see Equation (3.15)). In the following, we discuss some of the available results for fading models. We remark that, despite the simplified model, finding a closed-form analytical expression for the ergodic rate achieved by MCP for general fading statistics and finite number K of users per-cell remains an open problem.

4.1.2.1 Intra-Cell TDMA

Here we focus on the per-cell sum-rates achieved by intra-cell TDMA, or equivalently when $K = 1$, and Rayleigh fading with $\alpha = 1$. In this

[2] This result can be extended in a straightforward yet tedious manner to Wyner-like models with arbitrary L.

scenario, extending the discussion around Figure 3.1, it can be seen that the achievable per-cell sum-rate equals the capacity of an LTI system where the impulse response has two taps, corresponding to the fading coefficients toward a given BS. This scenario was studied by Narula in [78]. As also discussed in Appendix D, Narula's main observation was that the diagonal entries $\{d_m\}$ of the *Cholesky* decomposition applied to $\mathbf{G}_M = \mathbf{I}_M + P\mathbf{H}_M\mathbf{H}_M^\dagger$ form a discrete-time continuous space Markov chain with a unique ergodic stationary distribution as $M \to \infty$ given by

$$f_d(x) = \frac{\log(x)e^{-\frac{x}{P}}}{\mathrm{Ei}\left(\frac{1}{P}\right)P}, \quad x \geq 1, \tag{4.9}$$

where $\mathrm{Ei}(x) = \int_x^\infty \frac{\exp(-t)}{t}dt$ is the exponential integral function. Further, as proved in [78], the strong law of large numbers holds for the sequence $\{\log d_m\}$ as $M \to \infty$. Hence, the average per-cell sum-rate capacity of the intra-cell TDMA scheme ($K = 1$) can be proved to be given by

$$R_{\mathrm{MCP}}^{erg,\infty}(P) = \int_1^\infty \frac{(\log(x))^2 e^{-\frac{x}{P}}}{\mathrm{Ei}\left(\frac{1}{P}\right)P}dx. \tag{4.10}$$

Note that Narula's approach is based on an explicit calculation of the stationary distribution f_d, and is thus limited to Rayleigh fading. It has not be thus far possible to extend this result to other fading distributions or even to different values of the inter-cell interference factor α. An exception to this is the case with uniform phase fading, where the amplitude of all fading gains is (deterministically) equal to one and the phases are uniformly distributed. In this simpler scenario, considered in [43], the per-cell sum-rate capacity is shown to coincide with that of the corresponding Gaussian model, namely Equation (3.15), as $M \to \infty$.

4.1.2.2 Wide-Band Transmission

Following [78], it is also possible to obtain a nontrivial upper bound on the per-cell sum-rate of the WB scheme with finite K and infinite number of cells $M \to \infty$, in the presence of zero-mean unit power ($m_1 = 0$, $m_2 = 1$) fading (e.g., Rayleigh fading) and $\alpha = 1$. This bound is given

by

$$R_{\mathrm{MCP}}^{erg,\infty}(P) \leq \log\left(\frac{1+2P+\sqrt{(1+2P)^2-(4P^2/K)}}{2}\right). \qquad (4.11)$$

It is proved in [118] for $K=1$ (intra-cell TDMA protocol) and extended to an arbitrary K in [63]. The derivation is carried out by using the fact that the determinant of a Jacobi matrix is equal to a weighted sum of the determinants of its two largest principal sub-matrices.

It is noted that for $K=1$, the upper bound (Equation (4.11)) coincides with the per-cell sum-rate capacity (Equation (3.15)) of the Gaussian soft-handoff model. It follows that the presence of Rayleigh fading cannot increase the rates of the intra-cell TDMA protocol in the soft-handoff model with $\alpha=1$. This is in contrast to the Wyner model, as explained above. Nevertheless, it is shown in [118] that already for $K=2$ users the presence of fading may be beneficial at least for low SNR values. Note that the bound is tight for $K\to\infty$ since in this regime the strong law of large numbers can be invoked to show that the per-cell sum-rate is given by Equation (3.15).

In Figure 4.3 the per-cell sum-rate bound (Equation (4.11)) is shown for $K=1,2,5,\infty$, along with the exact rate expression for intra-cell TDMA (4.10), and a Monte-Carlo simulation for $K=2$ and $M=40$, as a function of the transmitted E_b/N_0.[3] It is noted that the bound for $K=1$ coincides with the corresponding nonfading per-cell sum-rate (Equation (3.15)). The tightness of the bounds with increasing numbers of users per-cell is apparent already for $K=2$. In addition, the fact that Rayleigh fading is beneficial for $K>1$ is also visible.

4.1.3 Quasistatic Fading

With quasistatic fading, the *outage capacity* is typically used as a performance measure [82]. This is, generally speaking, the maximum rate that guarantees reliable transmission for a given percentage of channel realizations (the complement of whose measure is referred to as outage

[3] Recall that $RE_b/N_0 = P$.

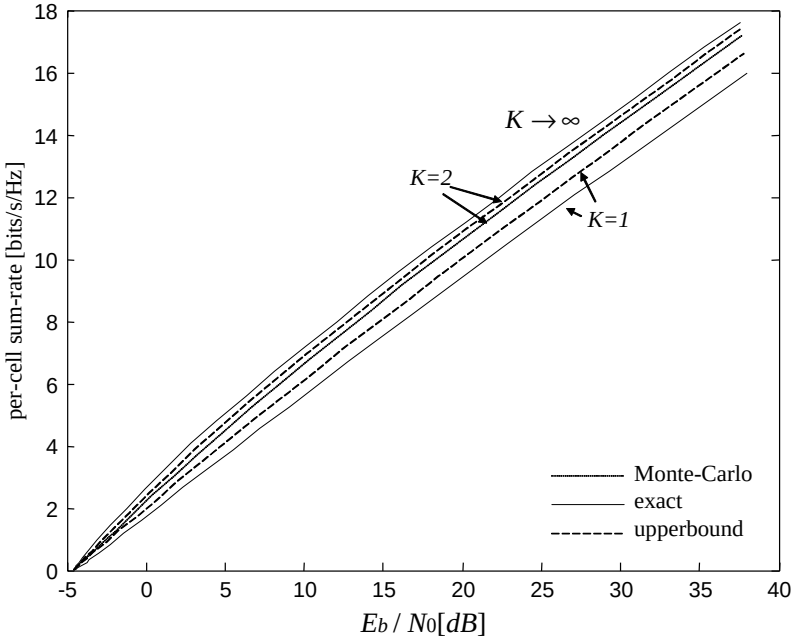

Fig. 4.3 Per-cell sum-rates versus E_b/N_0 for the uplink of the fading soft-handoff model (Rayleigh fading, $L = 1$, $\alpha = 1$).

probability). This setting implies either lack of CSI at the users (so that rate adaptation is not possible) or inelastic constant-rate applications. Using such a performance index in a large-scale cellular system with MCP proves to be challenging. In fact, on the one hand, defining outage as the event where *any* of the users' messages are not correctly decoded leads to uninteresting results as the number of cells M grows large; on the other hand, defining individual outage events, as studied in [77] for a two-user multiple-access channel, appears to be analytically intractable for large systems.

A tractable performance measure is instead obtained by considering the achievable per-cell sum-rate (Equation (3.8)) for given channel realizations in the limit as the number of users per cell K and/or the number of cells M grow large, where the limit is defined in an almost sure sense. It is noted that such per-cell sum-rate is achievable by appropriate choice of distinct rates by the MSs, and such choice depends on the

current realization of the channel matrices. The practical significance of this performance measure is thus limited to instances in which, thanks to appropriate signalling, such rate adaptation is possible.

The relations between rate expressions supported by quasistatic and ergodic channels in Wyner models is studied in [61]. As mentioned earlier, the rate $R_{MCP}(P)$ is shown to converge almost surely to $R_{MCP}^{erg,\infty}(P)$ as $M \to \infty$. Therefore, in the limit of $M \to \infty$, the outage probability equals one if the transmitted per-cell sum-rate is larger than $R_{MCP}^{erg,\infty}(P)$ and zero otherwise. Moreover, in [61], a central limit theorem is shown for $R_{MCP}(P)$ that characterizes the outage probability for large values of M. In particular, it is proved that the fluctuations of $R_{MCP}(P)$ decrease like $1/\sqrt{M}$ as M goes to infinity.

4.1.4 DS-CDMA

The previous sections focused mostly on the optimal performance in a cellular uplink with MCP, and showed that WB transmission with joint decoding is optimal over fading channels. Here, instead, we consider the performance of a practical and widely adopted transmission scheme, namely Direct Sequence-Code Division Multiple Access (DS-CDMA). We focus on the fading Wyner model and assume that the users employ binary spreading sequences with a *chip-level* interleaver and spreading factor, or processing gain, equal to N. Note that, in practical DS-CDMA systems, symbol-level interleaving may be used instead, but the current assumption simplifies the analysis, making it possible to employ random matrix tools (see Appendix D). Further discussion on this assumption can be found in [117].

We analyze the *asymptotic* regime in which both the number of users per cell K and the processing gain N go to infinity, while their ratio goes to some finite constant, which is denoted by $\beta = \lim_{K,N\to\infty} K/N$ and is referred to as the "cell-load". Moreover, in order to impose further practical constraints on the design, it is assumed that joint decoding is possible only among BSs belonging to the same cluster of M cells. More precisely, the signals received at each of the BSs within a cluster are jointly processed by exploiting the knowledge of signatures

and codebooks of all intra-cluster users, while the codebooks of out-of-cluster users are *unknown* and their signature sequences are known. This is an example of cluster decoding, which will be further considered in the next section. The CP employs either the optimum joint decoder or a more practical linear minimum mean square error (MMSE) decoder. We follow the treatment in [117].

Since transmission involves N chips per symbol, the baseband signal received by the CP for the cluster of concern consists of MN chip-level samples, which we write as

$$\boldsymbol{y}_1^M = \boldsymbol{H}\boldsymbol{x}_0^{M+1} + \boldsymbol{n}_1^M, \tag{4.12}$$

where $\boldsymbol{y}_1^M = (\boldsymbol{y}_1^T, \boldsymbol{y}_2^T, \ldots, \boldsymbol{y}_M^T)^T$ is the $MN \times 1$ vector received by the cluster M BSs, $\boldsymbol{x}_0^{M+1} = (\boldsymbol{x}_0^T, \boldsymbol{x}_1^T, \ldots, \boldsymbol{x}_M^T, \boldsymbol{x}_{M+1}^T)^T$ is the $(M+2)K \times 1$ transmitted symbols vector by the MK cluster users and the two interfering cells, and $\boldsymbol{n}_1^M = (\boldsymbol{n}_1^T, \boldsymbol{n}_2^T \ldots, \boldsymbol{n}_M^T)^T$ is the $MN \times 1$ additive complex Gaussian noise vector received by the cluster M cells sites. Note that the BSs in the cluster are numbered as $1, \ldots, M$, while the interfering, out-of-cluster, cells are 0 and M. The $MN \times (M+2)K$ channel transfer matrix $\boldsymbol{H}$ is given by

$$\boldsymbol{H} = \begin{pmatrix} \alpha \boldsymbol{S}_0 \circ \boldsymbol{H}_{1,0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \tilde{\boldsymbol{H}} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \alpha \boldsymbol{S}_{M+1} \circ \boldsymbol{H}_{M,M+1} \end{pmatrix}, \tag{4.13}$$

where $\circ$ stands for the Hadamard product, defined for arbitrary equal sized matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ as

$$[\boldsymbol{A} \circ \boldsymbol{B}]_{i,j} \triangleq [A]_{i,j}[B]_{i,j}, \tag{4.14}$$

and $\tilde{\boldsymbol{H}}$ is an $MN \times MK$ matrix given by

$$\begin{pmatrix} \boldsymbol{S}_1 \circ \boldsymbol{H}_{1,1} & \alpha \boldsymbol{S}_2 \circ \boldsymbol{H}_{1,2} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \alpha \boldsymbol{S}_1 \circ \boldsymbol{H}_{2,1} & \boldsymbol{S}_2 \circ \boldsymbol{H}_{2,2} & \alpha \boldsymbol{S}_3 \circ \boldsymbol{H}_{2,3} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \alpha \boldsymbol{S}_2 \circ \boldsymbol{H}_{3,2} & \boldsymbol{S}_3 \circ \boldsymbol{H}_{3,3} & \alpha \boldsymbol{S}_4 \circ \boldsymbol{H}_{3,4} & \cdots \\ & \ddots & \ddots & \ddots & \\ \boldsymbol{0} & \cdots & \boldsymbol{0} & \alpha \boldsymbol{S}_{M-1} \circ \boldsymbol{H}_{M,M-1} & \boldsymbol{S}_M \circ \boldsymbol{H}_{M,M} \end{pmatrix}. \tag{4.15}$$

The entries of the $N \times K$ matrices $\{\boldsymbol{H}_{m,n}\}$ are the channel fades affecting the signals transmitted by users in the n-th cell at a chip-level, as observed by the m-th BS. With the underlying chip-level interleaver assumption, the entries of $\boldsymbol{H}_{m,n}$ matrices become independent and identically distributed (i.i.d.) zero-mean circularly symmetric complex Gaussian random variables, with unit variances (corresponding to Rayleigh fading). The matrices are also assumed to be statistically independent for different values of m and n. The $N \times K$ matrices $\{\boldsymbol{S}_m\}$ denote the binary signature matrices, with the columns of the matrix $\boldsymbol{S}_m$ being the spreading sequences of the users operating in the m-th cell. It is assumed that the entries of the signature matrices are binary i.i.d. random variables taking the values $\{-1/\sqrt{N}, 1/\sqrt{N}\}$ with equal probability. Independence of the spreading sequences of different users is also assumed. It is noted that the matrices $\alpha \boldsymbol{S}_0 \circ \boldsymbol{H}_{1,0}$ and $\alpha \boldsymbol{S}_{M+1} \circ \boldsymbol{H}_{M,M+1}$ account for the inter-cell interference received by the cluster cell sites 1 and M, respectively.

4.1.4.1 Optimum Receiver

The per-cell ergodic sum-rate of the *optimum receiver*, which performs joint decoding, is given by [117]

$$R_{M,opt}^{erg}(P) = \Re_M - \frac{2}{M}\Re_I, \tag{4.16}$$

where

$$\Re_M = \frac{1}{M} \lim_{\substack{N,K\to\infty \\ \frac{K}{N}\to\beta}} \frac{1}{N} E\{\log\det(\boldsymbol{I} + P\tilde{\boldsymbol{H}}\tilde{\boldsymbol{H}}^{\dagger})\}, \tag{4.17}$$

and

$$\Re_I = \lim_{\substack{N,K\to\infty \\ \frac{K}{N}\to\beta}} \frac{1}{N} E\{\log\det(\boldsymbol{I} + \alpha^2 P\boldsymbol{G}\boldsymbol{G}^{\dagger})\}, \tag{4.18}$$

with the $N \times K$ matrix $\boldsymbol{G}$ in Equation (4.18) being the channel transfer matrix corresponding to users of cell 0, whose signals are received at cell site 1, given by

$$\boldsymbol{G} = \boldsymbol{S}_0 \circ \boldsymbol{H}_{1,0}. \tag{4.19}$$

Rate $\mathcal{R}_M$ can be interpreted as the average per-cell spectral efficiency in the case in which the receiver also tries to decode the transmissions of users in the two cluster-adjacent cells (assuming that their codebooks are now also known at the receiver). With the underlying assumption of binary spreading sequences, and independent circularly symmetric Gaussian channel fades, the entries of the channel transfer matrix $\boldsymbol{H}$ are marginally Gaussian. Furthermore, they are independent and hence jointly Gaussian (and uncorrelated). With this observation, a practical version of Girko's law (see Appendix D) can be used, obtaining

$$\begin{aligned}\mathcal{R}_M = & \frac{1}{M}\sum_{m=0}^{M+1}\beta\log(1+\Gamma_M(m)(1+2\alpha^2)\beta P) \\ & + \frac{1}{M}\sum_{m=1}^{M}\log\left(1+\sum_{k=m-1}^{m+1}\frac{[\mathbb{P}_M]_{m,k}\beta P}{1+\Gamma_M(k)(1+2\alpha^2)\beta P}\right) \\ & - \frac{1}{M}\sum_{m=0}^{M+1}\beta\Gamma_M(m)\Upsilon_M(m)\log e, \end{aligned} \tag{4.20}$$

where the discrete-index asymptotic power profile $M \times (M+2)$ matrix $\mathbb{P}_M$ is given by

$$\mathbb{P}_M = \begin{bmatrix} \alpha^2 & 1 & \alpha^2 & 0 & \cdots & & 0 \\ 0 & \alpha^2 & & \alpha^2 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \alpha^2 & & \alpha^2 \end{bmatrix} \tag{4.21}$$

with the rows of $\mathbb{P}_M$ being enumerated as $l = \{1,\ldots,M\}$ and the columns enumerated as $m = \{0,\ldots,(M+1)\}$. Moreover, the values of the discrete-index function $\{\Gamma_M(m)\}_{m=0}^{M+1}$ are given by the unique solutions to the following set of equations:

$$\Gamma_M(m) = \frac{1}{\beta(1+2\alpha^2)}\sum_{l=m-1}^{m+1}\frac{[\mathbb{P}_M]_{l,m}}{1+\sum_{k=l-1}^{l+1}\frac{[\mathbb{P}_M]_{l,k}\beta P}{1+\Gamma_M(k)(1+2\alpha^2)\beta P}}, \tag{4.22}$$

for $m = 0, 1, \ldots, M+1$, where "out-of-range" indices should be ignored. Finally, we have

$$\Upsilon_M(m) = \frac{(1+2\alpha^2)\beta P}{1+\Gamma_M(m)(1+2\alpha^2)\beta P}, \quad m = 0, 1, \ldots, M+1, \tag{4.23}$$

for $m = 0, 1, \ldots, M+1$. The quantity $\mathcal{R}_I$ may be interpreted as the spectral efficiency of an optimum receiver in a single isolated cell with homogeneous fading [126], and with a scaled SNR of $\alpha^2 P$:

$$\begin{aligned} \mathcal{R}_I = \; & \beta \log\left(1 + \alpha^2 P - \frac{1}{4}\mathcal{F}(\alpha^2 P, \beta)\right) \\ & + \log\left(1 + \alpha^2 P\beta - \frac{1}{4}\mathcal{F}(\alpha^2 P, \beta)\right) \\ & - \frac{\log e}{4\alpha^2 P}\mathcal{F}(\alpha^2 P, \beta) \end{aligned} \tag{4.24}$$

where

$$\mathcal{F}(x, z) \triangleq \left(\sqrt{x(1+\sqrt{z})^2 + 1} - \sqrt{x(1-\sqrt{z})^2 + 1}\right)^2. \tag{4.25}$$

Low-SNR Analysis

Having derived the general expressions above, we can now obtain more insight into the performance of the optimum receiver with DS-CDMA by analyzing the low-SNR behavior. The low-SNR characterization of the optimum receiver is summarized as follows [117]:

$$\begin{aligned} \frac{E_b}{N_0}_{\min} &= \frac{\log 2}{\left[1 + 2\alpha^2\left(1 - \frac{1}{M}\right)\right]}, \\ \mathcal{S}_0 &= \frac{2\beta}{1+\beta}\frac{\left[1 + 2\alpha^2(1 - \frac{1}{M})\right]^2}{\left[1 + 4\alpha^2\left(1 - \frac{1}{(1+\beta)M}\right) + 4\alpha^4\left(1 - \frac{(3+\beta)}{2(1+\beta)M}\right)\right]}. \end{aligned} \tag{4.26}$$

It can be seen that the optimum receiver is able to harness the beamforming gain, which approaches $(1 + 2\alpha^2)$ as the cluster sizes M increase. Moreover, it is observed that the slope $\mathcal{S}_0$ increases

monotonically with the cell load β. This implies that, in the low-SNR regime, it is optimal to increase the cell load as much as possible. This conclusion can be justified by noting that in the low-SNR regime the main performance impediment is set by noise and not by interference, so that increasing the number of users per cell is generally advantageous.

High-SNR Analysis

We now turn to the high-SNR regime. For $M \geq 2$, the multiplexing gain is given by [117]

$$\mathcal{S}_\infty = \begin{cases} \beta, & \beta \leq \dfrac{M}{M+2} \\ 1 - \dfrac{2}{M}\beta, & \dfrac{M}{M+2} < \beta \leq 1 \\ 1 - \dfrac{2}{M}, & 1 \leq \beta. \end{cases} \tag{4.27}$$

From this result, it is observed that taking $\beta \to \infty$ is no longer optimum in this regime. In fact, the optimum value of β in terms of the high-SNR slope approaches $\frac{M}{M+2}$ as the SNR grows. Moreover, it is interesting to note that when $M = 2$ (and also $M = 1$ [143]), taking $\beta \to \infty$ makes the performance *interference limited*, so that the multiplexing gain is zero. In contrast, for $M \geq 3$, the performance of the optimum receiver is no longer interference limited when $\beta \to \infty$, although, as noted above, an unbounded cell load is a strictly suboptimum choice in the high SNR regime.

4.1.4.2 Linear MMSE Receiver

We now consider the performance of a more practical linear MMSE receiver. Noting that the quantity $\Gamma_M(m)(1 + 2\alpha^2)\beta P$ is the Signal-to-Interference-plus-Noise Ratio (SINR) at the output of the linear MMSE receiver, we obtain the ergodic sum-rate as [117]

$$R^{erg}_{M,MMSE}(P) = \frac{1}{M}\sum_{m=1}^{M} \beta \log(1 + \Gamma_M(m)(1 + 2\alpha^2)\beta P). \tag{4.28}$$

Low-SNR Analysis

The low-SNR characterization of the linear MMSE receiver is summarized as follows for $M \geq 2$ [117]:

$$\begin{aligned}\frac{E_b}{N_0}_{\min} &= \frac{\log 2}{\left[1 + 2\alpha^2\left(1 - \frac{1}{M}\right)\right]} \\ \mathcal{S}_0 &= \frac{2\beta}{1+2\beta}\frac{\left[1 + 2\alpha^2(1 - \frac{1}{M})\right]^2}{\left[1 + 4\alpha^2\left(1 - \frac{1+\beta}{(1+2\beta)M}\right) + 4\alpha^4\left(1 - \frac{3+4\beta}{(1+2\beta)M}\right)\right]}.\end{aligned} \tag{4.29}$$

This analysis shows that the linear MMSE receiver is able to harness the same beamforming gains as the optimum receiver. Moreover, the slope of Equation (4.29) monotonically increases with β which establishes the optimality of taking $\beta \to \infty$ for the linear MMSE receiver in the low-SNR regime, as shown for the optimum receiver.

4.2 Downlink

We now turn to the analysis of the achievable per-cell sum-rate for the downlink in the presence of MCP with unlimited-capacity backhaul for both fading soft-handoff and fading Wyner models. Focusing on ergodic fading, the per-cell sum-capacity is given by $R_{MCP}^{erg}(P) = \mathbb{E}[R_{MCP}(P)]$, where $R_{MCP}(P)$ is obtained as Equation (3.20). Similar to $R_{MCP}(P)$ for Gaussian models, this rate is attained via DPC (see Appendix A). Evaluating this quantity is even more complex than for the uplink due to the min–max operation involved, which arises as a consequence of the per-BS power constraint, as revealed by the duality analysis reported in Appendix B. Therefore, as for the uplink, we will focus on obtaining analytical insights via approximations and asymptotic results in the following.

4.2.1 Soft-handoff Model

We start by analyzing the soft-handoff model with $L = 1$ and $\alpha = 1$. The choice $L = 1$ is made for analytical convenience in order to

overcome the difficulties identified above (extension to any α does not pose additional challenges).

4.2.1.1 Large Number K of Users Per Cell

We focus here on the scenario in which the number of users per cell, K, is large and derive upper and lower bounds on the ergodic per-cell sum-rate $R_{MCP}^{erg}(P)$ for Rayleigh fading. The discussion follows [118]. Before starting the treatment, it is useful to recall that for K large [98] the per-cell sum-rate of a system *without inter-cell interference* scales as $\log\log K$. This means that, as the number of MSs in the system increases, the system is able to transmit at a larger sum-rate, while keeping the transmit power and bandwidth constant. As further discussed below, this gain is enabled by selecting, and thus allocating power to, the MSs with the best channel conditions. This can be seen as another instance of the benefits of multiuser diversity. We emphasize that this particular multiuser diversity gain was not considered for the uplink, since we have assumed that the MSs do not have CSI and thus cannot perform channel-aware scheduling (although in principle channel-aware scheduling can be performed at the CP if channel coherence time permits).

We now study the effects of inter-cell interference on the achievable rate by deriving lower and upper bounds on the ergodic per-cell sum-rate $R_{MCP}^{erg}(P)$. From Equation (3.20), a lower bound on $R_{MCP}^{erg}(P)$ can be obtained by fixing a diagonal matrix $\mathbf{D}$ and minimizing over the noise covariance $\mathbf{\Lambda}$. Fixing a matrix $\mathbf{D}$ is equivalent to deciding on a particular power allocation across the MSs in the dual uplink (see Section 3.2.2). A convenient scheduling for the dual uplink is the following: only users received in the *dual* uplink channel with fade power levels exceeding some threshold L_{th} are allowed to transmit. As $K \to \infty$, the number of active users per cell with this scheme can be shown to crystallize to $K_0 \triangleq K\mathrm{e}^{-2L_{\text{th}}}$, so that all active users can transmit at equal powers $1/K_0$ to meet the power constraint $tr(\mathbf{D}) \leq M$ in Equation (3.20). The threshold L_{th} is selected so that, as $K \to \infty$, we have $K_0 \to \infty$, so that the law of large numbers can be invoked. As demonstrated in [118], minimizing Equation (3.20) when $\mathbf{D}$ is selected as

above is accomplished by choosing $\mathbf{\Lambda} = \mathbf{I}$. This way, we have obtained a lower bound on $R_{MCP}^{erg}(P)$, which selecting $K_0 = K\mathrm{e}^{-2L_{\mathrm{th}}} = K^{\varepsilon}$, and hence $L_{\mathrm{th}} = \frac{1-\varepsilon}{2}\log K$, where $0 < \varepsilon < 1$, is given by

$$R_{MCP}^{erg}(P) \geq \log_2(1 + P((1-\varepsilon)\log K + 2)). \tag{4.30}$$

We remark that the considered scheduling scheme for the dual uplink channel translates into a downlink DPC-based transmission scheme that achieves the corresponding sum-rate via the transformations derived in [127]. We also note that considering a scheduling scheme, where *all* users transmit simultaneously with equal powers, produces an achievable average per-cell sum-rate of $\log(1 + 2P)$. This shows that the lack of scheduling fails to produce the multiuser diversity gain factor of $(1-\varepsilon)\log K$, confirming the discussion above regarding the performance gains attainable by only serving the users with the best channel conditions.

An upper bound on $R_{MCP}^{erg}(P)$ can be instead obtained by fixing a matrix $\mathbf{\Lambda} = \mathbf{I}$ and then optimizing over $\mathbf{D}$. A further upper bound on the so obtained optimal value can be found by bounding the channel fades by the strongest fading gain received at each cell site (over all intra-cell users), and observing that the maximum of K i.i.d. exponentially distributed random variables scales as $\log K$ (neglecting smaller terms) for $K \gg 1$ [98]. This leads to

$$R_{MCP}^{erg}(P) \leq \log_2(1 + 2P\log K). \tag{4.31}$$

The above bounds are rather tight, and for $\varepsilon \ll 1$ the gap between the two bounds is less than 1 bit/s/Hz in the high-SNR region. Moreover, one can conclude from the two bounds that the per-cell sum-rate capacity scales as $\log\log K$ for $K \gg 1$, as for the interference-free system. In other words, interference, when MCP is used, does not hamper the gains achievable via multiuser diversity. In fact, a suboptimal scheme proposed in [112] is shown to achieve the same scaling law even without employing optimal DPC encoding. The scheme is based on zero-forcing (ZF) beamforming and a simple user selection (scheduling) rule whereby one user is served in each cell at any given time in an intra-cell TDMA fashion, and is discussed in Section 4.2.3.

4.2.1.2 Extreme-SNR Analysis

We now analyze the performance of MCP for the downlink in the extreme SNR regimes by leveraging the bounds obtained above. Recall that the results here hold for large K and that we have set $\alpha = 1$. We first remark that the multiplexing gain with MCP is unitary. This is consistent with the results obtained for the uplink and shows the ability of MCP to mitigate the effect of interference.

We then study the low-SNR regime. The low-SNR parameters can be characterized (for any number of cells $M \geq 3$) as

$$\frac{\log 2}{2 \log K} \leq \frac{E_b}{N_{0\,\min}} \leq \frac{\log 2}{(1-\varepsilon)\log K + 2} \tag{4.32}$$

and

$$\mathcal{S}_0 = 2. \tag{4.33}$$

We recall that in the absence of fading, for the Gaussian soft-handoff model at hand, we have $E_b/N_{0\min} = \log 2/2$ and $\mathcal{S}_0 = 4/3$ [118]. Note that the factor 2 in $\frac{E_b}{N_{0\,\min}} = \log 2/2$ accounts for the beamforming gain in the model at hand. Comparing with the results in the absence of fading, the equations above demonstrate the beneficial effect of fading on the downlink. Specifically, the minimum transmit $\frac{E_b}{N_0}$ that enables reliable communication is decreased at least by a factor of $2/((1-\varepsilon)\log K + 2)$, while the low-SNR slope is increased by a factor of 1.5 for large K. This is a manifestation of the gains of multiuser diversity. In particular, here the multiuser gains are shown to depend on $\frac{E_b}{N_{0\,\min}}$, and to increase with, K, unlike the uplink (cf. Equation (4.8)). This can be explained following the same reasoning used above to justify the rate scaling of $\log\log K$ with increasing numbers of users, demonstrating the advantages of scheduling only the best users.

4.2.2 Wyner Model

In this section, we study the fading Wyner model with Rayleigh fading, $L = 1$ and any value of α^2, following a similar approach to the one used for the soft-handoff model.

4.2.2.1 Large Number K of Users Per Cell

Using the same bounding techniques discussed above, for an arbitrary number of cells ($M > 2$) with many users ($K \gg 1$), the per-cell sum-rate capacity for the Wyner model is bounded by (ignoring small orders of $\log K$)

$$\begin{aligned} &\log_2(1 + P((1-\varepsilon)\log K + 1 + 2\alpha^2)) \\ &\quad \le R_{MCP}(P) \le \log_2(1 + (1+2\alpha^2)P\log K), \end{aligned} \tag{4.34}$$

for some $\varepsilon \to 0$, as $K \to \infty$. The bounds (Equation (4.34)) show the same $\log\log K$ multiuser diversity gain as for the soft-handoff (and the inter-cell interference-free) system.

4.2.2.2 Extreme-SNR Analysis

Similar results as for the soft-handoff model are also obtained when analyzing the extreme-SNR behavior. In fact, in the high-SNR regime, the multiplexing gain can be again shown to be unitary, whereas in the low-SNR regime we have

$$\frac{\log 2}{(1+2\alpha^2)\log K} \le \frac{E_b}{N_{0\,\min}} \le \frac{\log 2}{(1-\varepsilon)\log K + 1 + 2\alpha^2} \tag{4.35}$$

and

$$\mathcal{S}_0 = 2. \tag{4.36}$$

4.2.3 Distributed Zero-Forcing Beamforming

In the previous sections, we have derived bounds on the maximum achievable per-cell sum-rate. As was discussed, the optimal performance is attained by DPC techniques, which are generally complex to implement. Here, instead, we analyze the performance of a practical and well-established alternative to optimal DPC schemes, namely zero-forcing beamforming (ZFBF) with a simple scheduling strategy. As will be shown, ZFBF is an attractive scheme, since it attains noninterference-limited performance, while requiring only single-user coding/decoding techniques. Analysis for a general MIMO Gaussian broadcast channel with sum-power constraint was presented in [136], while this section

discusses the analysis for the Gaussian and fading Wyner models (with per-BS power constraints), following [112]. We refer to [5, 7] for related recent results.

We assume the system to be operated according to an intra-cell TDMA protocol. ZFBF is based on transmission via linear precoding from all the BSs. In particular, the signal $\mathbf{x}$ transmitted by the BSs (see Equation (2.6)) is given by $\boldsymbol{x} = \mathbf{B}\mathbf{u}$, where $\boldsymbol{u}$ is the $M \times 1$ vector of symbols intended for the M scheduled users, one per cell, and $\mathbf{B}$ is the $M \times M$ beamforming matrix. The received signal (Equation (2.6)) thus becomes

$$\boldsymbol{y} = \mathbf{H}^{\dagger}\boldsymbol{B}\boldsymbol{u} + \mathbf{z}, \tag{4.37}$$

where we have neglected the subscript M in the channel matrix for simplicity of notation. The beamforming matrix $\mathbf{B}$ is selected so as to eliminate inter-cell interference. This leads to

$$\mathbf{B} = \sqrt{\frac{MP}{\mathrm{trace}\left((\mathbf{H}\mathbf{H}^{\dagger})^{-1}\right)}}(\mathbf{H}^{\dagger})^{-1}. \tag{4.38}$$

Note that the definition (Equation (4.38)) only ensures the sum-power constraint, but it will be argued below that it also satisfies the required per-BS power constraints. Substituting Equation (4.38) into Equation (4.37), the received signal vector reduces to

$$\boldsymbol{y} = \sqrt{\frac{MP}{\mathrm{trace}\left((\mathbf{H}\mathbf{H}^{\dagger})^{-1}\right)}}\ \mathbf{u} + \mathbf{z}. \tag{4.39}$$

Thanks to ZFBF, as seen in Equation (4.39), the downlink channel has been decomposed into a set of M identical independent parallel single-user channels, one for each MS.

We now consider separately the Gaussian and fading Wyner models.

4.2.3.1 Gaussian Wyner Model

From Equation (4.39), the per-cell sum-rate achievable by ZFBF is given by

$$R_{\mathrm{zfbf}}(P) = \log_2\left(1 + \frac{MP}{\mathrm{trace}\,(\mathbf{H}\mathbf{H}^{\dagger})^{-1}}\right), \tag{4.40}$$

where the channel matrix $\mathbf{H}$ is a symmetric circulant Toeplitz matrix with first column given by $[1\ \alpha\ \mathbf{0}_{M-2}]$. Moreover, by symmetry, it can be seen that the transmission power per-BS, which is given by

$$[\boldsymbol{B}\boldsymbol{B}^{\dagger}]_{m,m} = \frac{MP[(\mathbf{H}\mathbf{H}^{\dagger})^{-1}]_{m,m}}{\mathrm{trace}\,(\mathbf{H}\mathbf{H}^{\dagger})^{-1}} \tag{4.41}$$

is equal for all BSs m.

Applying Szegö's theorem [30], the average per-cell sum-rate of the ZFBF scheme is given for $\alpha < 1/2$, by

$$R_{\mathrm{zfbf}}(P) \underset{M\to\infty}{=} \log_2(1 + (1-4\alpha^2)^{\frac{3}{2}}P). \tag{4.42}$$

Equation (4.42) clearly shows the noninterference-limited behavior of ZFBF. Moreover, comparing Equation (4.42) with Equation (3.1), it can be seen that the ZFBF scheme is superior to SCP with spatial reuse $F = 2$ when the SNR P is above a certain threshold

$$P_{\mathrm{t}}(\alpha) = \frac{2(1-(1-4\alpha^2)^{\frac{3}{2}})}{(1-4\alpha^2)^3}, \tag{4.43}$$

which is an increasing function of α. As a technical remark, it is noted that for $\alpha = 1/2$ the circulant channel transfer matrix $\boldsymbol{H}$ is singular and channel inversion methods such as ZFBF are not applicable. Moreover, $\mathbf{H}$ is not guaranteed to be nonsingular for $\alpha > 1/2$ and any finite number of cells M.

4.2.3.2 Fading Wyner Model

Having obtained some insight into the performance of ZFBF in a setting without fading, we now turn to the analysis of the fading Wyner model. In order to implement intra-cell TDMA while still leveraging multiuser diversity, we consider a simple scheduling algorithm that selects for each fading block (or TDMA slot) the user with the maximum channel gain for transmission in each cell. The resulting channel transfer matrix $\mathbf{H}^{\dagger}$ of this suboptimal scheduling consists of diagonal entries whose absolute values squared are the *maximum* of K independent exponentially distributed random variables (as for Rayleigh fading).

The other nonzero entries are on the two diagonals above and below the main diagonal and are complex Gaussian random variables with power α^2. It can be proved that this scheme ensures *in probability* an equal per-cell power (Equation (4.41)) equal to P, asymptotically with increasing number of users per-cell. We finally remark that, in case matrix $\mathbf{H}$ is ill-conditioned, the CP can start replacing the "best" users by their second "best" users until the resulting matrix $\mathbf{H}$ is well behaved. Since we assume that $K \gg 1$, the overall statistics are not expected to change by this user-replacing procedure.

We now need to evaluate the ergodic per-cell sum-rate

$$R_{\text{zfbf}}^{erg}(P) = E\left\{\log_2\left(1 + \frac{MP}{\text{trace}\left((\mathbf{H}\mathbf{H}^\dagger)^{-1}\right)}\right)\right\}. \tag{4.44}$$

Leveraging the special structure of the channel transfer matrix $\mathbf{H}^\dagger$ described above, it can be proved that the ergodic per-cell sum-rate capacity (Equation (4.44)) scales as $\log\log K$ as the number K of users per cell increases [112]. This demonstrates that the ZFBF is able to fully harness the multiuser diversity gains in the regime of large K. This result, can be intuitively explained by the fact that due to the scheduling process, $(\mathbf{H}\mathbf{H}^\dagger)$ "becomes" diagonal $(\log K\mathbf{I}_M)$ as K increases. Accordingly, for large K, $(\mathbf{H}\mathbf{H}^\dagger)^{-1}$ "behaves" like $(\mathbf{I}_M/\log K)$, and $R_{\text{zfbf}}^{erg}(P)$ in Equation (4.44) is approximated by

$$R_{\text{zfbf}}^{*} \underset{K\gg 1}{\cong} \log_2(1 + P\log K). \tag{4.45}$$

4.2.3.3 Numerical Results

In Figure 4.4 the ergodic per-cell sum-rate of the ZFBF scheme is compared with the sum-rate capacity (Equation (3.20)), obtained via Monte Carlo simulations, for $K = 100$ and $\alpha = 0.4$. The gap between the performance of ZFBF and the capacity is explained by the fact that the ZFBF scheme does not exploit antenna gains, but only attempts to eliminate inter-cell interference. Performance is also compared with SCP with $F = 1$ and $F = 2$.

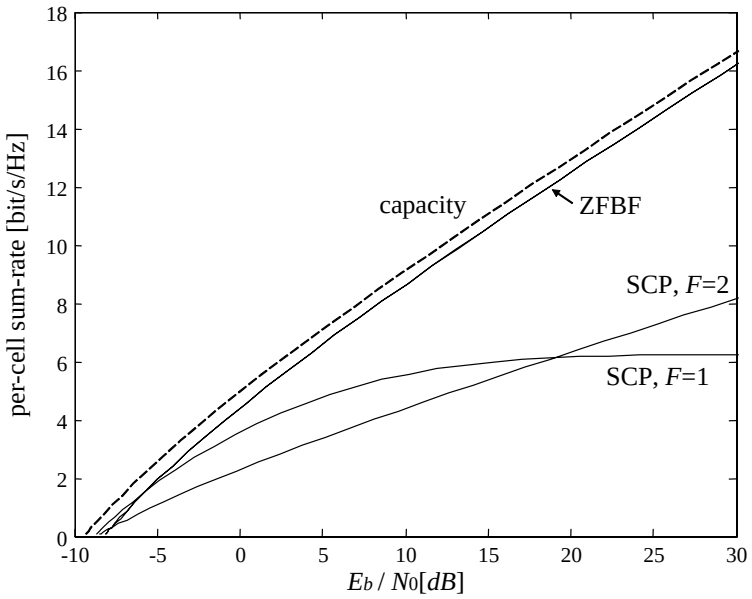

Fig. 4.4 Per-cell sum-rates versus E_b/N_0 for the downlink of the fading Wyner model (Rayleigh fading, $L = 1$, $\alpha = 0.4$).

4.3 Summary

In this section, we have analyzed the performance benefits of MCP in the presence of fading. A first observation has been that the beamforming and multiplexing gains that were identified in the previous section in the absence of fading continue to hold with fading. Moreover, we have seen that, with fading, additional gains become available due to *multiuser diversity.* The latter derives from the presence of multiple users per cell with independent channel realizations. Therefore, while without fading the number K of users per cell does not play a relevant role (under the assumption that the power remains constant with K), this is not the case in fading channels. Specifically, for the uplink, under the assumption of channel-independent scheduling, we have observed that multiuser diversity gains arise if K is sufficiently large, while for small K having fewer users per cell can be preferable. Instead, for the downlink, with channel-aware scheduling, MCP enables the system to harness a logarithmic gain as a function of K due to multiuser diversity for

all K and SNR levels. We emphasize that, if channel-aware scheduling is allowed for the uplink as well, then it is possible to obtain the same multiuser diversity gains obtained here for the downlink (see [118]). We have also studied the performance of practical system implementations in both the uplink, via DS-CDMA, and downlink, via ZFBF, and obtained performance comparable with that obtained under ideal conditions.

5

Constrained Multi-Cell Processing: Uplink

The previous sections analyzed the performance of MCP with unlimited-capacity backhaul links to a CP. The analysis has thus assessed the advantages of MCP with respect to SCP, with possible spatial reuse, under ideal conditions for the backhaul links. This leaves open the issue of addressing the actual performance advantages of MCP once limitations on the backhaul links are accounted for. This is the subject of this section, where we focus on the uplink. Discussion of the downlink is postponed to the next section. We follow the classification in Section 2.3 and consider first limited-capacity backhaul links to the CP, then finite-capacity links between adjacent BSs and finally clustered cooperation. We will consider Gaussian and (ergodic) fading Wyner and soft-handoff models.

5.1 Finite-Capacity Backhaul to a Central Processor

In this section, we consider the impact of finite-capacity backhaul between BSs and CP, as shown in Figure 2.4(a). Recall that we assume that each BS is connected to the CP via a backhaul link of capacity C (bits/s/Hz).

5.1.1 Upper Bound and Definitions

To provide a benchmark result, we observe that the achievable per-cell sum-rate can be upper bounded by the so-called cut-set upper bound (see, e.g., [17]). This bound dictates that the rate $R(P,C)$ achievable by any scheme with power constraint P and backhaul capacity C must satisfy

$$R(P,C) \leq R_{UB}(P,C) = \min\{C, R_{MCP}(P)\}, \tag{5.1}$$

where $R_{MCP}(P)$ is defined in Equation (3.8) and represents the per-cell sum-rate capacity with unlimited-capacity backhaul links. For ergodic fading, a similar relationship holds with $R^{erg}_{MCP}(P) = \mathrm{E}[R_{MCP}(P)]$ in lieu of $R_{MCP}(P)$. The cut-set bound (Equation (5.1)) merely states that the per-cell sum-capacity cannot be larger than the minimum of the (per-cell) backhaul capacity C and of the rate attainable with ideal backhaul.

We now study how close one can get to the upper bound (Equation (5.1)). We do this by considering separately two types of MCP strategies.

- ***Oblivious BSs***: The BSs are assumed to have no decoding power. This enables implementation with "light" BSs, that only need to have a radio transceiver and minimal baseband processing capabilities. The BSs are also not required to be informed about the MSs' codebooks. We refer to this scenario as "oblivious BSs". Note that, due to capacity limitations on the digital link to the CP, oblivious BSs cannot directly forward the received radio signal (say, using radio-over-fiber technologies), but instead have to compress the received signal before communicating to the CP.
- ***Informed BS***: The BSs have full decoding capabilities. We refer to this scenario as "informed BSs", since in this case the BSs must be aware of the MSs' codebooks in order to perform decoding.

5.1.2 Oblivious Base Stations

Here we study the performance achievable with oblivious BSs for Gaussian Wyner and Gaussian soft-handoff models. The corresponding fading models are studied in Section 5.6. For simplicity, we focus on the case $K = 1$, or equivalently intra-cell TDMA. This is without loss of generality, given the optimality of the intra-cell TDMA in Gaussian models (see Section 2).

We consider the following scheme, proposed in [85]. The basic idea is that each oblivious BS compresses the received signal to C (bits/s/Hz) and sends some function of the compressed index to the CP. The CP then performs decoding based on the information received from the BSs. To be more specific, design of the compression scheme can leverage the fact that the signals received by different BSs are correlated, since they are the result of the same set of signals propagating from MSs to BSs. This allows the BSs to reduce the rate, and thus the backhaul capacity, required to communicate the compression indices to the CP. Compression strategies that leverage signal correlation are typically referred to as distributed source coding and the compression problem at hand as the "CEO problem", which is introduced in Appendix C.

As discussed in Appendix C, distributed source coding strategies operate by transmitting a function (referred to as a "bin") of the compression index to the CP, thus reducing the required backhaul rate. Thanks to the signal correlation, the CP is in principle able to recover the compression indices, and hence the signals compressed by the RSs (i.e., the CP "decompresses"). Decoding can then take place at the CP based on the decompressed signals. Reference [85] instead proposes a different, and potentially better, approach, whereby the CP performs *joint* decompression and decoding of all compressed signals and messages of all active MSs. Incidentally, this approach to joint decompression and decoding has been also later found to be useful in many other scenarios [65]. We remark that the scheme at hand can also be defined as "Compress-and-Forward" (CF) following standard terminology for relaying [48] (see also Section 6).

References [85, 86] prove that the following per-cell rate is achievable by the scheme discussed above for a system characterized by the channel matrix $\mathbf{H}$ as $M \to \infty$:

$$R_{\text{obl}}(P,C) = \lim_{M\to\infty} \frac{1}{M} \log_2 \det(\mathbf{I} + (1-2^{-r^*})P\mathbf{H}\mathbf{H}^\dagger), \tag{5.2}$$

where r^* is the solution of the equation[1]

$$F(P,r^*) = C - r^*, \tag{5.3}$$

in which we have defined the function

$$F(P,r) \triangleq \lim_{M\to\infty} \frac{1}{M} \log_2 \det(\mathbf{I} + (1-2^{-r})P\mathbf{H}\mathbf{H}^\dagger). \tag{5.4}$$

Note that in this section, we do not explicitly denote the fact that the achievable rates correspond to the regime $M \to \infty$ in order to simplify the notation.

Since the result (Equation (5.2)) holds for any channel matrix $\mathbf{H}$, it applies in particular for the Gaussian Wyner and Gaussian soft-handoff models. Comparing the rate (Equation (5.2)) with the per-cell rate (Equation (3.8)) achievable with no limitations on the backhaul, it can be seen that a finite-capacity backhaul entails an *SNR loss* of $1/(1-2^{-r^*})$. Note in fact that, when $C \to \infty$, then we have $r^* \to \infty$ and Equation (5.2) reduces to the per-cell rate capacity (Equation (3.8)). The parameter r^* can be interpreted as the rate that the BSs have to invest for compressing the noise in the received signals [83]. In this regard, we emphasize that the BS cannot avoid compressing channel noise since it cannot decode the useful signal and thus separate signal and noise. The condition (Equation (5.3)) imposes the requirement that the achievable per-cell rate be equal to the amount of backhaul capacity left once the portion r^* wasted for noise compression is accounted for. Finally, we emphasize that $R_{\text{obl}}(P,C) \leq C$ due to (5.3), as it should be by the upper bound (5.1).

It can also be seen that if no inter-cell interference is present, i.e., $\alpha = 0$, the rate (Equation (5.2)) reduces to [85]

$$R_{\text{obl}}(P,C) = \log_2\left(1 + P\frac{1-2^{-C}}{1+P2^{-C}}\right). \tag{5.5}$$

[1] For finite C, the implicit (Equation (5.3)) is easily solved numerically, since $F(P,r)$ can be shown to be monotonic in r for the symmetric models at hand.

Comparison with the rate $\log_2(1+P)$ achievable with ideal backhaul links shows an SNR degradation of $\frac{1+P2^{-C}}{1-2^{-C}}$.

We can now specialize the general result (Equation (5.2)) to both the Gaussian Wyner and the Gaussian soft-handoff models with $L=1$. For the Wyner model, we get

$$\begin{aligned} R_{obl}(P,C) &= F(P,r^*) \\ &= \int_0^1 \log_2(1+P(1-2^{-r^*})(1+2\alpha\cos 2\pi\theta)^2)d\theta, \end{aligned} \tag{5.6}$$

where r^* is the unique solution to the fixed point equation $F(P,r^*) = C - r^*$. Alternatively, for the soft-handoff model, it is possible to get a more explicit expression as

$$\begin{aligned} R_{\text{obl}}(P,C) &= F(P,r^*) \\ &= \log_2\left(\frac{1+(1+\alpha^2)P+2\alpha^2 2^{-C}P^2}{2(1+2^{-C}P)(1+\alpha^2 2^{-C}P)}\right. \\ &\quad \left. + \frac{\sqrt{1+2(1+\alpha^2)P+((1-\alpha^2)^2+4\alpha^2 2^{-C})P^2}}{2(1+2^{-C}P)(1+\alpha^2 2^{-C}P)}\right). \end{aligned} \tag{5.7}$$

Note that the two results above reduce to Equations (3.12) and (3.15), respectively, for $C \to \infty$.

5.1.2.1 Low-SNR Analysis

Next we study the low-SNR characterization of the oblivious scheme discussed above. It is proved in [86] that the low-SNR parameters are given by

$$\frac{\mathrm{E_b}}{\mathrm{N_0}_{\min}} = \frac{\widetilde{\mathrm{E_b}}}{\mathrm{N_0}_{\min}} \frac{1}{1-2^{-C}} \tag{5.8}$$

and

$$\mathcal{S}_0 = \widetilde{\mathcal{S}_0} \frac{1}{1+\widetilde{\mathcal{S}_0}\frac{2^{-C}}{1-2^{-C}}}, \tag{5.9}$$

where $\frac{\widetilde{\mathrm{E_b}}}{\mathrm{N_0}_{\min}}$ and $\widetilde{S_0}$ are the minimum transmitted energy per bit required for reliable communication and the low-SNR slope of the

per-cell capacity with unlimited backhaul links, namely $R_{MCP}(P) = R_{\text{obl}}(P,\infty) = F(P,\infty)$.

This result confirms and quantifies for the low-SNR regime the insight offered above on the effect of a limited-capacity backhaul. In particular, it is seen that oblivious BSs are able to obtain the ideal performance with unlimited backhaul up to an SNR loss given by $1/(1-2^{-C})$. Therefore, for instance, by allocating at least $C \approx 3.2$ (bits/s/Hz) to the backhaul network, the minimum energy required for reliable communication of the limited channel will not increase by more than 0.5 (dB) when compared to that needed with unlimited backhaul.

5.1.2.2 High-SNR Analysis

Here we study the high-SNR characterization of the performance achievable with oblivious BSs. Similar to the results in the previous subsection, the high-SNR analysis is general and the results are applicable for Gaussian Wyner and soft-handoff models. We first observe from the upper bound (Equation (5.1)) that for a fixed backhaul capacity C and large SNR P, the per-cell rate is at most C. Note that this rate is achieved by the oblivious strategy for $P \to \infty$, as can be seen from Equations (5.2) and (5.3). It follows that, for fixed backhaul capacity C, no matter which strategy is used, the rate is finite and the system has a zero multiplexing gain (i.e., $S_\infty = 0$).

Based on the discussion above, the following question arises: How should the backhaul capacity C scale with P in order for the system to maintain a unit multiplexing gain (as in the inter-cell interference-free system)? Reference [86] shows that a scaling of $\log_2 P$ suffices to achieve unit multiplexing gain (and in fact also the high-SNR offset of an interference-free system).

5.1.2.3 Fading Channels

In this section, we address the performance of MCP with oblivious BSs for *fading* Wyner and soft-handoff models. We assume, as usual, Rayleigh fading with unit power. A first observation is that, as seen in the previous section, with fading, the intra-cell TDMA protocol is no

longer optimal and the WB protocol is to be preferred. We thus adopt the WB protocol in the following and set $K \geq 1$. Similar to Equation (5.2), it can be shown that the oblivious scheme, implemented along with the WB protocol, achieves the following ergodic per-cell sum-rate:

$$\begin{aligned} R_{obl}^{erg}(P,C) &= F(P,r^*) \\ &= \lim_{M\longrightarrow\infty} \max_{\mathrm{E}[r_m(\mathbf{H})]=r^*} E\left[\frac{1}{M}\log_2\left(\det\left(\mathbf{I}\right.\right.\right. \\ &\qquad \left.\left.\left. + \frac{P}{K}\mathrm{diag}(1-2^{-r_m(\mathbf{H})})_{m=1}^{M}\mathbf{H}\mathbf{H}^{\dagger}\right)\right)\right] \end{aligned} \tag{5.10}$$

where the functions $r_m(\mathbf{H})$ must be chosen so that condition (Equation (5.3)) holds. Note that in Equation (5.10), the rate $r_m(\mathbf{H})$ is to be interpreted as the rate that the m-th BS has to invest for compressing channel noise when the channel realization is $\mathbf{H}$. In order to simplify the evaluation of Equation (5.10), we will consider a lower bound on $R_{obl}^{erg}(P,C)$, which is achieved by setting $r_m(\mathbf{H}) = r^*$, irrespective of the channel realization.

Unfortunately, an explicit evaluation of Equation (5.10) is complicated in general. Therefore, in the following, we will consider various approximations and asymptotic regimes in order to get further insight into Equation (5.10). We consider the Wyner model and the soft-handoff model separately.

Wyner Model

For the Wyner model with $L = 1$, assuming that the number of users K per-cell is large (and also M is large as in Equation (5.10)), it is possible to evaluate Equation (5.10) as

$$R_{obl}^{erg}(P,C) = \log_2\left(1 + \frac{(1+2\alpha^2)P(1-2^{-C})}{1+(1+2\alpha^2)P2^{-C}}\right). \tag{5.11}$$

Hence, the rate of the limited network equals the rate of the single user Gaussian channel (Equation (5.5)) but with enhanced power $P(1+2\alpha^2)$. This complies with the observation made around Equation (4.4), to which Equation (5.11) reduces for $C \to \infty$. In particular, comparison with Equation (4.4) shows again that the effect

of finite capacity, even in the presence of fading, can be accounted for by an SNR loss, which in the regime of large K is given by $(1 + (1 + 2\alpha^2)P2^{-C})/(1 - 2^{-C})$.

Soft-Handoff Model

For the soft-handoff model with $L = 1$, reference [86] obtains an upper bound:

$$R_{obl}^{erg}(P,C) \leq \log_2 \left(\frac{1 + P(1+\alpha^2) + 2P^2\alpha^2 2^{-C}/K}{2(1 + P(1+\alpha^2)2^{-C} + P^2\alpha^2 2^{-2C}/K)} + \frac{\sqrt{\left(1 + P(1+\alpha^2)\right)^2 - 4P^2\alpha^2(1-2^{-C})/K}}{2(1 + P(1+\alpha^2)2^{-C} + P^2\alpha^2 2^{-2C}/K)} \right), \quad (5.12)$$

which can be shown to be tight in the regimes $C \to \infty$ and/or for $K \to \infty$. In the latter case, in particular, we obtain

$$R_{obl}^{erg}(P,C) = \log_2 \left(1 + P\frac{(1+\alpha^2)(1-2^{-C})}{1 + P(1+\alpha^2)2^{-C}} \right), \quad (5.13)$$

which shows the same SNR loss as compared to the unlimited capacity rate as found above for the Wyner model [86].

5.1.3 Informed Base Stations

As shown in the previous section, an implementation of MCP with oblivious BSs is able to closely approximate the performance achievable with an unlimited-capacity backhaul (as studied in the previous section) as long as the capacity C is large enough (see also Section 5.1.5). We now turn to assessing the further performance gains that can be accrued by endowing the BSs with decoding capabilities. These potential gains can be explained in light of Equations (5.2) and (5.10). In fact, it was discussed above that Equations (5.2) and (5.10) demonstrate that the performance with oblivious BSs is limited by the fact that the BSs inevitably have to invest some of the backhaul link capacity to compress noise, rather than signal, information. By

leveraging decoding capabilities the BSs may instead be able to distinguish, at least partially, signal from noise, thus decreasing the effective noise and increasing the efficiency of the backhaul usage. It is noted that the discussed performance gains come at the price not only of more processing power at the BSs, but also of the signaling necessary to inform the BSs of the codebooks selected by the MSs of interest (see below). Thus, this is the informed BSs situation mentioned above.

As an initial remark on the advantages of informed BSs, we note that local decoding at the BSs is expected to be beneficial when inter-cell interference is small and/or when the backhaul capacity C is small. In fact, if inter-cell interference is small, decoding can take place at a BS without hindrance from other cells' signals and, furthermore, no performance gains can be accrued by leveraging the signals received by other cells via MCP. Instead, if C is sufficiently small, as discussed above, it becomes imperative to use the backhaul more efficiently than with oblivious BSs. The analysis below quantifies these considerations.

To assess the benefits of informed BSs, we focus on Gaussian models and we evaluate the performance of a simple scheme that leverages decoding at the BSs. Analysis of the corresponding performance for fading models can be found in [86]. We will limit the analysis to one user per cell, $K = 1$, or equivalently to intra-cell TDMA. Analysis of a more sophisticated scheme is given in Section 5.1.4

According to the scheme considered here, each user splits its information bits (message) into two parts. It then transmits the sum of two signals, one encoding the first part of the message with power βP and the other encoding the second part with the remaining power $(1 - \beta)P$, with $0 \leq \beta \leq 1$. The first part of the message is intended to be decoded at the CP, while the latter is decoded by the same-cell BS. The idea is that the BS decodes part of the message of the same-cell MS and sends the decoded bits directly to the CP. The decoded signal can then be cancelled from the received signal, which is finally compressed as was done in the oblivious strategy presented above. We will refer to the messages to be decoded by the BS as "local" messages.

We denote the rate of the local message in each cell as $R_d(\beta)$, where we emphasize the dependence on the power split factor β. To evaluate

this rate, we assume that the BS either tries to decode the local message of the same-cell MS by treating all other signals as noise, or instead attempts joint decoding of all the local messages of the MSs, whose signal it receives. Specifically, in the latter case, for the Wyner model with $L = 1$, joint decoding of the signal of the same-cell MS and of the MSs in the two adjacent cells is performed, while for the soft-handoff model with $L = 1$ the MS of only one adjacent cell is received and decoded. By selecting the best decoder given the channel conditions, using well-known results on multiple access channels (see, e.g., [17]), we obtain, for the Wyner model [95]

$$\begin{aligned}\bar{R}_d(\beta) = \max\Bigg\{&\log_2\left(1+\frac{(1-\beta)P}{1+(\beta+2\alpha^2)P}\right),\\ &\min\Bigg\{\frac{1}{2}\log_2\left(1+\frac{(1-\beta)2\alpha^2P}{1+\beta(1+2\alpha^2)P}\right),\\ &\quad\frac{1}{3}\log_2\left(1+\frac{(1+2\alpha^2)(1-\beta)P}{1+\beta(1+2\alpha^2)P}\right)\Bigg\}\Bigg\},\end{aligned} \tag{5.14}$$

and for the soft-handoff model

$$\begin{aligned}\bar{R}_d(\beta) = \max\Bigg\{&\log_2\left(1+\frac{(1-\beta)P}{1+(\beta+\alpha^2)P}\right),\\ &\frac{1}{2}\log_2\left(1+\frac{(1-\beta)(1+\alpha^2)P}{1+\beta(1+\alpha^2)P}\right)\Bigg\}.\end{aligned} \tag{5.15}$$

The actual local message rate is

$$R_d(\beta) = \min\{\bar{R}_d(\beta), C\}, \tag{5.16}$$

since the decoded message must be sent to the CP over the backhaul, and thus enough capacity on the latter must be available.

Having decoded and cancelled the local message of the same-cell MS, each BS, as explained above, compresses the received signal as for the oblivious scheme. This leads to the following achievable per-cell rate:

$$R_{inf}(P,C) = \max_{0\le\beta\le1}\{F(\beta P, r^*) + R_d(\beta)\}, \tag{5.17}$$

where $F(\beta P, r^*)$ is defined in Equation (5.4) and evaluates as Equations (5.6) and (5.7) for the Gaussian Wyner and soft-handoff models with $L = 1$, respectively. Moreover, the rate r^* is the solution of the equation

$$F(\beta P, r^*) = C - R_d(\beta) - r^*. \tag{5.18}$$

The rate (Equation (5.17)) can be interpreted in light of the rate splitting strategy at hand: the overall per-cell rate is the sum of the rate $R_d(\beta)$ of the local message and the rate $F(\beta P, r^*)$ of the message to be decoded at the CP using the oblivious strategy. As for the latter, note that the amount of backhaul capacity left after transmission of the local message is $C - R_d(\beta)$, so that the condition (Equation (5.3)) is modified as in Equation (5.18). Moreover, the rate (Equation (5.17)) can be optimized with respect to the power split factor β. In this regard, if no inter-cell interference is present ($\alpha = 0$), it is easy to see that the power allocation $\beta = 0$ maximizes Equation (5.17), so that only local messages are transmitted. In fact, in the absence of interference, there is nothing to be gained by allowing centralized decoding at the CP and a multihop decode-and-forward scheme is to be preferred. Indeed, via comparison with the cut-set bound (Equation (5.1)), it can be seen that with $\alpha = 0$, the considered scheme with $\beta = 0$ is optimal. It is expected that increasing β becomes more and more advantageous as α increases, due to the potential advantages of MCP when the inter-cell channel gains are large enough (see Figure 3.2).

We finally remark that it turns out that the rate $R_{inf}(P, C)$ can be sometimes improved by a time-sharing technique, in which a fraction of time is devoted to transmitting only local messages ($\beta = 0$), while for the rest of time the oblivious strategy ($\beta = 1$) is used. In the following, further analysis of this time-sharing technique with optimized time-sharing fraction is provided in extreme-SNR regimes.

5.1.3.1 Low-SNR Analysis

Reference [86] shows that the low-SNR characterization of the time-sharing based scheme discussed above (which uses local decoding and

oblivious operation in different time-slots) is as follows:

$$\begin{aligned}\frac{\mathrm{E_b}}{\mathrm{N_0}\,\mathrm{min}} &= \frac{1}{\lambda_o\left(\frac{\mathrm{E_b}}{\mathrm{N_0}\,\mathrm{min}}^{\mathrm{d}}\right)^{-1}+(1-\lambda_o)\left(\frac{\mathrm{E_b}}{\mathrm{N_0}\,\mathrm{min}}^{\mathrm{obl}}\right)^{-1}} \quad \text{and} \\ \mathcal{S}_0 &= \frac{\left(\frac{\mathrm{E_b}}{\mathrm{N_0}\,\mathrm{min}}^{\mathrm{dec}}\right)^{-2}}{\lambda_o\left(\mathcal{S}_0^{\mathrm{d}}\right)^{-1}\left(\frac{\mathrm{E_b}}{\mathrm{N_0}\,\mathrm{min}}^{\mathrm{d}}\right)^{-2}+(1-\lambda_o)\left(\mathcal{S}_0^{\mathrm{obl}}\right)^{-1}\left(\frac{\mathrm{E_b}}{\mathrm{N_0}\,\mathrm{min}}^{\mathrm{obl}}\right)^{-2}},\end{aligned} \tag{5.19}$$

with optimized time-fraction

$$\lambda_o = 1 - \frac{C}{r^*} \tag{5.20}$$

and $r^* = \max\{C, \tilde{r}\}$, where $\tilde{r}$ is the unique solution of

$$2^{-\tilde{r}}(1+\tilde{r}\log 2) = \frac{2\alpha^2}{1+2\alpha^2}, \tag{5.21}$$

for the Gaussian Wyner model and

$$2^{-\tilde{r}}(1+\tilde{r}\log 2) = \frac{\alpha^2}{1+\alpha^2}\ , \tag{5.22}$$

for the Gaussian soft-handoff model. In Equation (5.19) the superscript $(\cdot)^{\mathrm{obl}}$ indicates the low-SNR parameters of the oblivious scheme obtained in Section 5.5, where r^* as defined above is written in place of C, and the notation $(\cdot)^{\mathrm{d}}$ indicates the low-SNR parameters of the local decoding rate $R_d(\beta)$ (Equation (5.16)). These can be easily calculated as

$$\frac{E_b}{N_0\,\mathrm{min}}^{\mathrm{d}} = \log 2; \quad \mathcal{S}_0^{\mathrm{d}} = \frac{2}{1+4\alpha^2}, \tag{5.23}$$

for the Wyner model and

$$\frac{E_b}{N_0\,\mathrm{min}}^{\mathrm{d}} = \log 2; \quad \mathcal{S}_0^{\mathrm{d}} = \frac{2}{1+2\alpha^2}, \tag{5.24}$$

for the soft-handoff model.

Equation (5.19) relates the low-SNR parameters of the time-sharing scheme at hand with those of the two strategies upon which it is based, namely transmitting only local messages ($\beta = 0$) and the oblivious

strategy ($\beta = 1$). The parameter λ_o in Equation (5.20) represents the optimized fraction of time that is devoted to transmitting local messages. From Equation (5.20), it can be thus seen that, as the backhaul capacity C increases, it becomes more advantageous to employ the oblivious scheme for a larger fraction of time. In particular, allocating resources (time) to transmitting local message is beneficial in the low-SNR regime only when C is below a certain threshold $\tilde{r}$ calculated from Equation (5.21) or Equation (5.22). The parameter $\tilde{r}$ can be seen to be a decreasing function of the intra-cell interference factor α, which implies that the range of values of C for which local decoding is advantageous decreases for increasing α. This quantifies the initial observations made above about the expected gains of informed BSs. As an example, when there is no inter-cell interference ($\alpha = 0$) then $\tilde{r} = \infty$ and local decoding is optimal for any C. In contrast, for $\alpha = 0.2$ numerical calculation reveals that $\tilde{r} \approx 2.15$. Hence, deploying informed BSs is beneficial when $C \lesssim 2.15$.

5.1.3.2 High-SNR Analysis

As seen above for $\alpha = 0$, using only local decoding is optimal. We thus focus on the case $\alpha > 0$. In this setting, it is easy to see that, as the SNR P and the capacity C grow large, it becomes less and less beneficial to transmit local messages. In fact, using only local decoding, the multiplexing gain would be limited to 1/3 for the Wyner model and 1/2 for the soft-handoff model, even for an unlimited-capacity backhaul (see, e.g., [119]). Instead, as seen in Section 5.1.2.2, oblivious BSs can achieve the interference-free multiplexing gain of one if C scales sufficiently fast with P.

We finally note that the qualitative conclusions given in this section on the performance of informed BSs carry over, with minor modifications, also to fading channels, as it is discussed in [86].

5.1.4 Informed Base Stations with Structured Coding

The potential performance gains of informed BSs were studied above based on the analysis of a simple transmission scheme, whereby each BS decodes parts of the signals transmitted by the MSs in their radio

range. In [79], instead, it is proposed that the BSs, rather than decoding the individual messages (or parts thereof) of the MSs, decode a *function* of such messages or, more precisely, of the corresponding transmitted codewords. The key idea that enables this operation is the use of a special class of codebooks that have the property that a sum of codewords is another codeword in the same codebook. This implies that decoding a single codeword from the codebook is as easy, or as difficult, as decoding a sum of two codewords. A class of codes that have this property is given by nested lattice codes. An introduction to nested lattice codes can be found in [145] and references therein.

To elaborate, each MS employs the same nested lattice code and the signal received at any m-th BS can be written for the Wyner Gaussian model from Equation (2.1) as $y_m = \sum_{k=-L}^{L} \alpha_k x_{m-k} + z_m$. Recalling that a lattice code is a discrete group, the (modulo[2]) sum of the lattice codeword x_{m-k}, weighted by *integer* coefficients, is still a codeword in the same lattice code and can thus be decoded by the m-th BS. The problem is that the channel coefficients α_k are generally not integers. The m-th BS can, however, decode an arbitrary linear combination $\sum_{k=-L}^{L} b_k x_{m-k}$ with $b_k \in \mathbb{Z}$ (and by symmetry $b_k = b_{-k}$) and $b_0 \neq 0$ and treat the remaining part of the signal as Gaussian noise. The index of the decoded codeword can then be sent to the CP, which decodes based on all received linear combinations. This leads to the achievable rate [79][3]

$$R_{LAT}(P,C) = \min\left\{C, \max_{(b_0,\ldots,b_L)\in\mathcal{B}} -\log_2\left(b_0^2 + 2\sum_{k=1}^{L} b_k^2 - \frac{P(b_0 + 2\sum_{k=1}^{L}\alpha_k b_k)^2}{1 + P(1 + 2\sum_{k=1}^{L}\alpha_k^2)}\right)\right\}, \tag{5.25}$$

where $\mathcal{B} = \{(b_0,\ldots,b_L) \in \mathbb{Z} :\ b_0 \neq 0 \ \text{ and } \ b_0^2 + 2\sum_{k=1}^{L} b_k^2 \leq 1 + P(1 + 2\sum_{k=1}^{L}\alpha_k^2)\}$.

A low-complexity implementation of the scheme outlined above, which is often referred to as "compute-and-forward" is proposed in [36].

[2] The modulo operation is taken with respect to the coarse lattice forming the nested lattice code.

[3] [79] considers the special case $L = 1$.

Reference [79] also proves that the rate (Equation (5.25)) can be improved upon by superimposing additional "private" messages to be decoded at the local BS to the lattice codewords. This modifies the interference pattern so that the sharp performance degradations observed for some values of the inter-cell channel gains (see Figure 5.2) are mitigated.

5.1.5 Numerical Results

We now provide some numerical examples that corroborate the results in this section pertaining to the effect of a limited-capacity backhaul. Consider a Gaussian Wyner model with $L = 1$, $K = 1$ user per cell and $P = 10\,\text{dB}$. In Figure 5.1, the per-cell sum-rate is shown versus the inter-cell gain α for the case of unlimited-capacity backhaul, oblivious BSs and informed BSs with $C = 6$ (bits/s/Hz). It can be seen that the limitation on the capacity of the backhaul leads to a fairly small loss of per-cell sum-rate even for values of backhaul capacity that are less than double the rate transmitted per cell. Moreover, as expected from

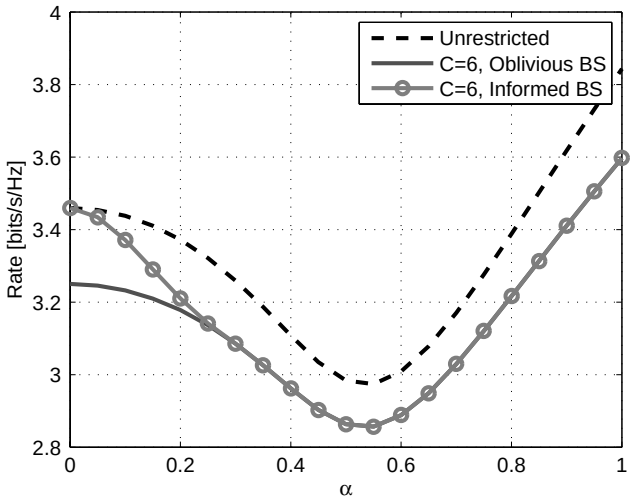

Fig. 5.1 Per-cell sum-rates versus the inter-cell gain α for unlimited-capacity backhaul, oblivious BSs and informed BS for a Gaussian Wyner model with $L = 1$, $K = 1$ user per cell, $C = 6$ (bits/s/Hz) and $P = 10$ dB.

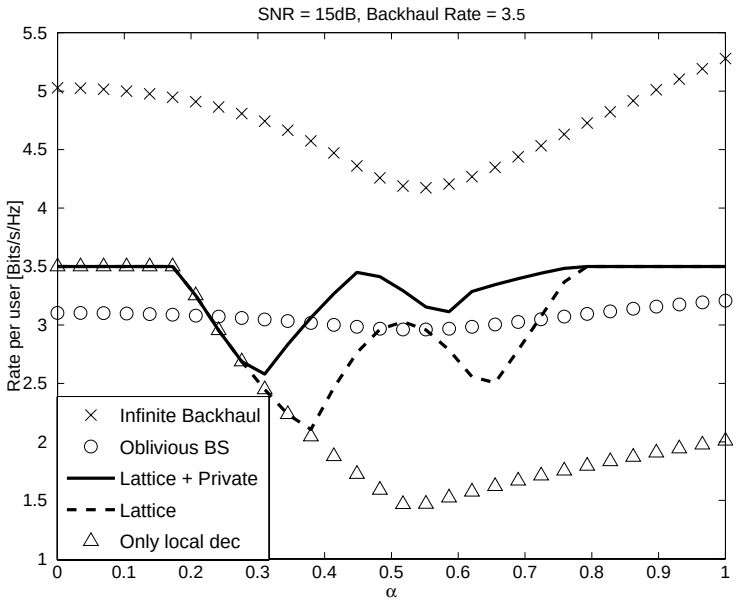

Fig. 5.2 Per-cell sum-rates versus the inter-cell gain α for unlimited-capacity backhaul, oblivious BSs, lattice-based schemes and informed BS with transmission of only local messages ("Only local dec") for a Gaussian Wyner model with $L = 1$, $K = 1$ user per cell, $C = 3.5$ (bits/s/Hz) and $P = 15\,\text{dB}$.

the analysis, deploying informed BSs leads to a performance gain for sufficiently small inter-cell interference α.

For the same channel with $P = 15\,\text{dB}$ and $C = 3.5$ (bits/s/Hz), we then add comparison with the lattice-based scheme mentioned above in Figure 5.2. It can be seen that the lattice-based scheme, and its enhancement (with "private" messages), outperforms the oblivious scheme for sufficiently low or high inter-cell interference α. The figure also shows the performance of a scheme with informed BSs where only local messages are transmitted ("Only local dec" in the figure), demonstrating the poor performance of local decoding for large inter-cell interference.

5.2 Local Backhaul between Adjacent BSs

In this section, we study the performance of MCP as afforded by the presence of local backhaul links between BSs (Figure 2.4(b)). Where not

stated otherwise, we focus on a Gaussian soft-handoff model with $L = 1$ and a large number of cells ($M \to \infty$). As will be discussed below, the soft-handoff model enables us to isolate key ideas and techniques, and to highlight the main conceptual differences with the scenario studied above of MCP via global (though possibly limited-capacity) connections to a CP. We also consider $K = 1$ or equivalently intra-cell TDMA.

Assume that each BS, say the m-th, is connected via backhaul links to the $(m + 1)$-th BS, as shown in Figure 5.3. Each backhaul link has capacity C (bits/s/Hz). Thanks to the backhaul links, each m-th BS can decode based not only on the locally received signal y_m but also on the information received on the backhaul from the BSs on the left. We will later consider briefly also the case in which the backhaul link between two adjacent BSs can also be used for communication from the $(m + 1)$-th BS to the m-th one.

The reason for emphasizing the backhaul connections toward the right has to do with the following fact. The m-th BS receives (with unitary gain) the signal of the MS that affects the $(m + 1)$-th BS. Therefore, thanks to the backhaul links with the given rightward direction, information about the interference can be communicated. Instead, a backhaul link in the opposite direction could only provide additional information about the signal, but not the interference. Therefore, a backhaul structure such as in Figure 5.3 has the potential to mitigate the interference and thus to overcome the limitations of SCP at high SNR, while a backhaul link in the opposite direction cannot be as effective in combating interference. Clearly, in a more realistic scenario, or even in the Wyner model, each BS measures a signal correlated with both the useful signal and the interference of the BSs to which it is

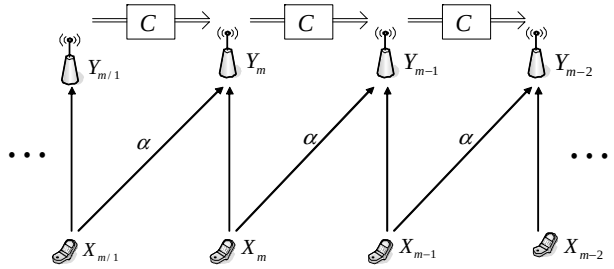

Fig. 5.3 A soft-handoff model with $L = 1$, $K = 1$, and local backhaul links to adjacent BSs.

connected. The soft-handoff model thus enables the aspect of interference management via local backhaul connections to be isolated.

Following the discussion above about the role of codebook information, we will consider both the scenario in which each BS is aware only of the local (same-cell) codebooks and that in which BSs can also exploit the knowledge of the codebooks used in adjacent (interfering) cells. Note that a fully oblivious cell as considered in the previous section does not apply to the scenario of interest here since decoding must be performed at the BS level. We will show below that, unlike the backhaul setting of the previous section, here major gains (even in terms of multiplexing gain) can be harnessed with enhanced codebook information. This points to the fact that the cost for signaling codebook information in the presence of a local backhaul topology is more likely to be worthwhile from a system design viewpoint than in the presence of global backhaul connectivity to a CP.

5.2.1 Performance Bounds and High-SNR Analysis

In this section, we discuss bounds on the per-cell rate achievable for the setting of Figure 5.3. Moreover, to emphasize the capability of different strategies to mitigate interference, we will also focus on the high-SNR regime and address performance in terms of multiplexing gain.

5.2.1.1 Upper Bound

A first observation made in [108] is that, irrespective of the codebook information available at the BSs, the per-cell rate can be upper bounded as

$$R(P,C) \leq \min\{R_{MCP}(P),\ R_U(P,C)\}, \tag{5.26}$$

where $R_{MCP}(P)$ is in Equation (3.15) and

$$R_U(P,C) = \frac{1}{2}\left[\log_2(1+P) + \log_2\left(1+\frac{P}{1+\alpha^2 P}\right)\right] + C. \tag{5.27}$$

As shown in [108], the latter upper bound is obtained by noticing that the backhaul cannot increase the sum-rate by more than C (bits/s/Hz), and that the sum-rate with $C = 0$ is upper bounded by

the sum-capacity of a Z-interference channel (which is obtained in [87] and given by $\log_2(1+P)+\log_2\big(1+P/(1+\alpha^2P)\big)$).

Similar to what was observed above for the setting of global backhaul connectivity, we note from Equation (5.26) that the multiplexing gain with fixed backhaul capacity C is limited to $1/2$, and is achieved by SCP with spatial reuse $F=2$. Instead, if we let the backhaul capacity scale with P as $\beta\log P$ with $0\leq\beta\leq 1$, then, from Equation (5.26), we see that the multiplexing gain is limited by $\min\{1,\, 1/2+\beta\}$. According to this bound, a scaling of $\frac{1}{2}\log P$ is thus potentially sufficient to achieve unitary multiplexing gain. This contrasts with global backhaul connectivity, where, it was shown above, it is necessary and sufficient for C to scale as $\log P$ to achieve such a result, and thus perfect interference mitigation in the high-SNR regime. We will discuss below to what extent the gains promised by Equation (5.27) are achievable.

5.2.1.2 Local Codebook Information

Consider now the per-cell achievable rate in the case in which each BS knows only the codebook employed by the same-cell MS. Following [108], we consider the following strategy based on successive decoding. Namely, starting with the first ($m=1$) BS and ending with the last ($m=M$) BS (recall Figure 5.3), we operate as follows. Once the m-th BS has decoded its local message, it can compress the decoded codeword via a rate-C quantization codebook and send the corresponding index over the backhaul link to the $(m+1)$-th BS. The latter then proceeds to decode its local message based on the received signal and the quantized codeword received over the backhaul link. The procedure repeats similarly for all the BSs. We refer to this scheme as Codeword Compression (CC). The rationale, as explained above, is that the compressed codeword provides information about the interference signal for the recipient BS. It is proved in [108] that this scheme achieves the per-cell sum-rate

$$R_{CC}(P,C)=\log_2\left(1+\frac{P}{1+\alpha^2P\left(1-\frac{1}{1+\frac{1}{1+\frac{\alpha^2P}{1+P}}\frac{1}{2^C-1}}\right)}\right). \qquad (5.28)$$

To obtain further insight into the interference mitigation capabilities of this scheme, we analyze its multiplexing gain. We note that CC leads to interference-limited performance (if not coupled with spatial reuse [108]), unless we let $C \to \infty$. In fact, with $C \to \infty$, perfect interference cancellation is possible, and $R_{CC} \to \log_2(1+P)$; therefore, in the regime $C \to \infty$, the CC scheme is not interference-limited and achieves a multiplexing gain of one.

We now remark that the CC scheme has the drawback of requiring a large delay for BSs with large indices due to the successive decoding procedure. It is therefore of interest to consider the performance of a scheme that has instead zero delay. This is accomplished by having each m-th BS compress and forward the received signal (instead of the decoded codeword). In this case, the backhaul quantization codebook is then used for the purpose of compressing the received signal. We refer to this transmission strategy as Signal Compression (SC). It was shown in [108] that the following rate is achievable with SC:

$$R_{SC}(P,C) = \log_2\left(1 + \frac{P}{1+\alpha^2 P\left(1 - \frac{1}{1+\alpha^2+(1+\sigma^2)/P}\right)}\right), \qquad (5.29)$$

with

$$\sigma^2 = \frac{P(1+\alpha^2) + 1 - \alpha^2 P^2(1+P(1+\alpha^2))}{2^C - 1}. \qquad (5.30)$$

This result emphasizes the importance of successive decoding for proper interference mitigation. In fact, by taking $C \to \infty$, we see that, unlike CC, the SC scheme is interference-limited. We recall that this is because the interference from other cells cannot be decoded and cancelled due to the presence of only local codebook information.

5.2.1.3 Enhanced Codebook Information

In this section, we investigate the performance advantage that can be accrued with enhanced codebook information at the BSs. Namely, we assume that the channel codebook employed by a given m-th MS is known not only at the local m-th BS but also at the $(m+1)$-th. As will be discussed in the sequel, this further information allows: (i) to

perform joint decoding of the local message and of (possibly part of) the interfering message (of the $(m+1)$-th MS) at the m-th BS; and (ii) to adopt more sophisticated quantization strategies on the backhaul link that exploit the side information available at the receiving BS regarding the channel codebook [20].

We first investigate a successive decoding strategy that differs from the CC technique described above in that: (i) joint decoding of the messages of the m-th and $(m+1)$-th is carried out at each m-th BS; and (ii) instead of compressing the decoded codeword, any m-th BS bins (compresses) directly the decided message, exploiting the fact that the channel codebook is known at the $(m+1)$-th BS [20]. To elaborate, assume at first that $R > C$. The $(m-1)$-th BS decodes the local message and sends the index of the bin in which the message falls to the m-th BS. The m-th BS then jointly decodes as explained above based on the received signal and the bin index received over the backhaul link. If $R \leq C$, then the entire message W_{m-1} can be sent over the backhaul link and the interference-free rate $R = \log(1+P)$ is achievable. More generally, we obtain the achievable per-cell rate

$$R_{DC}(P,C) = \min\left\{\begin{matrix}\log_2(1+P), & \log_2\left(1+\alpha^2 P\right)+C, \\ & \frac{1}{2}\log_2(1+(1+\alpha^2)P)+\frac{C}{2}\end{matrix}\right\}, \qquad (5.31)$$

where we refer to the scheme at hand as Decision Compression (DC).

Due to the joint decoding carried out at each BS, the DC scheme is noninterference-limited for any fixed value of C, even without employing spatial reuse $F = 2$. This is unlike the approaches discussed above, in which joint decoding was ruled out by the absence of information about the interfering MS's codebook. Moreover, assume now that the backhaul capacity scales with P as $\beta \log P$ with $0 < \beta \leq 1$. It is then easy to see from Equation (5.31) that the multiplexing gain of DC becomes $(\beta+1)/2$, which is generally less than the upper bound $\min\{1, 1/2+\beta\}$ derived above, but it achieves the maximum multiplexing gain of unity as long as the capacity C scales as $\log P$ $(\beta = 1)$. It is currently not known whether the multiplexing gain of $\min\{1, 1/2+\beta\}$ is achievable for any $0 < \beta \leq 1$. Finally, we remark that, with DC, for any fixed P, the interference-free performance $\log(1+P)$ is achieved

from (5.31) if

$$C \geq \max\{\log_2((1+P)/(1+\alpha^2 P)),\ \log_2((1+P)^2/(1+(1+\alpha^2)P))\}$$
$$= \log_2((1+P)^2/(1+(1+\alpha^2)P)). \tag{5.32}$$

Notice that this contrasts with the CC scheme where it was necessary to take $C \to \infty$ to achieve rate $\log(1+P)$.

We finally remark that reference [108] also considers the case in which the backhaul links can be used in both directions and obtains transmission strategies that generalize the ones presented above.

5.2.1.4 Numerical Results

Here, we further corroborate the results discussed above via numerical results. Figure 5.4 shows the derived achievable rates versus the SNR P for $\alpha^2 = 0.6$ and $C = 3$. The interference-limited behavior of the schemes based on local codebook information, namely CC and SC, is apparent, unlike DC. Moreover, the performance gain of CC over SC measures the advantages of allowing for some decoding delay.

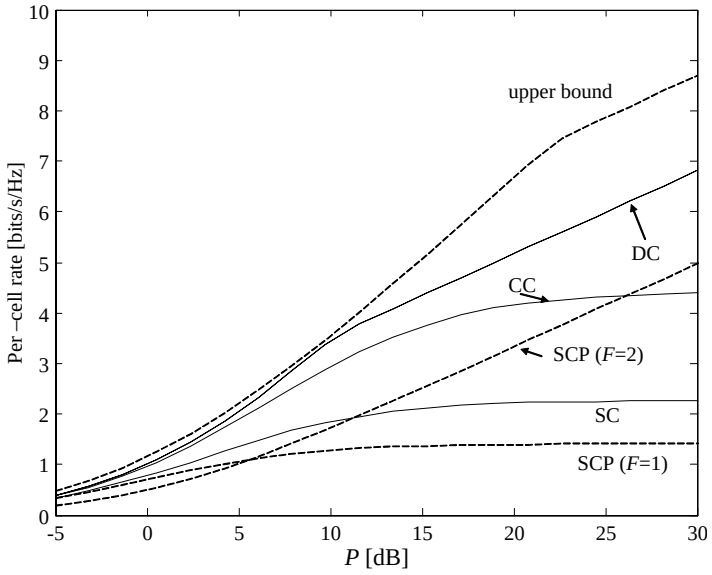

Fig. 5.4 Per-cell achievable rates versus the SNR P for a system with local backhaul connections ($\alpha^2 = 0.6$ and $C = 3$ bits/s/Hz).

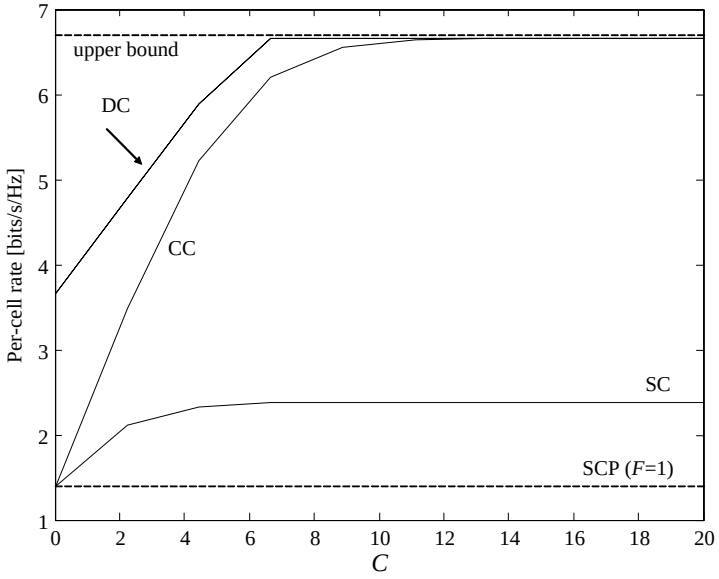

Fig. 5.5 Per-cell achievable rates versus the backhaul capacity C (bits/s/Hz) for a system with local backhaul connections ($\alpha^2 = 0.6$ and $P = 3\,\text{dB}$).

Figure 5.5 shows the achievable rates versus the capacity C for $\alpha^2 = 0.6$ and $P = 3\,\text{dB}$. It is confirmed that for sufficiently large C, both CC and DC-based schemes are able to achieve the interference-free rate $\log(1 + P) \simeq 1.59$, which approaches the upper bound (Equation (5.27)). Moreover, again, the advantages of allowing decoding delays are clear in the gain of the CC and DC techniques over SC.

5.3 Clustered Cooperation

We finally consider the clustered decoding scenario of Figure 2.4. In this scenario, clusters of BSs are connected to a different CP that is tasked with decoding the signals of a specified subset of MSs. Clusters of BSs can be overlapping or not, and no communication is possible between different CPs. Each CP must thus decode the assigned users based only on the signals received on the backhaul links from the BSs.

An example of clustered decoding is shown in Figure 5.6 for a Gaussian soft-handoff model with $L = 1$ and $K = 1$. In the example,

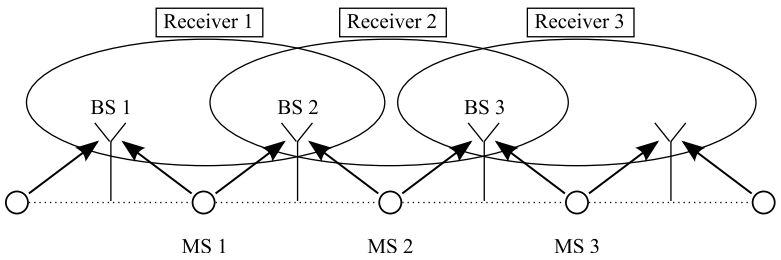

Fig. 5.6 A soft-handoff model with $L = 1$, $K = 1$, and clustered decoding.

clusters contain two cells with an overlap of one BS and each CP decodes a single MS. In particular, as shown in the figure, the first CP ("Receiver 1" in the figure) is connected to the BSs in the first and second cells and is tasked with decoding the MS in the first cell, the second CP is connected to the second and third cell and decodes the second MS, and so on. This particular clustering choice enables each CP to "see" the signal of the MS of interest as it is received from two BSs (plus inter-cell interference as noise). We refer to the previous section for further discussion on this point. Note that this scenario is in fact somewhat related to the setting of local backhaul studied above, since each CP has access to information from the BS in the same cell of the MS to be decoded and from the one on its right. The main difference is that, unlike the local backhaul case the CP cannot collect received signal information from BSs other than the ones to which it is directly connected. Recall that we assume unlimited-capacity backhaul links in this section.

In Section 5.3.2 we consider a more general setting, in which the CP tasked with decoding the m-th MS has access to the signal received, not only by the m-th BS, but also by j_r BSs on the right and j_ℓ BSs on the left.

5.3.1 Performance Bounds and High-SNR Analysis

We first discuss a lower bound on the achievable per-cell rate for the scenario in Figure 5.6 that was derived in [59]. This is obtained by adapting transmission strategies similar to the standard techniques proposed in [32] for transmission over interference channels. Specifically, as done

with informed BSs, each m-th transmitter (MS) splits its information bits into two independent parts, which are then encoded over independent codebooks and transmitted by superposition. The idea is that one part, referred to as "private", is to be decoded only by the m-th receiver (i.e., the BS in the same cell), while the other part, referred to as "public" is to be decoded by BSs $m-1$, m, and $m+1$. Note that BS $m-1$ does not "see" the m-th MS directly but only through the signal obtained via the backhaul by the m-th BS. The private message is sent with power βP, while the common message is transmitted with the remaining power $(1-\beta)P$ for a power split parameter $0 \leq \beta \leq 1$. Expression of the corresponding achievable rate can be found in [59].

Reference [59] also obtains the following upper bound on the achievable rates with clustered decoding by adapting the analysis in [47, Theorem 2] and [97]:

$$R(P) \leq \frac{1}{3}\log_2\Bigg[(1+(1+\alpha+\alpha^2)P) \cdot \frac{(1+(1-\alpha+\alpha^2)P)(1+(1+\alpha^2)P)^2}{(1+P)(1+\alpha^2 P)}\Bigg]. \quad (5.33)$$

Figure 5.7 shows the performance obtained with clustered cooperation in terms of the scheme discussed above (labeled with "lower" in the figure) and of the upper bound (Equation (5.33)). For reference, the per-cell sum-rate (Equation (3.15)) attained with ideal MCP is also shown, along with the per-cell sum-rate achievable with local backhaul connections via the DC scheme. We set $\alpha^2 = 0.6$. Recall that the DC scheme only requires knowledge of the codebook used by one adjacent BS, whereas the decoder for each cluster here needs to have access to the codebook of an additional BS. The figure shows that the local backhaul architecture is potentially advantageous as long as the capacity C is large enough. In fact, as explained in the previous section, by leveraging local backhaul connections, one can propagate information from remote BSs, which can be beneficial for interference management.

5.3.2 High-SNR Analysis

Some further insight can be obtained by studying the high-SNR regime. In this regard, a simple observation is that the multiplexing gain should

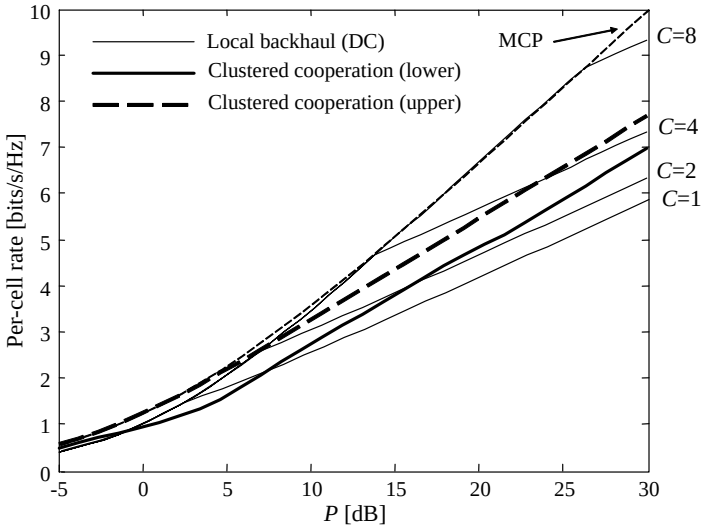

Fig. 5.7 Per-cell sum-rate of clustered cooperation, of ideal MCP (Equation (3.15)), and of local backhaul connections via the DC scheme with various backhaul capacities C ($L = 1$, $K = 1$, $\alpha^2 = 0.6$).

be at least 2/3. Indeed, it is enough to shut off one MS out of every three at any time. This way, the two BSs neighboring the silent MS are able to receive at least one signal free of interference. Since two out of three MSs can transmit without any interference, a multiplexing gain of 2/3 is attained with this simple scheduling scheme. The upper bound (Equation (5.33)) shows that this is also the maximum multiplexing gain, proving the high-SNR optimality of the scheduling strategy mentioned above in terms of multiplexing gain.

We now turn to the more general model in which each CP is connected with j_ℓ BSs on the left and j_r BSs on the right. Reference [58] obtains the maximum multiplexing gain of this setting as

$$\mathcal{S}_\infty = \frac{j_\ell + j_r + 1}{j_\ell + j_r + 2}. \tag{5.34}$$

Note that this result recovers the conclusion of the discussion above for $j_r = 1$ and $j_\ell = 0$. The result (Equation (5.34)) shows that with local backhaul connections the multiplexing gain is always less than one, but it is larger than 1/2 as achievable with SCP and spatial reuse $F = 2$ (see Section 3.1.1.1). Moreover, this shows that (for the soft-handoff

model), left or right side information has the same impact on the multiplexing gain. Finally, multiplexing gain (Equation (5.34)) is shown to be achieved by *successive interference cancellation* at the BSs, where BSs exchange information about the decoded signals, similar to the CC technique discussed above. Therefore, this scheme requires knowledge of the codebooks used in adjacent cells by the BSs.

5.4 Summary

In this section, we have analyzed the impact of the backhaul architecture and of limitations on the available backhaul capacity on the performance of MCP for the uplink. We have first considered global, but finite-capacity, backhaul connectivity to a CP. In this case, it was shown that the approach of deploying oblivious BSs, which merely act by compressing the received signal, has the potential to achieve a large fraction of the gains promised by MCP with reasonable backhaul capacities, namely of the order of the operating per-cell sum-rate. In particular, the analysis has shown that the performance loss of oblivious BSs with respect to ideal MCP can be quantified by an SNR degradation that vanishes exponentially with the backhaul capacity C. It was also pointed out that performance can be further improved by leveraging some knowledge of the MSs' codebooks, especially through structured coding strategies.

We have then considered different types of backhaul architectures in which the BSs are either connected to one another, or are connected to a CP that serves only a cluster of cells. In both cases, unlike the backhaul architecture considered above, the critical role of codebook information at the BSs, or at the cluster CP, was pointed out, especially in the high-SNR regime. Moreover, it was found that, if the backhaul capacity is sufficiently large, the capability of BSs to propagate interference information through inter-BSs backhaul links makes it possible to achieve rates, especially in the high-SNR regime, that cannot be attained by clustered decoding.

6

Constrained Multi-Cell Processing: Downlink

In this section, we turn to the downlink and address the performance of MCP in the presence of limitations in the backhaul links. As in the previous section, we will consider first limited-capacity backhaul to a CP, then finite-capacity links between adjacent BSs and finally clustered cooperation (see Section 2.3). We focus on Gaussian models, leaving the impact of fading to future work.

6.1 Finite-Capacity Backhaul to a Central Processor

Let us start by considering the setting in Figure 2.4(a). Here, the CP, which has the available information messages intended for the MSs in all cells, is connected to each BS via a backhaul link with capacity C (bits/s/Hz). We focus on the soft-handoff model with $L = 1$. The per-cell sum-rate capacity $R_{MCP}(P)$ for the ideal case of unlimited C was discussed in Section 3.2.2 and is given by Equation (3.15), as obtained via duality arguments (see Appendix B). It is recalled that the capacity is achieved by means of DPC (see Appendix A). With an arbitrary value for C, considerations similar to the uplink show that the per-cell

sum-rate is limited, as for the uplink, by

$$R(P,C) \leq R_{UB}(P,C) = \min\{C, R_{MCP}(P)\}. \tag{6.1}$$

Therefore, the same arguments as used in Section 5.1 apply, demonstrating that the optimal multiplexing gain of one can be achieved only if C scales as $\log P$.

We now study how close one can get to the upper bound (Equation (6.1)). As for the uplink, we do this by considering separately two types of MCP strategies. In the first, the BSs are assumed to have no *encoding* capabilities. This enables implementation with "light" BSs, that only need to have a radio transceiver and minimal baseband processing. The BSs are also not required to be informed about the codebooks used by the CP to communicate with the MSs. Following the nomenclature in the previous section, we refer to this scenario as that with oblivious BSs. We then investigate the performance with BSs that have some encoding capabilities. Again, we refer to this scenario as that with informed BSs, since in this case the BSs must be aware of the codebooks employed in other cells.

6.1.1 Oblivious Base Stations

We start by considering the rate achievable with oblivious BSs. For this scenario, reference [107] proposes the following scheme, assuming intra-cell TDMA ($K = 1$). The CP performs encoding using DPC *as if* the backhaul were of unlimited capacity (see Section 3.2.2). This operation produces one codeword per BS. To convey such codewords to the BSs over the finite-capacity backhaul links, each is quantized to C (bits/s/Hz). The index of the compressed codeword is then sent to the corresponding BS. Each BS simply forwards the received compressed codeword, after appropriate scaling so as to satisfy the power constraint. In designing DPC at the CP, the CP accounts for the fact that the BSs forward inevitably also quantization noise, which leads to an SNR loss. The following per-cell rate is proved to be achievable:

$$R_{obl}(P,C) = \log_2\left(\frac{1+(1+\alpha^2)\tilde{P}+\sqrt{1+2(1+\alpha^2)\tilde{P}+(1-\alpha^2)^2\tilde{P}^2}}{2}\right), \tag{6.2}$$

where

$$\tilde{P} = \frac{P}{\frac{1+(1+\alpha^2)P}{2^C-1}+1}. \tag{6.3}$$

Comparing the per-cell rate Equation (6.2) with Equation (3.15), which is the maximum achievable with unlimited backhaul links, reveals that the performance loss due to a finite C translates into an SNR degradation of a factor $\frac{1+(1+\alpha^2)P}{2^C-1}+1$. Note that for $C \to \infty$, this SNR loss vanishes, and the rate (Equation (6.2)) reduces to Equation (3.15). Recall that similar conclusions were obtained also for the uplink in Section 5.1.2. The parallel between the two situations is further discussed below by focusing on the extreme-SNR analysis.

Similar to the uplink, a coding strategy based on structured codes can also be devised, as proposed in [37].

6.1.1.1 Extreme-SNR Analysis

Let us start with the high-SNR regime. Substituting $C = \beta \log_2 P$ in Equation (6.2), 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 scale as $\log_2 P$, which is optimal according to our discussion about the upper bound.

In the low-SNR regime, characterization of the performance is given by Equations (5.8) and (5.9), exactly as for the uplink. This result confirms the parallels between the considered schemes for uplink and downlink. It is recalled that Equation (5.8) implies that the SNR loss with respect to the ideal performance with unlimited backhaul in the low-SNR regime is given by $1/(1-2^{-C})$.

6.1.2 Informed Base Stations

In the previous section, it was shown that, as for the uplink, oblivious BSs are able to achieve capacity if the backhaul capacity C is large enough. However, if a more efficient use of the backhaul is necessary, deploying informed BSs becomes mandatory. Moreover, if the inter-cell interference α is small, the SNR loss due to the oblivious BS implementation overcomes the gains that are possible with MCP,

and thus informed BSs become especially beneficial. In this section, we study these issues by studying the performance of a scheme proposed in [107].

We observed that design of a transmission scheme with informed BSs is less straightforward than for the uplink. In fact, for the downlink adapting the encoding strategy at a given BS based on the encoding operations (codebooks) of another has a chain effect on all other BSs whose operation depends on the codebook of the BS at hand. This issue is further explored in [107], where a number of additional techniques are proposed and evaluated. Here we bypass this issue by considering only information about the codebook intended for the same-cell BS. We will see that the simple technique considered here addresses, at least partly, the issues mentioned above.

The idea is the following. Consider a soft-handoff model with $L = 1$. If a BS knew the signal transmitted by the BS that interferes with the MS in the same cell, it could use DPC and cancel such interference (see Appendix B). In the considered scheme, the CP thus sends to each BS a quantized version of the codeword of the interfering BS along with the message intended for the same-cell BS. Note that the CP can precalculate all the codewords transmitted by all BSs. Each BS can then perform DPC over the quantized interfering signal. Reference [107] shows that the following per-cell rate is achievable:

$$R_{\text{inf}}(P,C) = \begin{cases} C & \text{if } C \le \log_2\left(1 + \frac{P}{1+\alpha^2 P}\right) \\ R'_{\text{inf}}(P,C) & \text{otherwise} \end{cases} \tag{6.4}$$

where

$$R'_{\text{inf}}(P,C) = \log_2\left(1 - \frac{2^C}{\alpha^2 P} + \sqrt{1 + \frac{2^{C+1}}{\alpha^2}\left(2 + \frac{1}{P}\right) + \frac{2^{2C}}{\alpha^4 P^2}}\right) - 1 \tag{6.5}$$

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

A first observation is that if C is large, this scheme achieves $\log_2(1 + P)$, thus fully mitigating interference, but failing to harness the array gains achievable with MCP. However, following the discussion above, if $C \le \log_2(1 + P/(1 + \alpha^2 P))$ or $\alpha = 0$, then the scheme is optimal. In these cases, in fact, the upper bound (Equation (6.1))

can be achieved simply by having the CP forward only local message information to the BS and each MS decodes by treating inter-cell interference (if present) as noise. However, the scaling law of C required to achieve a multiplexing gain of one can be seen to be $2\log P$, which is double what is required by the oblivious strategy (see more on this in the next section).

6.1.3 Numerical Results

Figures 6.1 and 6.2 show the per-cell rates achievable by the oblivious implementation of MCP and by the simple strategy based on informed BSs, discussed above versus the backhaul capacity C and the inter-cell gain α^2, respectively, for $P = 10\,\text{dB}$. Specifically, for $\alpha = 1$, Figure 6.1 suggests that, just as we concluded for the uplink, a backhaul capacity that is 2–3 times the per-cell transmission rate is enough to achieve performance close to the upper bound. Moreover, for this setting, the

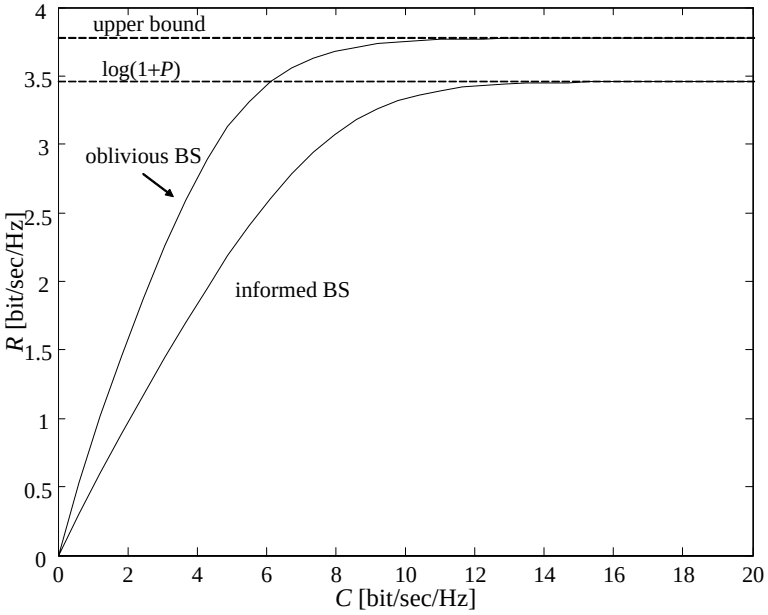

Fig. 6.1 Per-cell achievable rates for the downlink of a soft-handoff model with the oblivious and informed implementations of MCP versus the backhaul capacity C (L =1, $P = 10\,\text{dB}$, $\alpha = 1$).

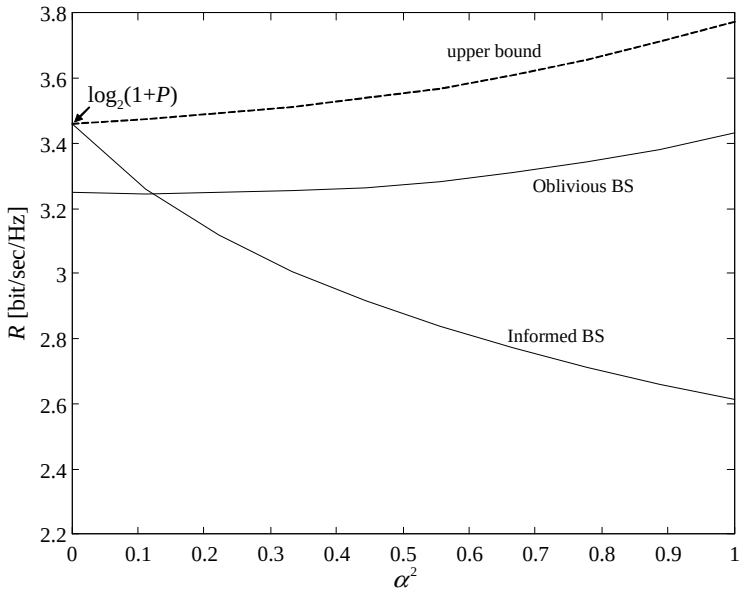

Fig. 6.2 Per-cell achievable rates for the downlink of a soft-handoff model with the oblivious and informed implementations of MCP versus the inter-cell gain α^2 (L =1, $P = 10\,\text{dB}$, $C = 6$ (bits/s/Hz)).

inter-cell factor α is very large, so that MCP, even if implemented with oblivious BSs, outperforms the simple informed technique also for small backhaul capacity C. Further insight into the comparison between oblivious and informed implementation can be gathered from Figure 6.2, which for $C = 6$ (bits/s/Hz), shows that for small inter-cell interference gains, an informed implementation is beneficial.

6.2 Local Backhaul between Adjacent BSs

In this section, we study the performance of MCP as afforded by the presence of local backhaul links between adjacent BSs as shown in Figure 2.4(b). We focus on a Gaussian soft-handoff model with $L = 1$ and $K = 1$. Our main aim here, following the discussion above, is to illustrate some of the main differences between designing transmission strategies for the uplink and downlink in this scenario. In particular, we recall that in Section 5.2.1.3, a decoding strategy was devised for the counterpart uplink model (Figure 5.3) that, when backhaul

capacity scales with P as $\log P$, achieves the maximum multiplexing gain of one. This uplink strategy is based on successive decoding and hinges on the availability at each m-th BS of the codebook used also in the $(m-1)$-th cell (which creates interference). We now discuss the performance of a related strategy based on successive *encoding* based on the proposal in [107].

A first observation is that in the model at hand, while interference from the m-th cell affects the $(m+1)$-th in the uplink, the opposite is true in the downlink, so that the signal of the $(m+1)$-th BS affects the m-th MS. Therefore, if the goal is to successively remove interference, encoding should be performed from the last cell ($m = M$) toward the first ($m = 1$), in the opposite order with respect to the successive decoding technique used in the uplink. A similar phenomenon is observed also in the context of uplink–downlink duality as explained in Appendix B. Therefore, we assume here that the backhaul links are operated from the $(m+1)$-th to the m-th BS. Moreover, as discussed above, while in the uplink a change in the decoding strategy at one BS does not affect the others, a modification in the encoding function of one BS does affect the performance in other cells via interference. This complicates the design and has implications on the codebook information necessary at each BS.

To elaborate, consider the following scheme. The M-th BS, not suffering from any interference, transmits with a standard codebook to its MS. The $(M-1)$-th BS receives the message sent by the M-th BS over the backhaul, and, based on the knowledge of the codebook of the M-th BS, can use DPC and obtain an interference-free rate. Now, in order for the $(M-2)$-th BS to know the interference affecting its MS, both the message of BS $M-1$ and BS M must be received over the backhaul, and the corresponding codebooks must also be known. This is because the signal sent by the interfering $(M-1)$-th BS depends also on the message of the M-th. In general, for DPC at the m-th BS to be successful, the messages of BSs $m+1, \ldots, M$ must be received over the backhaul and the corresponding codebooks known. This should be contrasted with the successive decoding strategy of Section 5.2.1.3, in which it was sufficient to transmit only one message over the backhaul and to know only one additional codebook.

In order not to create a bottleneck on the backhaul link, we assume that the scheme described above can be implemented only in J successive cells. This way, each BS should only know the encoding functions, and receive the messages, of the $J-1$ BSs preceding it according to the encoding order. As pointed out in [53], this can be implemented by "turning off" one out of every $(J+1)$ BS and consider the clusters of J BSs in between silent BSs. The following rate can thereby be achieved:

$$R(P,C) = \max_{J} \min \left\{ \frac{C}{J-1}, \left(\frac{J}{J+1} \right) \log_2(1+P) \right\}. \tag{6.6}$$

Note that the rates in all cells can be made uniform by time sharing of different scheduling patterns [107]. Moreover, in Equation (6.6), the first term is due to the requirements for communication over the backhaul, while the second accounts for the fact that, at any given time, only J out of $J+1$ cells are active.

We now address the requirements of this scheme in terms of backhaul in order to achieve the optimal multiplexing gain of one. It can be seen that, for any fixed cluster size J, the maximum multiplexing gain is $J/(J+1) < 1$, irrespective of the scaling of C. In particular, achieving this rate scaling requires the backhaul capacity C to grow as $(J-1)\log_2 P$. Therefore, a unitary multiplexing gain can only be approached at the cost of increasing the backhaul capacity much faster than required for the uplink of the same scenario.

The discussion above points to the inefficiency, at least in the regime of high SNR, of successive encoding based on "hard" information sent over the backhaul links. A related strategy based on transmission of "soft" information can hence be devised similar to the development in Section 6.1.2. In this strategy, each m-th BS sends to the $(m-1)$-th a compressed version of the transmitted codeword over the backhaul. This strategy, following [107], can be proved to achieve

$$R(P,C) = \log_2 \left(1 + \frac{P}{1 + \frac{\alpha^2 P}{2^C}} \right). \tag{6.7}$$

It can thus be seen that the optimal multiplexing gain of one is achieved with the scaling of $\log P$, as for the uplink.

6.3 Clustered Cooperation

We finally consider the clustered decoding scenario of Figure 2.4 and focus again on the Gaussian soft-handoff model with $L = 1$ and $K = 1$. We assume, similar to Section 5.3, that each BS can encode based not only on the local message but also of the messages of j_ℓ cells on the left and j_r cells on the right. We concentrate on the high-SNR regime and ask whether the result (Equation (5.34)) obtained for the uplink still holds for the downlink. Given the differences between uplink and downlink highlighted above, this is not immediately clear.

References [52, 57] show that the multiplexing gain (Equation (5.34)), reported here for convenience,

$$\mathcal{S}_\infty = \frac{j_\ell + j_r + 1}{j_\ell + j_r + 2}, \tag{6.8}$$

is indeed achievable (and also optimal). This is shown by exploiting a technique that operates in a similar fashion as the one presented in the previous section. In particular, the idea is to silence some BSs, thereby splitting the network into noninterfering subnetworks which can be treated separately. In each subnetwork, some of the BSs use simple single-user encoding schemes and some of the transmitters use DPC. The multiplexing-gain per user in Equation (6.8) shows an equivalence between having message of BSs on the right or on the left.

6.4 Summary

In this section, we have studied the impact of backhaul architecture and of limitations on the backhaul capacity on the downlink by focusing on Gaussian channels. For Gaussian channels, Section 2 demonstrated that, thanks to duality, the maximum per-cell sum-rate achievable in the downlink is the same as that in the uplink. This section has lent evidence to the fact that the same fact holds to some extent also in the presence of limitations in the backhaul, although a full theoretical understanding on this point is still lacking. To elaborate, considering at first global, but finite-capacity, connectivity from BSs to the CP, we have seen that an oblivious deployment of BSs has a similar effect as in the uplink. It is, in fact, able to approximate the ideal performance with

an SNR loss that decays exponentially with the backhaul capacity C. With local backhaul connections between BSs, it was argued that, in order to perform interference mitigation as effective (in the high-SNR regime) as for the uplink, BSs should exchange soft information about the transmitted signals, rather than hard information about the transmitted messages. The latter approach was instead seen to provide excellent performance in the uplink in the previous section. This asymmetry is due to the fact that, in the downlink, the need to control interference via precoding ties the transmission strategies of all the BSs. Therefore, the interference experienced at an MS does not depend solely on the messages communicated to the MSs in the adjacent cells, but also on those of cells further apart, which are not directly interfering with the MS at hand. Instead, in the uplink, interference depends only on the message transmitted by the MSs in the adjacent cells (since the MSs do not cooperate). Finally, with clustered BS cooperation, an uplink–downlink duality was shown to hold in the high-SNR regime in terms of multiplexing gain.

7

Dedicated Relays

In this section, we address the performance of cooperation in cellular networks in the presence of dedicated relay stations, following the uplink model discussed in Section 2.4.1 (see Figure 2.5). We consider the performance of two classes of strategies. The first, referred to as "nonregenerative" relaying, includes strategies in which the RSs do not decode the signal transmitted by the MSs, but instead forward "soft" information toward the BSs. In this class, we will consider techniques based on Amplify-and-Forward (AF) and Compress-and-Forward (CF) relaying techniques. The second class, referred to as "regenerative" relaying, is such that the RSs decode, at least partially, the signal received from the MSs before forwarding. This second class includes techniques based on Decode-and-Forward (DF) relaying. For an introduction to the protocols DF, AF, and CF, we refer the reader to [48].

Central to our discussion will be the interplay between the relaying technique and the implementation of SCP or MCP decoding. As will be concluded, the design choice of the best relaying technique depends critically on whether SCP or MCP is implemented. In order to simplify the analysis, when considering MCP we will assume unlimited backhaul links to a CP. The effect of limitations in the backhaul can be in

principle studied by combining the analysis below with that presented in Section 5. We finally remark that, while we assume full-duplex relays, the analysis can be easily adapted to half-duplex relays (see [102]).

7.1 Upper Bound

We first derive an upper bound on the achievable per-cell rate. We do so by leveraging cut-set arguments as already done in previous sections, see, e.g., Section 5.1. To this end, observe that the per-cell sum-rate can never be larger than the per-cell sum-capacity of the first hop (MSs-to-RSs) for the case in which the RSs are able to fully cooperate for the decoding. In other words, the achievable rate cannot be better than that of an MCP system in which the CP is directly connected to all relays. Using the definition (cf. Equation (3.14))

$$R_{\mathrm{w}}(a,b,\rho) = \int_0^1 \log_2\left(1+\rho H(\theta)^2\right) d\theta \ , \tag{7.1}$$

where $H(\theta) = b + 2a\cos(2\pi\theta)$, this implies the following upper bound on the achievable per-cell rate:

$$R(P) \leq R_{\mathrm{w}}(\alpha,\beta,\rho_1), \tag{7.2}$$

where $\rho_1 = P/\sigma_1^2$, and σ_1^2 is the noise power for the first hop. Moreover, the per-cell sum-rate cannot be greater than that achievable on the second hop (RSs-to-BSs) for the case in which the RSs are fully cooperative for transmission to the BSs. Note that this statement would not be correct if one did not allow cooperation at the RS level. This is because, as can be seen via cut-set arguments, the relays can potentially correlate their transmissions thanks to the (correlated) signals received from the MSs. This argument leads to the upper bound

$$R(P) \leq R_{\mathrm{w}}^{\mathrm{wf}}(\eta,\gamma,\rho_2), \tag{7.3}$$

where we define

$$R_{\mathrm{w}}^{\mathrm{wf}}(\eta,\gamma,\rho_2) = \max \int_0^1 \log_2\left(1+\rho_2 S(\theta) H(\theta)^2\right) d\theta, \tag{7.4}$$

with $H(\theta) = \eta + 2\gamma\cos(2\pi\theta)$ and $\rho_2 = Q/\sigma_2^2$, and where the maximization is taken over all functions $S(\theta)$ such that $\int_0^1 S(\theta)d\theta = 1$. The rate

(Equation (7.4)) can be seen, from the discussion in Section 3.1.2, to correspond to the maximum per-cell capacity of a Gaussian Wyner model with channel gains as in the second hop and with full transmitter cooperation. The optimization of Equation (7.4) is well-known to be given by the so-called waterfilling solution, which leads to

$$R_{\mathrm{w}}^{\mathrm{wf}}(\eta,\gamma,\rho_2) = \int_0^1 \log_2\left(1+\rho_2\left(\nu-\frac{1}{H(\theta)^2}\right)^+ H(\theta)\right)d\theta$$
$$\text{s.t.} \int_0^1 \left(\nu-\frac{1}{H(\theta)^2}\right)^+ d\theta = 1. \tag{7.5}$$

Equation (7.5) can be expressed in a closed-form expression for a certain range of the channel parameters (see [105]).

Overall, we obtain that the per-cell sum-rate is upper bounded as

$$R(P) \leq R_{UB}(P) = \min\{R_{\mathrm{w}}(\alpha,\beta,\rho_1), R_{\mathrm{w}}^{\mathrm{wf}}(\eta,\gamma,\rho_2)\}. \tag{7.6}$$

We remark that since both arguments of Equation (7.6) increase with SNR it is easily verified that $R_{\mathrm{UB}} \underset{\rho_1\to\infty}{\to} R_{\mathrm{w}}^{\mathrm{wf}}(\eta,\gamma,\rho_2)$ and that $R_{\mathrm{UB}} \underset{\rho_2\to\infty}{\to} R_{\mathrm{w}}(\alpha,\beta,\rho_1)$. This means that if the SNR in the first hop, ρ_1, is large, the performance is limited by the upper bound on the per-cell sum-rate of the second hop, and vice versa when $\rho_2 \to \infty$.

7.2 Nonregenerative Relaying

In this section, we study the performance of nonregenerative techniques, namely CF and AF, for the system presented in Section 2.4.1. With nonregenerative strategies, the RSs need not know the codebooks used by any MS. Therefore, using the nomenclature employed in previous sections, we can say that the RSs are oblivious. This section follows mostly reference [114]. Throughout, we assume a large number of cells ($M \to \infty$) and follow the model presented in Section 2.4.1.

7.2.1 Amplify-and-Forward (AF)

In this section, we assess the performance of an AF-based scheme, whereby the RSs amplify and forward the received signal with an integer delay of $\lambda \geq 1$ symbols. More specifically, with $(\cdot)^{(1)}$, $(\cdot)^{(2)}$ denoting

the association to the first (MS-to-RS) and second (RS-to-BS) hops, as in Section 2.4.1, the m-th RS transmits $x^{(2)}_{m,t} = g y^{(1)}_t$ with $y^{(1)}_t$ given in Equation (2.9) and $g \geq 0$ being the amplification gain. This is selected to satisfy the average power limitation

$$\sigma_r^2(g) \triangleq \mathbb{E}\{|x^{(2)}_{m,t}|^2\} \leq Q,$$

which satisfies the per-block power constraint (Equation (2.12)). Note that, for simplicity of notation, here we include the time index $t \in [1, n]$ in the subscript, unlike the notation used so far. Using Equation (2.9), the signal transmitted by the m-th RS can thus also be written as

$$\begin{aligned} x^{(2)}_{m,t} = g \cdot (&\beta x^{(1)}_{m,t-\lambda} + \alpha x^{(1)}_{m-1,t-\lambda} + \alpha x^{(1)}_{m+1,t-\lambda} \\ &+ \mu x^{(2)}_{m-1,t-\lambda} + \mu x^{(2)}_{m+1,t-\lambda} + z^{(1)}_{m,t-\lambda}). \end{aligned} \tag{7.7}$$

We now study the performance of this scheme with MCP or SCP with reuse $F = 1$.

7.2.1.1 Multi-cell Processing

In this section, we assume that the signals received at all BSs are jointly decoded by the CP. The CP is connected to the BSs via ideal backhaul links and is assumed to be aware of the codebooks of all the MSs. It is noted that using similar arguments as in [132], it can be shown that in this setup an intra-cell TDMA protocol is optimal.

Extending the LTI interpretation of the Wyner model given in Section 3.1.2 (see Figure 3.1), the received signal $y^{(1)}_{n,t}$ at the RSs (Equation (2.9)) and the received signal $y^{(2)}_{n,t}$ at the BSs (Equation (2.10)) can be interpreted, given Equation (7.7), as a *two-dimensional* LTI system with input $x^{(1)}_{m,t}$. Note that this is unlike the LTI system used for the Wyner model in Figure 3.1 in which the single dimension corresponds to the BS index. Here in fact the two dimensions correspond to the BS index (m) and the time index (t). We emphasize that the need to account for the temporal dimension follows from the feedback introduced by the fact that the relays overhear each other through the channel gain μ.

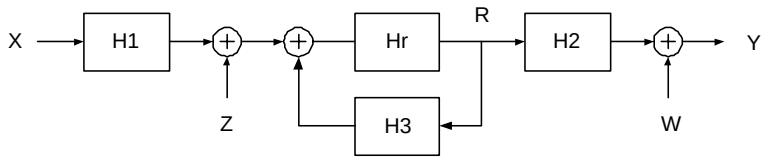

Fig. 7.1 Equivalent two-dimensional LTI channel.

The block diagram of this equivalent two-dimensional LTI system is depicted in Figure 7.1 where the impulse responses are given by

$$\begin{aligned} h_{1m,t} &= \delta_t(\alpha\delta_{m-1} + \beta\delta_m + \alpha\delta_{m+1}) \\ h_{2m,t} &= \delta_t(\eta\delta_{m-1} + \gamma\delta_m + \eta\delta_{m+1}) \\ h_{rm,t} &= g\delta_{t-\lambda}\delta_m \quad \text{and} \\ h_{3m,t} &= \mu\delta_t(\delta_{m-1} + \delta_{m+1}), \end{aligned} \tag{7.8}$$

where we recall that δ_n denotes the Kronecker delta function. The corresponding frequency responses are given by

$$\begin{aligned} H_1(\theta,\varphi) &= \beta + 2\alpha\cos 2\pi\theta \\ H_2(\theta,\varphi) &= \gamma + 2\eta\cos 2\pi\theta \\ H_r(\theta,\varphi) &= ge^{-j2\pi\lambda\varphi} \quad \text{and} \\ H_3(\theta,\varphi) &= 2\mu\cos 2\pi\theta. \end{aligned} \tag{7.9}$$

where θ and φ are the spatial frequency (as in Section 3.1.2) and the temporal frequency, respectively. Since the noise processes $z^{(1)}$ and $z^{(2)}$ are zero mean i.i.d. complex Gaussian and statistically independent of each other and of the input signal $x^{(1)}$, the received signals (Equation (2.10)) are at the BSs can be expressed as

$$y^{(2)}_{m,n} = s_{m,n} + q_{m,n}, \tag{7.10}$$

where $s_{m,n}$ and $q_{m,n}$ are zero mean wide sense stationary (WSS) statistically independent processes representing the useful part of the signal and the noise, respectively. Therefore, using a two-dimensional formulation of Szegö's theorem [132], the per-cell rate achievable by the

considered scheme is given by

$$R_{\mathrm{AF-MCP}}(P) = \int_0^1 \int_0^1 \log_2\left(1 + \frac{\mathcal{S}_S(\theta,\varphi)}{\mathcal{S}_N(\theta,\varphi)}\right) d\varphi\, d\theta, \tag{7.11}$$

where $\mathcal{S}_S(\theta,\varphi)$ and $\mathcal{S}_N(\theta,\varphi)$ are the power spectral density of S and N, respectively. These can be obtained as

$$\mathcal{S}_S(\theta,\varphi) = P\,|H_S(\theta,\varphi)|^2 = P\left|\frac{H_1 H_r H_2}{1 - H_r H_3}\right|^2, \tag{7.12}$$

and

$$\mathcal{S}_N(\theta,\varphi) = \sigma_1^2\,|H_N(\theta,\varphi)|^2 + \sigma_2^2 = \sigma_1^2\left|\frac{H_r H_2}{1 - H_r H_3}\right|^2 + \sigma_2^2, \tag{7.13}$$

where the transfer functions H_1, H_2, H_r, and H_3 are defined in Equation (7.9).

Analytical evaluation of Equation (7.11) leads to [114, Appendix A]

$$R_{AF-MCP}(P) = \int_0^1 \log_2\left(\frac{A + B + \sqrt{(A+B)^2 - C^2}}{B + \sqrt{B^2 - C^2}}\right) d\theta, \tag{7.14}$$

where

$$\begin{aligned} A &\triangleq Pg^2(\beta + 2\alpha\cos 2\pi\theta)^2(\gamma + 2\eta\cos 2\pi\theta)^2 \\ B &\triangleq \sigma_1^2 g^2(\gamma + 2\eta\cos 2\pi\theta)^2 + \sigma_2^2(1 + 4g^2\mu^2\cos^2 2\pi\theta) \quad \text{and} \\ C &\triangleq 4\sigma_2^2 g\mu\cos 2\pi\theta. \end{aligned}$$

Furthermore, the optimal relay gain g_{o} is the unique solution to the equation $\sigma_r^2(g) = Q$ where

$$\sigma_r^2(g) = \frac{(P\beta^2 + \sigma_1^2)g^2}{\sqrt{1 - (2\mu g)^4}} + \frac{4P\alpha^2 g^2}{\sqrt{1 - (2\mu g)^2} + 1 - (2\mu g)^2}. \tag{7.15}$$

It can be proved, as implied by the result above, that the optimal relay gain g_o is such that the RSs use their full power Q, and that $g_{\mathrm{o}} \underset{Q\to\infty}{\longrightarrow} 1/(2\mu)$. Moreover, it can be seen that the performance of AF is not interference limited, achieving the full multiplexing gain of one, and that it is independent of the actual RS delay λ.

Some further remarks are in order. In the case where the RSs employ directional antennas pointed toward their local BSs and they are in

sufficiently elevated positions to avoid multipath (see also discussions in [103, 104]), there is no inter-relay interference and $\mu = 0$. In this case, the general expression (Equation (7.14)) reduces to

$$R_{\mathrm{AF-MCP-DA}}(P) = \int_0^1 \log_2\left(1+\frac{Pg^2(\beta+2\alpha\cos 2\pi\theta)^2(\gamma+2\eta\cos 2\pi\theta)^2}{\sigma_1^2 g^2(\gamma+2\eta\cos 2\pi\theta)^2+\sigma_2^2}\right)d\theta. \quad (7.16)$$

In addition, by setting $\mu = 0$ in Equation (7.15) we obtain the optimal RS amplification gain as

$$g_{\mathrm{o}}^2 = \frac{Q}{P(\beta^2+2\alpha^2)+\sigma_1^2}. \quad (7.17)$$

Finally, we remark that, in the case of half-duplex operation, the RSs are not capable of simultaneous receive–transmit operation. Accordingly, the time is divided into equal slots: during odd numbered slots the MSs are transmitting with power $2P$ and the RSs only receive, while during even numbered slots the MSs are silent and the RSs transmit. It is easily verified that the per-cell sum-rate in this case is given by multiplying the rates obtained above by $1/2$ while replacing P and Q, respectively, with $2P$ and $2Q$.

7.2.1.2 Single-cell Processing

We now assess the performance of the AF strategy when a conventional SCP scheme with spatial reuse factor $F = 1$ is implemented. Recall that for SCP each BS must be aware of the codebooks of its own MS only, as it treats all other cells' signals as interference. In this case, the output signal (Equation (2.10)) can be expressed as

$$y_{m,t}^{(2)} = s_{Um,t} + s_{Im,t} + q_{m,t}, \quad (7.18)$$

with the following definition. The useful part of the output signal s_U is defined as

$$s_{Um,t} = \sum_{l=-\infty}^{\infty} h_{S0,t-l}x_{m,l}^{(1)}.$$

We define h_S and h_N as the signal and noise space–time impulse response functions whose frequency responses are given in

Equations (7.12) and (7.13), respectively. The interference part of the output signal s_I is defined as

$$s_{I\,m,t} = \sum_{\substack{l_1=-\infty \\ l_1 \neq m}}^{\infty} \sum_{l_2=-\infty}^{\infty} h_{S\,m-l_1,t-l_2} x^{(1)}_{l_1,l_2},$$

and the equivalent noise is defined as

$$q_{m,t} = \sum_{l_1=-\infty}^{\infty} \sum_{l_2=-\infty}^{\infty} h_{N\,m-l_1,t-l_2} z^{(1)}_{l_1,l_2} + z^{(2)}_{m,t}.$$

Since $x^{(1)}$, $z^{(1)}$, and $z^{(2)}$ are independent of each other, zero-mean complex Gaussian and i.i.d. in space and time, it is easily verified that s_U, s_I, and q are independent and zero-mean complex Gaussian as well. It is also evident that for each m the processes are WSS along the time axis t. Accordingly, the received signal (Equation (7.18)) at the m-th BS can be seen as the output of an LTI system with additive colored independent interference and noise. Accordingly, the per-cell sum-rate of SCP with AF relaying is given for an arbitrary relay gain $0 < g < g_{\mathrm{o}}$, by [114, Appendix B]

$$R_{AF-SCP}(P) = \int_0^1 \log_2 \left(1 + \frac{\mathfrak{S}_U(\varphi)}{\mathfrak{S}_I(\varphi) + \mathfrak{S}_N(\varphi)}\right) d\varphi, \tag{7.19}$$

where $\mathfrak{S}_U(\varphi)$, $\mathfrak{S}_I(\varphi)$, and $\mathfrak{S}_N(\varphi)$ are the power spectral densities of the useful signal, interference, and noise, respectively:

$$\begin{aligned}
\mathfrak{S}_U(\varphi) &= P \left| \int_0^1 H_S(\theta,\varphi) d\theta \right|^2 \\
\mathfrak{S}_I(\varphi) &= P \int_0^1 |H_S(\theta,\varphi)|^2 \, d\theta - P \left| \int_0^1 H_S(\theta,\varphi) d\theta \right|^2 \quad \text{and} \\
\mathfrak{S}_N(\varphi) &= \sigma_1^2 \int_0^1 |H_N(\theta,\varphi)|^2 \, d\theta + \sigma_2^2.
\end{aligned}$$

We note that, in contrast to the MCP scheme studied above, the performance of the AF strategy with SCP is interference limited. Moreover, it is also easy to verify that the achievable rate (Equation (7.19))

is independent of the actual RS delay value λ. Finally, it will be verified below that, with SCP, due to the effect of interference, it turns out that it is not always optimal to employ the full transmission power Q at the RSs.

7.2.2 Compress-and-Forward (CF)

We now consider a transmission scheme that still requires no codebook information at the RSs, as for the AF scheme discussed above, but in which more processing capabilities are needed at the RSs. The CF-based strategy organizes transmission into successive blocks. The idea is that the RSs compress the signal received in any block and forward it in the next block to the BSs. We detail the strategy for MCP and SCP below.

7.2.2.1 Multi-cell Processing

With MCP, compression is done as if the RSs were connected by finite-capacity backhaul links directly to the CP. Note that this is clearly not the case since the RSs have to communicate with the BSs, which are in turn connected with unlimited-capacity links to the CP. The compression scheme is the same as discussed in Section 5.1.2. In particular, using the distributed coding strategies discussed therein, each RS in each block produces a compression index, which is binned to produce a bin index as explained in Appendix C. This is sent to the BSs by treating the RS-to-BS hop as a regular Wyner model. The CP then will decode over the second hop using standard MCP techniques and recover the bin indices. From the bin indices, using the same processing as for the CEO problem (see Appendix C), the CP can perform joint decompression and decoding as done in the oblivious scheme of Section 5.1.2.

The previous discussion hides a difference between the CF scheme considered here and the oblivious strategy of Section 3. In fact, here, by decoding the RSs' codewords, the CP acquires some information that is correlated with the signals received and compressed by the RSs. In fact, note that the RSs' received signals (Equation (2.10)) are correlated with the RSs' transmitted codewords $\{x_m^{(2)}(t)\}$. Therefore, while

compressing and binning their received signals, the RSs can leverage the fact that the CP will have correlated side information when decompressing. This further improves the efficiency of the use of the backhaul links.

The following per-cell rate can be shown to be achievable by the CF-MCP scheme [114, Appendix C]:

$$R_{\mathrm{CF-MCP}}(P) = R_{\mathrm{w}}(\alpha, \beta, \rho_1(1 - 2^{-r^*})), \tag{7.20}$$

where $r^* \geq 0$ is the unique solution to the following fixed point equation:

$$R_{\mathrm{w}}(\alpha, \beta, \rho_1(1 - 2^{-r^*})) = R_{\mathrm{w}}(\eta, \gamma, \rho_2) - r^*. \tag{7.21}$$

Note that since $R_{\mathrm{w}}(\alpha, \beta, \rho_1(1 - 2^{-r}))$ is monotonic in r, Equation (7.21) is easily solved numerically.

The rate (Equation (7.20)) can be easily explained by recalling the expression (Equation (5.6)) for the per-cell rate achievable by the oblivious scheme of Section 3 and related discussion. In fact, as for Equation (5.6), the rate (Equation (7.20)) shows that the CF scheme with MCP attains the maximum rate achievable if the RSs were all connected with ideal backhaul links to the CP (i.e., the bound (Equation (7.2))), but with an SNR loss given by $(1 - 2^{-r^*})$. The parameter r^*has the interpretation of the per-cell RS-to-BS rate that the RSs have to waste transmitting observation noise.

The analysis above reveals another important fact: the considered CF-MCP performs as if there is no inter-relay interference (i.e., $\mu = 0$). This is a consequence of exploiting the side information about the RSs' codewords when designing the compression and decompression/decoding strategy. Moreover, it can be verified that, as with the AF-MCP scheme, the use of the full relay power Q is optimal. Finally, it is easily verified that when $\rho_1 \to \infty$ then $r^* \to 0$, and $R_{CF-MCP}(P)$ does not achieve the upper bound (Equation (7.6)). This is since $R_{\mathrm{CF-MCP}} \underset{\rho_1 \to \infty}{\to} R_{\mathrm{w}}(\eta, \gamma, \rho_2) \leq R_{\mathrm{w}}^{\mathrm{wf}}(\eta, \gamma, \rho_2)$. On the other extreme when $\rho_2 \to \infty$ then $r^* \to \infty$, and $R_{\mathrm{cf-mcp}}$ achieves the upper bound $R_{\mathrm{CF-MCP}} \underset{\rho_2 \to \infty}{\to} R_{\mathrm{w}}(\alpha, \beta, \rho_1)$. In other words, CF-MCP works well when the performance is limited by the SNR on the first hop, while the second hop has a large SNR.

7.2.2.2 Single-Cell Processing

We now analyze the performance of the CF scheme in the presence of SCP with reuse factor $F = 1$. Recall that in this case each BS must be aware of its local MS and RS codebooks only, and decodes its signals treating other BSs' signals as noise. Each RS quantizes its received signal and transmits it to its local BS. Before describing the scheme in detail let us define the SINR at each RS and each BS $\tilde{\rho}_1$, and $\tilde{\rho}_2$, respectively, as

$$\tilde{\rho}_1 \triangleq \frac{P\beta^2}{\sigma_1^2 + 2\alpha^2 P + 2\mu^2 Q} \quad \text{and} \quad \tilde{\rho}_2 \triangleq \frac{Q\gamma^2}{\sigma_2^2 + 2\eta^2 Q}. \tag{7.22}$$

In addition, we denote the capacity of the single-user Gaussian channel with SNR ρ to be $R_g(\rho) = \log_2(1 + \rho)$ for simplicity.

According to the considered CF-SCP scheme, each RS compresses the received signal and then bins the compressed index. The bin index is transmitted to the same-cell BS. Each BS performs joint decompression/decoding of the message transmitted by the same-cell MS. It is easily verified that applying the CF-SCP scheme, as explained above, each cell (MS, RS, and BS) becomes equivalent to a multihop channel where CF is employed at the RS. This channel was studied in [85]. Using the results therein, we can show that an achievable per-cell rate is given by

$$R_{\mathrm{CF-SCP}} = R_{\mathrm{g}}(\tilde{\rho}_1(1 - 2^{-r^*})), \tag{7.23}$$

where $r^* \geq 0$ is the unique solution to the following fixed point equation:

$$R_{\mathrm{g}}(\tilde{\rho}_1(1 - 2^{-r^*})) = R_{\mathrm{g}}(\tilde{\rho}_2) - r^*, \tag{7.24}$$

where $\tilde{\rho}_1$ and $\tilde{\rho}_2$ are the SINRs at each RS and each BS, respectively, as defined above. In contrast to the MCP scheme analysis, here we can analytically solve Equation (7.24). Specifically, the achievable rate of Equation (7.23) is explicitly given by

$$R_{\mathrm{CF-SCP}}(P) = \log_2\left(\frac{(1 + \tilde{\rho}_1)(1 + \tilde{\rho}_2)}{1 + \tilde{\rho}_1 + \tilde{\rho}_2}\right). \tag{7.25}$$

As a remark, it is noted that the rate (Equation (7.25)) can also be achieved without joint decompression/decoding, but instead with

standard binning and successive decompression and decoding; see, e.g., [18]. In other words, with SCP, there is no need for more complex joint decoding/decompression strategies.

Examining the rate (Equation (7.25)) it is easily verified that it increases with the local path gains (β^2, γ^2) while it decreases with the inter-cell path gains $(\alpha^2, \eta^2, \mu^2)$. Hence, in contrast to the MCP scheme, the inter-relay interference is deleterious for the SCP scheme, which is interference-limited. In addition, the rate increases with the MSs' power and using the full power P is beneficial. On the other hand increasing the RSs' power unboundedly reduces $\tilde{\rho}_1$ to zero which drives $R_{\mathrm{CF-SCP}}(P)$ to zero as well. In particular, fixing the MSs' power P, the optimal RSs' power that maximizes the rate is given by

$$Q_{\mathrm{o}} = \min\left\{Q, \sqrt{\frac{((2\alpha^2+\beta^2)P+\sigma_1^2)\sigma_2^2}{2\mu^2(2\eta^2+\gamma^2)}}\right\}. \tag{7.26}$$

This shows that, as with AF-SCP, also with CF-SCP it is generally not optimal for the RSs to use all the available power.

7.3 Regenerative Relaying

In this section, following [102], we consider alternative relaying schemes for the system described in Section 2.4.1. Unlike the nonregenerative strategies studied in the previous section, the techniques analyzed here are regenerative in the sense that the RSs decode the signals sent by (some of) the MSs. Therefore, we refer to the strategies considered here as DF-based. Note that in order for DF-based strategies to be employed, the RSs must be informed about the codebooks of the MSs whose signals are to be decoded (see details below). Another relevant difference between the techniques proposed here and the nonregenerative strategies is that the latter do not require any change at the MS level, while the former generally do. In other words, with nonregenerative strategies, the RSs are transparent to the MSs, while this is in general not the case for regenerative strategies.

We organize the rest of this section as follows. First, we discuss the operation in the first hop. Then, we turn to the second hop and discuss per-cell achievable rates for SCP and finally for MCP. Throughout, to

simplify the analysis, we set $\mu = 0$ (i.e., no inter-relay interference). Achievable schemes with $\mu > 0$ can be obtained by assuming that each RS treats the signals from other RSs as additive noise, thus raising the noise level by $2\mu^2 P$. More sophisticated strategies where the inter-relay links are used to aid relay cooperation are possible, but will not be considered here.

7.3.1 Operation in the First Hop

Similar to the informed BS strategy of Section 5.3, the idea here is to use rate splitting [32], whereby the information bits of each MS are split into two parts. The first part, referred to as "private," is transmitted with power $\beta_1 P$ and is decoded only by the same-cell BS, with $0 \le \beta_1 \le 1$. The second part, referred to as "public" or "common", is transmitted with power $(1-\beta_1)P$ and is decoded not only at the local BS but also at the adjacent BSs. Therefore, each m-th RS jointly decodes four messages: the private message and the common message of the same-cell MS, and the common messages of the two adjacent-cell MSs, namely $m-1$ and $m+1$. The private messages of the two adjacent-cell MSs are instead considered as the interference terms with power $2\alpha^2\beta_1 P$. The channel seen at any m-th RS is then a four-user Multiple Access Channel (MAC) with equivalent Gaussian noise with power $\sigma_1^2 + 2\alpha^2\beta_1 P$. Accordingly, using well-known results on the capacity of MACs (see, e.g., [17]), for each choice of the power allocation factor β_1, the achievable per-cell rates of the private message part (R_{1p}) and of the common part (R_{1c}) are limited by the 15 inequalities defining the corresponding MAC capacity region, which are easily shown to reduce to

$$R_{1p} \le R_{\mathrm{g}}\left(\frac{\beta_1\rho_1}{\nu^2 + 2\alpha^2\beta\rho_1}\right) \triangleq R_{1p}^{\max}(\beta_1) \tag{7.27}$$

$$R_{1c} \le \min\left\{\frac{1}{2}R_{\mathrm{g}}\left(\frac{2\alpha^2(1-\beta_1)\rho_1}{\nu^2 + 2\alpha^2\beta_1\rho_1}\right), \tag{7.28}$$

$$\frac{1}{3}R_{\mathrm{g}}\left(\frac{(2\alpha^2+1)(1-\beta_1)\rho_1}{\nu^2 + 2\alpha^2\beta\beta_1\rho}\right)\right\} \tag{7.29}$$

$$\triangleq \min\{R_{1c}^{\max,1}(\beta_1),\ R_{1c}^{\max,2}(\beta_1)\} \tag{7.30}$$

$$R_{1p} + 2R_{1c} \leq R_{\mathrm{g}}\left(\frac{\beta_1\rho_1 + 2\alpha^2(1-\beta_1)\rho_1}{\nu^2 + 2\alpha^2\beta_1\rho_1}\right) \tag{7.31}$$

$$\triangleq R_1^{sum,1}(\beta_1) \tag{7.32}$$

$$R_{1p} + 3R_{1c} \leq R_{\mathrm{g}}\left(\frac{\beta_1\rho_1 + (2\alpha^2+1)(1-\beta_1)\rho_1}{\nu^2 + 2\alpha^2\beta_1\rho_1}\right) \tag{7.33}$$

$$\triangleq R_1^{sum,2}(\beta_1), \tag{7.34}$$

where $\nu^2 = 1$ and we recall the definition $R_g(P) = \log_2(1+P)$. Notice that in writing the conditions above we have removed dominated inequalities. We refer to the region of rates (R_{1p}, R_{2p}) satisfying the inequalities above as $\mathcal{R}_1(\beta_1)$.

The rate region identified by the inequalities above for a given power allocation β_1 is a polyhedron. It can be shown that for each power allocation β, the maximum rate $R_{1p} + R_{1c}$ attainable in the first hop is obtained in one of the vertices of this polyhedron [102]. Moreover, the rate pair (R_{1p}, R_{1c}) corresponding to the maximum sum-rate can be achieved via successive interference cancellation by first jointly decoding the common messages, treating the private information as noise, then cancelling the decoded common messages and finally decoding the same-cell private message. Specifically, the maximum rate for a given power allocation β_1 is given by

$$R_1^{\max}(\beta_1) = R_{1p}^{\max}(\beta_1) \tag{7.35}$$

$$+\min(R_{1c}^1(\beta_1),\ R_{1c}^2(\beta_1)), \tag{7.36}$$

with Equation (7.27) and

$$R_{1c}^1(\beta_1) = \frac{1}{2}R_g\left(\frac{2\alpha^2(1-\beta_1)\rho_1}{\nu^2 + (2\alpha^2+1)\beta_1\rho_1}\right) \quad \text{and} \tag{7.37}$$

$$R_{1c}^2(\beta_1) = \frac{1}{3}R_g\left(\frac{(2\alpha^2+1)(1-)\rho_1}{\nu^2 + (2\alpha^2+1)\beta_1\rho_1}\right), \tag{7.38}$$

where $\nu^2 = 1$.

7.3.2 Single-Cell Processing

Having described the system operation for the first hop, we now turn to the second hop. We describe two techniques assuming that each BS

performs decoding independently from other BSs according to SCP with spatial reuse $F = 1$.

7.3.2.1 Noncooperative Relay Stations

As explained, with rate splitting in the first hop, each RS, say the m-th, decodes in each block the private message and the common message of the same-cell MS, along with the common messages of the adjacent cells. A simple approach is for the m-th relay to neglect the knowledge of the common messages of adjacent cells and to simply retransmit to the local BS the private and common messages of the same-cell MS. This is done using rate splitting and interference cancellation exactly as explained in the previous section for the first hop. Note that the total rate $R_{1p} + R_{1c}$, delivered to the RSs by the MSs in the first hop, can be now split into two streams, one private and one common, in a generally different share with respect to the first hop. In particular, a different power allocation between private and common parts, say β_2, can be used in the second hop, so that the private message is transmitted with power $\beta_2 Q$ and the common message with power $(1 - \beta_2)Q$. Note that, in order to implement this scheme, each BS must be informed of the private codebook used by the same-cell MS and of the common codebooks employed by the same-cell MSs and the two adjacent-cell MSs.

Following the previous section, we can thus obtain the maximum per-cell rate that can be transmitted in the second hop as (recall Equation (7.35))

$$R_2^{\max}(\beta_2) = R_{2p}^{\max}(\beta_2) \tag{7.39}$$

$$+\min(R_{2c}^1(\beta_2),\ R_{2c}^2(\beta_2)), \tag{7.40}$$

where $R_{2p}^{\max}(\beta_2)$, $R_{2c}^1(\beta_2)$, and $R_{2c}^2(\beta_2)$ are obtained from Equations (7.27), (7.37) and (7.38), respectively, where subscript "2" should be substituted for "1" and parameters (γ^2, η^2) should be written in lieu of (ν^2, α^2).

Since with rate splitting in both hops the two hops are operated independently, the optimal strategy is to transmit in both hops at the maximum sum-rates $R_1^{\max}(\beta_1)$ and $R_2^{\max}(\beta_2)$ for given power allocations β_1 and β_2. It follows that, optimizing over the power allocation

on both hops, the rate achievable with rate splitting in both hops is

$$R_{DF-SCP-NO-COOP}(P) = \min_{i=1,2} \max_{0 \leq \beta_i \leq 1} R_i^{\max}(\beta_i). \tag{7.41}$$

7.3.2.2 Cooperative Relay Stations

In this section, we investigate the performance of an alternative transmission scheme for the second hop that leverages the common information gathered at the RSs as a by-product of the use of rate splitting in the first hop. This contrasts with the naive scheme discussed above whereby the common messages from adjacent cells were neglected when transmitting in the second hop.

The rate splitting-based scheme discussed above for transmission from RSs to BSs fails to exploit the knowledge of the common messages of adjacent cells at any m-th RS. Based on this side information, any m-th cell could cooperate with the adjacent cells $m-1$ (and $m+1$) in order to deliver the common messages of these cells to the intended BS in cell $m-1$ (and $m+1$). To this end, we consider a superposition scheme whereby RSs cooperate for transmission of common information toward the goal of achieving coherent power combining at the BSs.

The private and common messages are the ones sent in the first hop by the MSs and therefore have rates R_{1p} and R_{1c}, respectively. We focus on a simple power allocation among the transmitted codewords by the RSs, whereby the total power Q is divided according to a parameter β_2 as above, so that power $\beta_2 Q$ is used for transmission by the RSs of the private part and the power $(1-\beta_2)Q$ is equally shared among the three cooperative common signals. As in the previous section, each BS is assumed to know the private codebook used by the same-cell MS and of the common codebooks employed by the same-cell MSs and the two adjacent-cell MSs in order to enable joint decoding. Specifically, each BS performs joint decoding of four messages: the private message and the common message of the same-cell MS, and the common messages of the two adjacent-cell MSs, namely $m-1$ and $m+1$. The common messages of cells $m-2$ and $m+2$ are considered as interference by the m-th BS.

It can be seen that any m-th BS observes a four-user MAC with equivalent noise power $\sigma_2^2 + 2\eta^2(\beta_2 Q + (1-\beta_2)Q/3)$. Therefore, similarly to the discussion above, the achievable rates (R_{1p}, P_{1c}) of the private and common information must satisfy

$$\begin{aligned}
R_{1p} &\leq R_g\left(\frac{\gamma^2\beta_2\rho_2}{1+2\eta^2(\beta_2\rho_2+(1-\beta_2)Q/3)}\right)\\
R_{1c} &\leq \min\left\{\frac{1}{2}R_g\left(\frac{2(\gamma+\eta)^2(1-\beta_2)\rho_2}{1+2\eta^2(\beta_2\rho_2+(1-\beta_2)\rho_2/3)}\right),\right.\\
&\qquad \left.\frac{1}{3}R_g\left(\frac{(2(\gamma+\eta)^2+(\gamma+2\eta)^2)(1-\beta_2)\rho_2}{1+2\eta^2(\beta_2\rho_2+(1-\beta_2)\rho_2/3)}\right)\right\}\\
R_{1p}+2R_{1c} &\leq R_g\left(\frac{\gamma^2\beta_2\rho_2+2(\gamma+\eta)^2(1-\beta_2)\rho_2}{1+2\eta^2(\beta_2\rho_2+(1-\beta_2)\rho_2/3)}\right)\\
R_{1p}+3R_{1c} &\leq R_g\left(\frac{\gamma^2\beta_2\rho_2+(2(\gamma+\eta)^2+(\gamma+2\eta)^2)(1-\beta_2)\rho_2}{1+2\eta^2(\beta_2\rho_2+(1-\beta_2)\rho_2/3)}\right).
\end{aligned}$$

We refer to the polytope of rates (R_{1p}, R_{1c}) satisfying the inequalities above as $\mathcal{R}_{COOP}(\beta_2)$. It is noted that the effect of RS cooperation is seen above in the beamforming (array) gains due to coherent combining, which are reflected in the effective channel gains $(\gamma+\eta)^2$ and $(\gamma+2\eta)^2$. Focusing on the m-th BS, the first, $(\gamma+\eta)^2$, accounts for the effective channel gains of the common message of cells $m-1$ and $m+1$, which are received both directly and via the signal transmitted by the m-th RS. The second, $(\gamma+2\eta)^2$, accounts for the effective channel gain of the common message of the m-th RS, which is received, not only through direct transmission, but also via the signals relayed by the $(m-1)$-th and the $(m+1)$-th RSs.

The maximum achievable rate with rate splitting in the first hop and cooperative transmission in the second hop, according to the coding scheme described above, can be then found by solving the following optimization problem:

$$R_{DF-SCP-COOP}(P) = \max_{R_{1p},R_{1c},\beta_1,\beta_2} R_{1p}+R_{1c} \tag{7.42}$$

$$\text{s.t.} \begin{cases} 0\leq\beta_1,\beta_2\leq 1 \\ (R_{1p},R_{1c})\in \mathcal{R}_1(\beta_1)\cap \\ \qquad\mathcal{R}_{COOP}(\beta_2). \end{cases} \tag{7.43}$$

Notice that for each choice of the power allocation (β_1, β_2), the optimization problem (Equation (7.42)) can be solved by linear programming.

7.3.3 Multi-Cell Processing

In this section we consider the performance of DF-based schemes with MCP. We assume the use of rate splitting in the first hop, whereas in the second hop the cooperative transmission scheme discussed above, which aims at coherent power combining at the BSs for the common messages, is employed. Following this scheme, similar to what was done above, we can interpret the received signal at the BSs as an equivalent LTI system with inputs given by the signals transmitted by the RSs. In particular, defining as $x_{p,m}^{(2)}$ and $x_{c,m}^{(2)}$ the codewords transmitted by the RSs that encode the private and common messages of the m-th MS, one can easily prove that the received signal at the m-th BS is given by

$$y_m^{(2)} = h_{p,m} * x_{p,m}^{(2)} + h_{c,m} * x_{c,m}^{(2)} + z_m^{(2)}, \tag{7.44}$$

where "$*$" denotes convolution and the finite-impulse response filters $h_{nc,m}$ and $h_{c,m}$ are given by

$$h_{p,m} = \eta\delta_{m+1} + \gamma\delta_m + \eta\delta_{m-1} \quad \text{and} \tag{7.45}$$

$$\begin{aligned} h_{c,m} = {} & \eta\delta_{m+2} + (\gamma+\eta)\delta_{m+1} + (\gamma+2\eta)\delta_m \\ & + (\gamma+\eta)\delta_{m-1} + \eta\delta_{m-2}, \end{aligned} \tag{7.46}$$

with δ_m denoting the Kronecker delta function. The channel (Equation (7.44)) is a Gaussian MAC with inter-symbol interference, so that the achievable rates (R_{1p}, R_{1c}) in the second hop must satisfy the conditions [13]

$$\begin{aligned} R_{1p} &\leq R_w(\eta, \gamma, \beta_2 Q) \\ R_{1c} &\leq \int_0^1 R_g\left(\frac{(1-\beta_2)Q}{3}(\gamma + 2\eta + 2(\gamma+\eta)\cos(2\pi\theta)\right. \\ &\quad \left. + 2\eta\cos(4\pi\theta))^2\right) d\theta \end{aligned}$$

$$R_{1p} + R_{1c} \leq \int_0^1 R_g \Big(\beta_2 Q(\gamma + 2\eta\cos(2\pi\theta))^2 + \frac{(1-\beta_2)Q}{3}(\gamma + 2\eta + +2(\gamma+\eta)\cos(2\pi\theta) + 2\eta\cos(4\pi\theta))^2 \Big) d\theta,$$

where we used the definition (7.1). We refer to the polytope of rates (R_{1p}, R_{1c}) described by the inequalities above as $\mathcal{R}_{MCP}(\beta_1, \beta_2)$.

Finally, accounting for both first and second hops, the rate achievable with rate splitting, relay cooperation and multi-cell processing can be obtained by solving the following optimization problem:

$$R_{mcp} = \max_{R_{1p}, R_{1c}, \beta_1, \beta_2} R_{1p} + R_{1c} \tag{7.47}$$

$$\text{s.t.} \begin{cases} 0 \leq \beta_1, \beta_2 \leq 1 \\ (R_{1p}, R_{1c}) \in \mathcal{R}_1(\beta_1, \beta_2) \cap \\ \qquad \mathcal{R}_{MCP}(\beta_1, \beta_2). \end{cases} \tag{7.48}$$

Notice again that, for fixed power allocation (β_1, β_2), problem (Equation (7.47)) can be solved by linear programming.

7.4 Numerical Results

We first provide a performance comparison of the DF-based techniques that we have just presented and then discussed a comparison among all schemes. Figure 7.2 shows the achievable per-cell rates by the considered regenerative schemes for $\rho_1 = 2$, $\rho_2 = 1$, $\sigma_1^2 = \sigma_2^2 = 1$, $\gamma^2 = 1$, $\mu^2 = 0$ versus the inter-cell gains $\alpha^2 = \eta^2$. We compare the per-cell rates achievable with SCP with and without RS cooperation and with MCP. As for the DF-SCP technique without RS cooperation, we also show the special case $\beta_1 = \beta_2 = 1$, which corresponds to no rate splitting, i.e., to transmission of only private messages. This scheme is of interest for its simplicity and because it can be implemented by maintaining the deployment of relays transparent to the MSs, since the latter do not have to change their transmission strategies. For reference, we plot the maximum rate achievable on the first hop with rate splitting and optimal power allocation $R_1^{\max}$ (Equation (7.35)). This provides an upper

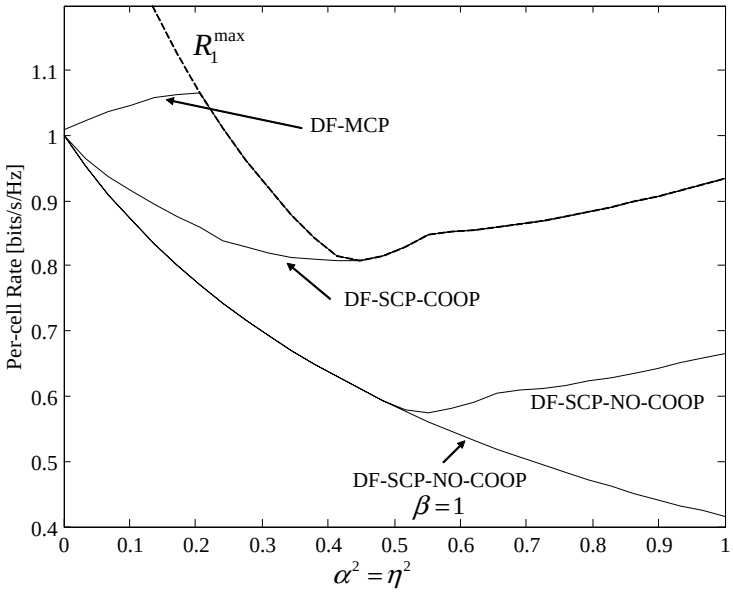

Fig. 7.2 Per-cell achievable rates for the DF-based (regenerative) schemes versus the inter-cell gains $\alpha^2 = \eta^2$ ($\rho_1 = 2$, $\rho_2 = 1$, $\sigma_1^2 = \sigma_2^2 = 1$, $\gamma^2 = 1$, $\mu^2 = 0$).

bound on the per-cell achievable rate for the considered DF-based techniques in the scenario at hand where the SNR in the second hop creates the performance bottleneck (since $\rho_2 = \rho_1/2$).

It can be seen that rate splitting is advantageous with respect to single-rate transmission (i.e., $\beta_1 = \beta_2 = 1$) if the inter-cell gains $\alpha^2 = \eta^2$ are large enough. Moreover, cooperation at the relays provides relevant performance gains and allows the network to achieve the upper bound $R_1^{\max}$ for $\alpha^2 = \eta^2$ large enough. Finally, MCP allows the upper bound $R_1^{\max}$ to be achieved for a larger range of $\alpha^2 = \eta^2$ than SCP with RS cooperation.

We now present a comparison among the regenerative and the non-regenerative schemes discussed in this section for $\gamma^2 = 1$, $\alpha^2 = \eta^2 = 0.2$, $\sigma_1^2 = \sigma_2^2 = 1$, $\mu^2 = 0.1$, $\rho_1 = 10$ versus the ratio ρ_2/ρ_1 in Figure 7.3. For DF, we consider strategies with relay cooperation. A first critical observation is the interplay between the deployment of SCP or MCP and the choice of relaying strategies. Specifically, it can be seen that if SCP is employed, DF is advantageous with respect to CF, and also with

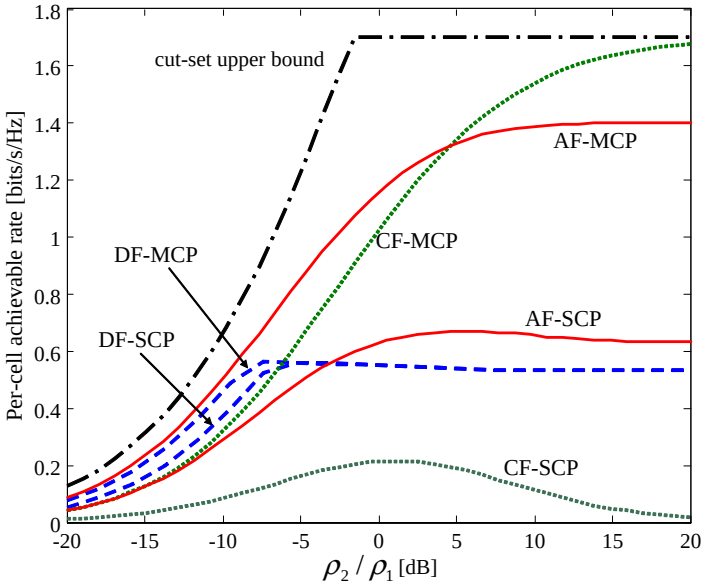

Fig. 7.3 Per-cell achievable rates for regenerative and the nonregenerative schemes versus the ratio ρ_2/ρ_1 ($\gamma^2 = 1$, $\alpha^2 = \eta^2 = 0.2$, $\sigma_1^2 = \sigma_2^2 = 1$, $\mu^2 = 0.1$, $\rho_1 = 10$).

respect to AF, if the power of the MSs is sufficiently larger than that of the RSs and thus ρ_2/ρ_1 is sufficiently larger than one. It is noted that CF performs very poorly due to its inability, unlike DF and AF, to beamform the users' signals toward the BSs. However, if MCP is in place, the situation is remarkably different in that DF is outperformed by both CF and AF unless the MSs' power is sufficiently larger than that of the relays. This is because DF is limited by the performance bottleneck due to the need to decode at the RSs, which prevents the system from benefiting from MCP. Finally, it is seen that the proposed CF scheme performs close to optimal if the relay power is sufficiently large.

7.5 Summary

In this section, we have studied the performance of cellular systems with RSs and with conventional SCP or with MCP. We have focused on Gaussian channels and considered the deployment of RSs as a means

to extend the system coverage for the uplink. We have investigated both regenerative techniques, which require the RSs to know the MSs' codebooks and are typically nontransparent to the MSs, and nonregenerative techniques, which do not require any codebook knowledge at the RSs and are transparent to the MSs. We have seen that the design of the relaying techniques depends critically on whether SCP or MCP is implemented. In fact, if SCP is employed, regenerative techniques (DF) are seen to be advantageous with respect to nonregenerative strategies, unless the relay power is sufficiently large, in which case AF has comparable performance. Instead, if MCP is in place, DF is outperformed by both nonregenerative techniques, CF and AF, unless the MSs' power is sufficiently larger than that of the relays. This is because the decoding requirement at the relays of DF can set the performance bottleneck, thus nullifying the potential gains attainable with MCP.

8

Mobile Station Cooperation

In this section, we address the impact of cooperation in cellular networks when there is cooperation at the MS level. We follow the uplink model discussed in Section 2.4.2 and consider first out-of-band cooperation and then in-band cooperation. We recall that in both cases we study cooperation between MSs in adjacent cells. As explained in Section 2.4.2, inter-cell MS cooperation is useful if MCP is deployed, since, thanks to MCP, the signal relayed by an adjacent-cell MS can be exploited by the CP for decoding. For this reason, we will focus solely on MCP here. It is noted that some analysis of intra-cell cooperation can be found in [105]. We remark that MS cooperation raises issues related to privacy and altruistic versus selfish behavior that are not considered here (see, e.g., [50]).

8.1 Upper Bound

We start by first deriving an upper bound on the per-cell rate achievable by exploiting inter-MS cooperation for the model presented in Section 2.4.2. Specifically, as pointed out in Section 2.4.2, an upper bound on the per-rate achievable with inter-MS cooperation is given by

$$R(P) \leq R_{UB}(P) = R_{\mathrm{w}}^{\mathrm{wf}}(\alpha, 1, P), \tag{8.1}$$

where $R_{\mathrm{w}}^{\mathrm{wf}}(\alpha,1,P)$ is as defined in Equation (7.5), where it was found to be equal to

$$R_{\mathrm{w}}^{\mathrm{wf}}(\alpha,1,P) = \int_0^1 \log_2\left(1+\left(\nu-\frac{1}{H(\theta)^2}\right)^+ H(\theta)\right)d\theta$$
$$\text{s.t.} \int_0^1 \left(\nu-\frac{1}{H(\theta)^2}\right)^+ d\theta = P, \tag{8.2}$$

with $H(\theta)=1+2\alpha\cos(2\pi\theta)$. This bound follows again from cut-set arguments.

In order to interpret Equation (8.1) and obtain insights into the possible gains attainable with inter-cell MS cooperation, it is useful to recall that $R_{\mathrm{w}}^{\mathrm{wf}}(\alpha,1,P)$ is given by

$$R_{\mathrm{w}}^{\mathrm{wf}}(\alpha,1,P) = \max \int_0^1 \log(1+P\cdot H(\theta)^2 S(\theta))d\theta, \tag{8.3}$$

where the maximum is taken over the input power spectral density $S(\theta)$ such that the power constraint $\int_0^1 S(f)df = 1$ is satisfied. (This can be shown to be given by the "waterfilling" solution $S(\theta)=\left(\nu-\frac{1}{PH(\theta)^2}\right)^+$ where the parameter $\nu\geq 0$ is selected so as to satisfy the power constraint.) Now, from the discussion in Section 3.1.2, we know that the power spectral density $S(\theta)$ accounts for the correlation among the signals sent by the MSs. The MSs are able to correlate their signal only if they have exchanged information so as to have common data on which correlation can be built. Note that this is the same mechanism that allowed us to obtain RS cooperation in Section 7. Also, recall that a constant $S(\theta)$ represents noncooperative MSs, while a nonconstant $S(\theta)$ is possible only if cooperation at the MSs is enabled.

In light of the discussion above and of Equation (8.3), the optimal correlation of the MSs' signals corresponds to the waterfilling solution. By examining this solution and its corresponding time-domain transform, one can conclude that, in order for such a power spectral density to be realized, correlation needs to span a large number of MSs, especially if α is large enough. In other words, many MSs would have to exchange information in order for the upper bound to be achieved. Note that clearly if $\alpha=0$ neither MCP nor inter-cell MS cooperation

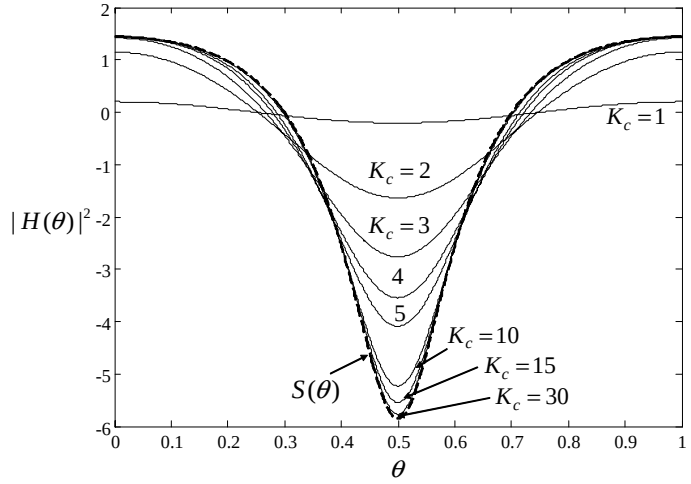

Fig. 8.1 Comparison between the optimal (waterfilling) correlation of the signals transmitted by the MSs and approximations obtained where each MS can cooperate only with K_c on the left and on the right ($\alpha = 0.2$ and $P = 2$).

can provide performance gains. An example for $\alpha = 0.2$ and $P = 2$ is shown in Figure 8.1. There, the optimal waterfilling power spectral density $S(\theta)$ is compared with the best approximation achievable if each MS is able to cooperate only with K_c MSs on the left and on the right. It is seen that, even for a relatively small value of α, the number of cooperating MSs must be rather large in order to approximate the optimal input correlation and thus the optimal cooperative strategy. Further details on how this example was generated will be provided below.

8.2 Out-of-Band Cooperation

In this section, we analyze the performance gains attainable with out-of-band MS cooperation. From Section 2.4.2, it is recalled that out-of-band cooperation amounts to having each MS connected to the two MSs in adjacent cells via finite-capacity links of capacity $C^{(MS)}$ (bits/s/Hz). Moreover, transmission on each inter-MS out-of-band link can be performed in parallel and without interference to/from the transmissions to the BSs. This implies that, by arranging transmissions in successive blocks, the MSs can exchange information on the inter-MS

links in a given block, while transmitting to the BSs, and then use the exchanged information in the next block. By pipelining transmissions, this guarantees that in each block the MS can cooperate when transmitting to the BSs based on the signals exchanged over the inter-MS links.

8.2.1 An Achievable Rate

In this section, we derive an achievable rate for the model at hand and discuss some of the implications of this result. The scheme, proposed in [105], works as follows. We employ again *rate splitting* so that each m-th MS splits its information bits into two parts, private and common, of rates R_c and R_p (bits/s/Hz), respectively. Then, it shares the common part with the $2K_c$ neighboring MSs in cells $m+i$ with $i=-K_c,-K_c+1,\ldots,-1,1,\ldots,K_c$, i.e., with K_c MSs on the left and K_c on the right. This is realized by sending the common message over all the inter-MS links between the m-th MS and the $(m+i)$-th MS. Notice that, after this information exchange, each m-th MS is aware of the $2K_c$ common messages beside its own. Moreover, the need to send these messages on the inter-MS links is easily seen to impose the constraint

$$R_c \leq \frac{C^{(MS)}}{K_c}. \tag{8.4}$$

In the block following the information exchange above, each common message can then be transmitted cooperatively by all the $2K+1$ MSs that have acquired the information in the conferencing phase. Moreover, superimposed on the codewords encoding the common messages, each MS also sends its private message. Transmission is performed in a way similar to the way RSs format their transmissions in the cooperative strategy discussed in Section 7.3.2.2. The following per-cell rate is achieved:

$$\begin{aligned} R(P,C^{(MS)}) = \max_{\beta,\mathbf{h}_c} \min \Bigg\{ & \int_0^1 \log(1+\beta PH(\theta)^2 \\ & +(1-\beta)PH(\theta)^2|H_c(\theta)|^2)d\theta, \\ & \int_0^1 \log(1+\beta PH(\theta)^2)d\theta + \frac{C^{(MS)}}{K_c} \Bigg\}, \end{aligned} \tag{8.5}$$

where $0 \leq \beta \leq 1$ is the power allocation coefficient between private and common messages, $\mathbf{h}_c = [h_{c,-K_c} \cdots h_{c,K_c}]^T$ is a symmetric complex vector of $2K_c + 1$ entries with unitary norm, $\|\mathbf{h}_c\|^2 = 1$, and

$$H_c(\theta) = \sum_{m=-K_c}^{K_c} h_{c,m} \exp(-j2\pi\theta m). \tag{8.6}$$

As shown by the result above, the impact of inter-cell MS cooperation, using the considered scheme, is equivalent to that of allowing *precoding* (*pre-equalization*) of the common information by a $2K_c + 1$ finite-impulse filter $\mathbf{h}_c$ with frequency response $H_c(\theta)$ Equation (8.6) on the LTI system $H(\theta)$ that accounts for the Wyner model. This allows the correlation of the signals transmitted by the MSs, as far as the common messages are concerned, to be given by $|H_c(\theta)|^2$. As can be inferred from Equation (8.5), instead, the private messages are sent with a constant power spectral density (i.e., without cooperation). We emphasize that, while the number of taps of the filter $\mathbf{h}_c$ increases with the number of cooperative MSs, the overall achievable rate may suffer according to Equation (8.5) due to the need to exchange more information among MSs. We further explore this trade-off below with a numerical example.

8.2.2 Asymptotic Optimality

The scheme considered above is easily seen to be optimal for $C \to \infty$, $K_c \to \infty$ and $\frac{C^{(MS)}}{K_c} \geq R_{UB}(P)$. This can be proved by noting that the rate (Equation (8.5)) equals the upper bound (Equation (8.1)) under the conditions in the proposition above by setting $\beta = 0$ and recalling that the optimal (waterfilling) power spectral density $S(\theta)$ can be approximated arbitrarily well by the frequency response $|H_c(\theta)|^2$ in Equation (8.6) as the number of taps $2K_c + 1$ increases unboundedly; see Figure 8.1 for an illustration. Note that, in particular, this argument implies that, under the given asymptotic conditions, it is optimal to allocate all the power to the common messages.

We also remark that, while here we do not consider fading channels, it is apparent from the discussion above that the advantages of inter-cell MS cooperation are related to the possibility of optimizing

the transmission strategy based on the knowledge of the channel structure at the MSs. Therefore, inter-cell conferencing is expected not to provide any performance gain over fading channels in the absence of channel state information at the MSs. This claim can be substantiated by using the results in [121], where it is shown that, in the case of independent fading channels, even in the presence of *statistical* channel state information at the transmitters (i.e., at the MSs) the optimal power allocation is asymptotically uniform so that cooperation at the MSs does not provide any advantage.[1]

8.2.3 Low-SNR Analysis

We now further analyze the performance of the inter-cell MS cooperative scheme at hand in the low-SNR regime. The goal is to assess to what extent the minimum energy per bit required for reliable communication obtained with ideal MS cooperation is attainable with the scheme discussed above. First, we remark that from Equation (8.1), with ideal MS cooperation, we can ideally achieve

$$\frac{E_b}{N_0}_{\min,UB} = \frac{\ln 2}{(1+2\alpha^2)^2}. \tag{8.7}$$

The latter can be proved by noticing that, when the SNR tends to zero ($P \to 0$), it is optimal to allocate all the available power around the maximum value of the channel transfer function, $\max_\theta H(\theta)^2 = (1+2\alpha)^2$, which occurs at $\theta = 0$. In other words, the optimal waterfilling power allocation is $S(\theta) = \delta(\theta)$, where $\delta(\theta)$ is a Dirac delta function. This result implies, when compared with the minimum energy per bit (Equation (3.16)) attainable without MS cooperation, that MS cooperation can lead, ideally, to a power gain of $1 + 2\alpha^2$.

We now analyze Equation (8.5). In [105], the following approximation is derived:

$$\frac{E_b}{N_0}_{\min} \simeq \frac{\ln 2}{(1+2\alpha)^2\left(1 - \frac{8\alpha\pi^2}{3(1+2\alpha)K_c^2}\right)}. \tag{8.8}$$

[1] This result holds for channels with column-regular gain matrices (see the definition in [121]). The channel considered here belongs to this class when $M \to \infty$.

This shows that the minimum energy per bit achievable with inter-cell MS cooperation is a decreasing function of the number K_c of cooperating MSs and, as expected from the discussion above, tends to the optimal performance (8.7) by letting $K_c \to \infty$.

8.2.4 Numerical Results

As discussed above, increasing K_c is always beneficial in obtaining a better approximation of the optimal (waterfilling) power spectral density. However, due to the finite inter-MS capacity $C^{(MS)}$, it is not necessarily advantageous in terms of the achievable rate (Equation (8.5)). To show this, Figures 8.2 and 8.3 present the achievable per-cell rate (Equation (8.5)) versus the inter-cell gain α along with the lower bound obtained by setting $C^{(MS)} = 0$ (which corresponds to Equation (3.12)) and the upper bound (Equation (8.1)) for $C^{(MS)} = 1$ and $C^{(MS)} = 10$ (bits/s/Hz), respectively, with $P = 2$. Figure 8.2 shows that, with $C^{(MS)} = 1$, while increasing the cooperating MSs from $K_c = 1$ to 2 increases the achievable rate, further increments of the

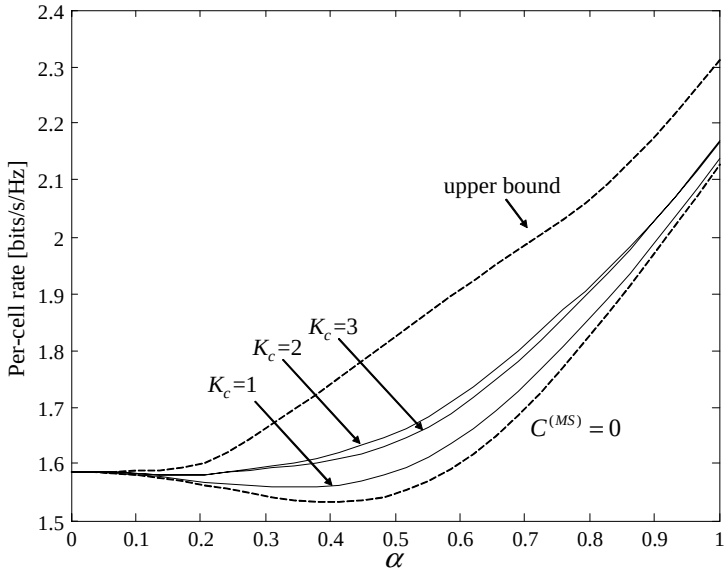

Fig. 8.2 Achievable per-cell rates and bounds versus the inter-cell gain α for out-of-band MS cooperation with $C^{(MS)} = 1$ (bits/s/Hz) ($P = 2$).

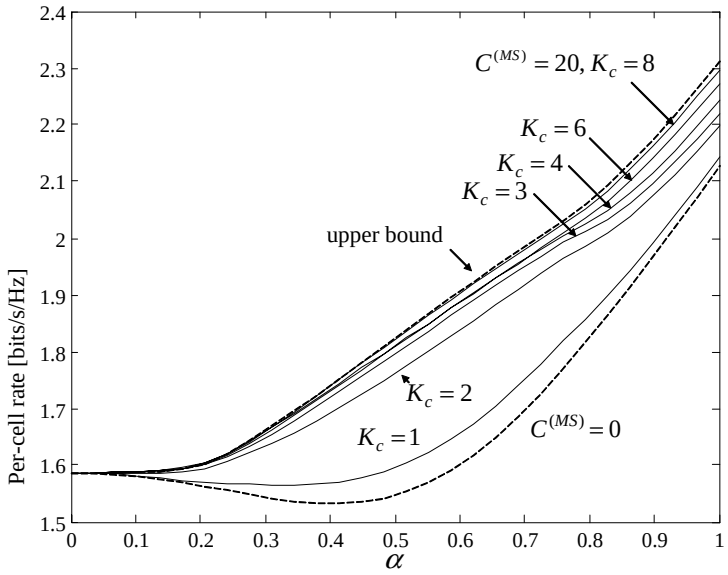

Fig. 8.3 Achievable per-cell rates and bounds versus the inter-cell gain α for out-of-band MS cooperation with $C^{(MS)} = 10$ (bits/s/Hz) ($P = 2$).

inter-MS capacity $C^{(MS)}$ are disadvantageous, according to the trade-off mentioned above. With a larger capacity $C^{(MS)} = 10$, Figure 8.3 shows that very relevant performance gains can be harnessed by increasing the number of cooperating MSs, especially from $K_c = 1$ to $K_c = 2$. Moreover, as expected from the discussion above, having sufficiently large inter-MS capacity $C^{(MS)}$ and number of cooperating MSs K_c (with $C^{(MS)}/K_c \geq R_{UB}$) enables the upper bound (Equation (8.1)) to be approached.

8.3 In-Band Cooperation

The previous section showed that with out-of-band links of sufficiently large capacity, inter-MS cooperation is able to achieve the performance promised by the upper bound obtained under the assumption of ideal MS cooperation. These gains are substantial, as precisely quantified above in the low-SNR regime. In this section, we study how much of the promised gains can be instead realized if out-of-band links are not available for inter-cell MS cooperation. Instead, the MSs are assumed

to be able to receive on the same bandwidth they use for transmission to the BSs, as detailed in Section 2.4.2 (see Figure 2.7). This reception capability of the MSs can be capitalized upon by letting the MSs convey information to adjacent MSs by encoding such information on the same signal used for transmission to the BSs.

To elaborate, we focus on a technique proposed in [106]. The technique attempts to mimic the approach used for out-of-band MS cooperation in that each MS performs rate splitting into private and common parts, and then shares the common part with K_c MSs on the right and K_c on the left. Note that such information exchange must occur on the same bandwidth used for communication to the BSs so that, apart from the power allocated for transmission of private and common messages, a fraction of the power should also be used for *signalling* the common information to nearby users. To illustrate this point, we take a brief detour and discuss a MAC with only two users and a destination.

8.3.1 Gaussian Multiple Access Channel With In-Band Cooperation

Consider a symmetric Gaussian MAC, in which two users communicate with a destination. The signal received by the destination at time instant t is given by

$$y(t) = x_1(t) + x_2(t) + z(t), \quad i = 1, 2, \ldots, n, \tag{8.9}$$

where the noise sequence $z(t)$ is i.i.d. complex Gaussian with unit power and we enforce a power constraint of P on both transmitters. The m-th MS, similar to Equation (8.10), receives the following signal at time t:

$$y_m^{(MS)}(t) = \kappa x_{m'}(t) + z_m^{(MS)}(t), \tag{8.10}$$

where $m' = 2$ if $m = 1$ and $m' = 1$ if $m = 2$. Note that the noises at the destination and at the corresponding MS may be correlated, as this will not affect our results [106]. From Equation (8.10), each MS overhears the signal transmitted by the other with a channel gain $\kappa \geq 0$.

We start by observing that, with ideal cooperation, the per-user rate

$$R_{UB}(P) = \frac{1}{2}\log_2(1 + 4P) \tag{8.11}$$

can be achieved, since with full cooperation, a beamforming (array) gain of 2 can be attained. The rate (Equation (8.11)) provides an upper bound on the achievable rate with any form of cooperation. To assess to what extent this upper bound can be realized, we study below an achievable scheme.

Using standard tools (see, e.g., [131]), the following per-user rate R (bits/s/Hz) can be shown to be achievable:

$$R(P) = \max_{0\leq\beta,\nu\leq 1} \min \left\{ \begin{array}{l} \frac{1}{2}\log_2(1+2\beta P) + R_c, \\ \\ \frac{1}{2}\log_2(1+2P+2\nu(1-\beta)P) \end{array} \right\}, \tag{8.12}$$

with

$$R_c = \log\left(1+\frac{\kappa^2(1-\nu)(1-\beta)P}{1+\kappa^2\beta P}\right). \tag{8.13}$$

The rate (Equations (8.12) and (8.13)) is achieved again by rate splitting. However, here a further level of power allocation is necessary in order to account for inter-MS signalling. Specifically, each MS splits its rate and powers between a *private* and a *common* part, so that the rate R is split as $R = R_p + R_c$ (as above, subscripts denote private "p" and common "c" parts), and power βP is devoted to private message transmission and $(1-\beta)P$ to common message transmission. As for the out-of-band strategy, the private part is sent to the BS without any cooperation from the other MS, while transmission of the common part benefits from the cooperation with the other MS. Unlike the out-of-band strategy, in order to enable cooperative transmission, a fraction $(1-\nu)$ of the common power $(1-\beta)P$ must be devoted to *signalling* the common message to the other MS (while the rest is used for transmission to the BSs). To be precise, transmission is organized into blocks so that the information signalled in a certain block is used in the next for cooperation. Condition (Equation (8.13)) guarantees that each MS is able to decode the signalling message (of rate R_c) from the other MS. The remaining power $\nu(1-\beta)P$ is then employed for cooperative transmission to the BS.

It can be seen from Equation (8.12) that, if κ is large enough, it is optimal to invest all the power in the common message ($\beta \simeq 0$) and to

dedicate a vanishingly small portion of such power to signalling (and $\nu \simeq 1$). This way, the considered scheme approaches the upper bound Equation (8.12) set by the full-cooperation scenario.

8.3.2 An Achievable Rate

We now aim at extending the strategy presented above for a two-user Gaussian model to the Wyner model with in-band cooperation presented in Section 2.4.2 (see Figure 2.7). Using a block-based strategy similar to the one presented above, we obtain the following. As for the out-of-band strategy of the previous section, we define an integer $K_c > 0$, and a complex symmetric vector $\mathbf{h}_c = [h_{c,-K_c} \cdots h_{c,K_c}]^T$ with unit norm, $\|\mathbf{h}_c\|^2 = 1$, and its Fourier transform $H_c(\theta)$ as in Equation (8.6). As we will see, these quantities have analogous definitions as for the out-of-band strategy. Moreover, we recall that $H(\theta) = 1 + 2\alpha\cos(2\pi\theta)$ is the transfer function accounting for the operation of the Wyner model on the signals transmitted by the MSs. It is proved in [106] that the following per-cell rate is achievable:

$$R(P) = \max_{\beta,\nu_1,\nu_2,\mathbf{h}_c} \min \left\{ \begin{array}{l} \displaystyle\int_0^1 \log_2(1+\beta P H(\theta)^2)d\theta + R_c, \\ \displaystyle\int_0^1 \log_2(1+\beta P H(\theta) \\ \qquad +\nu_1(1-\beta)P|H_c(\theta)|^2 H(\theta)^2)df \end{array} \right\} \tag{8.14}$$

with $0 \le \beta,\ \nu_1, \nu_2 \le 1$ and

$$R_c = \min \left\{ \begin{array}{l} \displaystyle\frac{1}{2}\log_2\left(1+\frac{\kappa^2 2(1-\beta)P(1-\nu_1)(1-\nu_2)}{1+\kappa^2(2|h_{c,K_c}|^2(1-\beta)P\nu_1+2\beta P)}\right), \\ \displaystyle\frac{1}{2(K_c-1)}\log_2\left(1+\frac{\kappa^2(1-\beta)P(1-\nu_1)\nu_2}{1+\kappa^2(2|h_{c,K_c}|^2(1-\beta)P\nu_1+2\beta P)}\right), \\ \displaystyle\frac{1}{2K_c}\log_2\left(1+\frac{\kappa^2(1-\beta)P(1-\nu_1)(2-\nu_2)}{1+\kappa^2(2|h_{c,K_c}|^2(1-\beta)P\nu_1^2+2\beta P)}\right), \\ \displaystyle\frac{1}{2K_c-1}\log_2\left(1+\frac{\kappa^2 P_c(1-\nu_1)}{1+\kappa^2(2|h_{c,K_c}|^2(1-\beta)P\nu_1+2\beta P)}\right), \\ \displaystyle\frac{1}{1+K_c}\log_2\left(1+\frac{\kappa^2(1-\beta)(P/2)(1-\nu_1)(4-3\nu_2)}{1+\kappa^2(2|h_{c,K_c}|^2(1-\beta)P\nu_1+2\beta)P)}\right) \end{array} \right\}. \tag{8.15}$$

The rate (Equations (8.14) and (8.15)) is achieved via a block-based scheme that uses rate and power splitting in a way that resembles the scheme for the two-user Gaussian MAC described above. Specifically, as for that basic scenario, each MS divides its resources between transmission of private and common information, where the latter is transmitted cooperatively by the MSs to the BSs. Moreover, in order to enable cooperation, the MSs exchange signalling information about the common messages. The main issue is how to perform this task in an effective manner. As explained below, this can be done by using decode-and-forward techniques and exploiting the side information available at each MS regarding the signals generated at the MS itself or already decoded by it.

In Equations (8.14) and (8.15), the common-part power $(1-\beta)P$ is divided between the power used for cooperative transmission to the BSs, given by $\nu_1(1-\beta)P$, and the power used for signalling to other MSs so as to enable cooperation, given by $(1-\nu_1)(1-\beta)P$. The latter power is in turn split between the power employed to forward signalling information received from neighbors ($\nu_2(1-\beta)P(1-\nu_1)$) and locally generated common information ($(1-\nu_2)(1-\beta)P(1-\nu_1)$). We emphasize that the parameter ν_1 is especially critical as it accounts for the trade-off between power used for cooperative transmission to the BSs and that used for signalling among MSs. We also remark that, as shown in [106], selecting the common rate R_c as in Equation (8.15) guarantees correct decoding of the signalling messages at the MSs, while condition (Equation (8.14)) enables correct decoding at the CP.

It was noted above that for a two-user Gaussian MAC, if the MS measurements are of good enough quality, the scheme at hand is able to attain the upper bound corresponding to full cooperation. The same conclusion applies to the rate (Equations (8.12) and (8.13)) for the Wyner model at hand. In fact, if κ is large enough, by devoting all power to common message transmission ($\beta \simeq 0$) and dedicating a vanishingly small portion of such power to signalling ($\nu_1 \simeq 1$), one can get arbitrarily close to the upper bound Equation (8.1). The reasons for this are analogous to the ones provided above for out-of-band MS cooperation.

8.3.3 Numerical Results

In this section, we provide some numerical results to obtain insight into the performance and limitations of the achievable rate derived above. For comparison, we consider the rate (Equation (3.12)) achievable with no MS cooperation ($\kappa = 0$), which sets a lower bound, and the upper bound (Equation (8.1)) corresponding to ideal MS cooperation. Figure 8.4 shows the achievable rate (Equation (8.5)) versus the SNR P as compared to lower and upper bounds for $\alpha = 0.8$, $\kappa^2 = 20\,\mathrm{dB}$ and for different values of the number of cooperating terminals $K_c = 1, 2,$ and 3. It is noted that the optimal fraction of common power v_1 used for cooperative transmission decreases with K_c (not shown), as expected, since more power is required for signalling as the number of common messages to be delivered increases. We also remark that increasing the number of cooperating MSs beyond $K_c = 3$ is deleterious in terms of achievable rates in this example, due to the limitations in terms of resources for signalling. This is consistent with the analysis above for

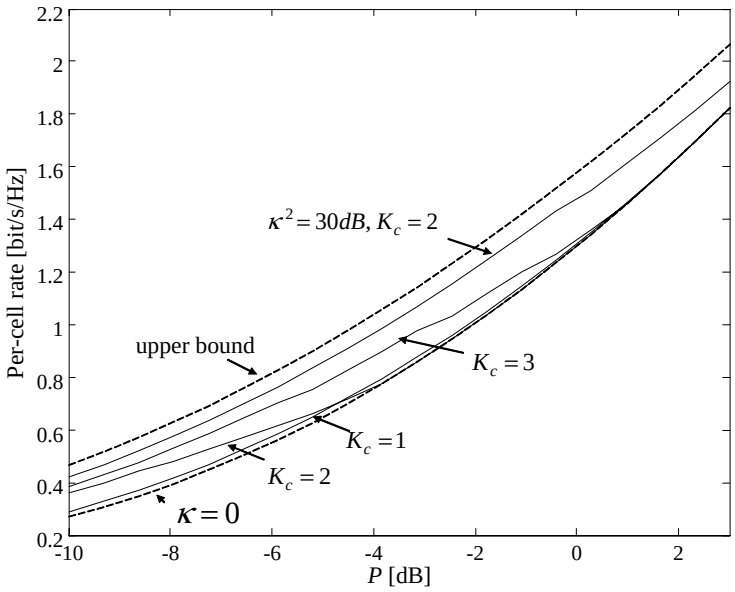

Fig. 8.4 Achievable per-cell rates and bounds versus the SNR P for in-band MS cooperation ($\alpha = 0.8$, $\kappa^2 = 20\,\mathrm{dB}$).

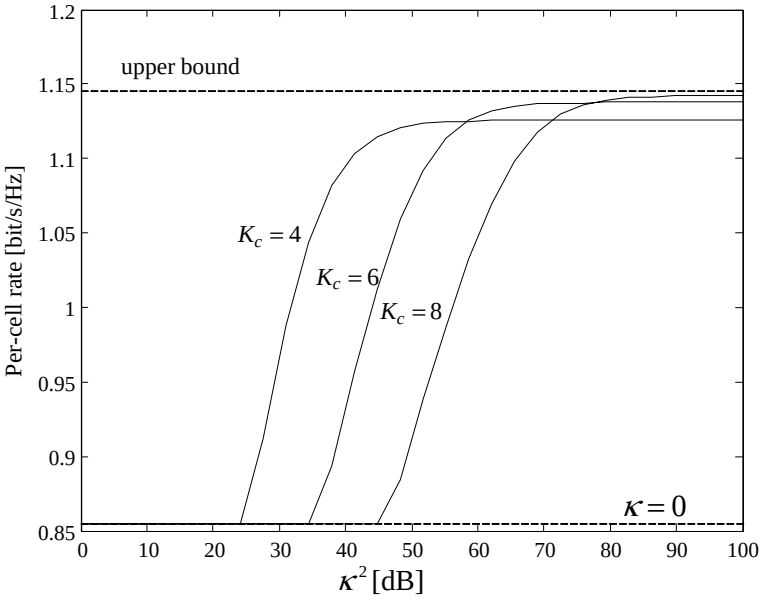

Fig. 8.5 Achievable per-cell rates and bounds versus the inter-MS channel gain κ for in-band MS cooperation ($\alpha = 0.6$ and $P = -2\,\text{dB}$).

out-of-band cooperation. Also shown is the case $\kappa^2 = 30\,\text{dB}$, $K_c = 2$. It is seen that, if κ^2 is sufficiently large, the proposed scheme enables relevant rate gains with respect to no cooperation, and allows the system to partially bridge the gap to the upper bound corresponding to full cooperation.

This fact is further investigated in Figure 8.5 where the rates discussed above are shown versus κ^2 for $\alpha = 0.6$ and $P = -2\,\text{dB}$. Figure 8.5 confirms that, for sufficiently large κ and K_c, the performance of the proposed scheme attains the upper bound of full cooperation.

8.4 Summary

In this section, we have studied the potential performance gains achievable with inter-cell MS cooperation and MCP for the uplink, assuming Gaussian channels. Via a low-SNR analysis, we have seen that, ideally, MS cooperation offers additional SNR (beamforming) gains on top of what is already achieved thanks to MCP. We have then investigated

to what extent such gains can be realized. Considering at first the availability of out-of-band inter-MS links, we have seen that relevant performance gains can be attained even with inter-MS links with capacities of the order of the desired per-cell sum-rates. We have then argued that, if the inter-MS links have to share the same bandwidth as the uplink transmission, then MS cooperation provides significant gains only in the presence of very strong inter-MS links.

9

Concluding Remarks

This monograph has investigated, via information-theoretic arguments, the advantages of cooperation in cellular systems. Cooperation among the BSs, or MCP, has been shown to be able to potentially increase the capacity of the network by an amount that depends critically on the inter-cell interference span. While initial work demonstrated these performance benefits under idealistic conditions, including absence of fading and unrestricted backhaul links, the work reviewed here has confirmed the promise of MCP under more practical conditions. The performance benefits of cooperation at the MS level have been reviewed as well, along with considerations regarding the strong interplay between the design of relaying strategies and of MCP techniques. The presentation has also briefly touched upon the potential gains achievable by exploiting novel transmission strategies such as structured codes [120].

As per our goals set forth in Section 1, the treatment has focused on simple models that allowed insightful conclusions to be reached via analysis. We remark that, while the models at hand may be considered naive, comparisons with more complex, and analytically involved, models has often found the insights obtained from the simple models at

hand valuable and reasonably accurate (see, e.g., [135]). In this regard, we emphasize that the models considered in this monograph overcome some of the limitations of the original Wyner-like models in [33, 132]. For instance, different users' locations can be accounted for, at least to some extent, by modifying appropriately the distribution of the fading channel gains – a fact that is sometimes overlooked in related literature. In the treatment, we have emphasized the role of architectural constraints, such as limitations on the backhaul links and on the relay deployment, on the performance gains achievable through cooperative techniques. The reviewed research activity, while admittedly incomplete, provides a clear picture of the potential of cooperation in cellular systems. However, it also leaves open a number of critical issues that more refined, but likely less compact, analyses need to address. We list below some of the current research challenges.

1. *More complex Wyner-type models*: While the analysis has focused on linear models, more complex scenarios such as planar models or models with random interfering spans are of interest in practice. We note that planar models have received some attention, starting with the original paper by Wyner [132] (see also [10, 99]), but many of the aspects covered in this text for linear models are yet to be addressed. We also remark that treatment of such models motivates further mathematical research in random matrix theory, e.g., on band random matrices.
2. *Channel State Information (CSI) availability*: CSI is known to constitute a bottleneck on the system performance in some operating regimes (see, e.g., [120]). For instance, CSI issues may affect the ability of the network to enlarge the size of cell clusters for MCP. Bounds on the capacity under channel uncertainty are then needed. Moreover, the implications of limited backhaul capacity on the collection and distribution of CSI are critical aspects yet to be explored. Novel concepts such as retrospective CSI, which allows "completely stale" CSI to be profitably used, are potentially of interest as well [70, 72, 134].

3. *Interference alignment*: The current state-of-the-art around the idea of interference alignment, summarized in [40], leaves open the question as to what is the potential role of such techniques in complex systems with practical limitations on CSI and coordination.
4. *MIMO cellular systems*: In the monograph, we have assumed single-antenna BSs and MSs. Extension of the analysis to multiantenna nodes is timely and of great practical interest. We note that the analysis of this setting is related to the study of DS-CDMA presented in Section 4.
5. *Structured codes*: While some initial effort in applying structured codes to cellular system has been reported in Section 5.1.4, further work is needed in order to assess the relevance of this technology under more general conditions.
6. *Green networking and energy efficiency*: Toward the goal of reducing the energy expenditure of the network [34], various techniques become of interest, such as cell-site duty cycling. Analysis of such scenarios could benefit from the techniques developed in [60, 115].

A

Gelfand–Pinsker Precoding and Dirty Paper Coding

Channels with random states known at the transmitter have attracted considerable attention by the research community in view of their central role in many settings of practical interest (see, e.g., [44] for an elaborate review). Here we review the so-called Gel'fand and Pinsker (GP) problem, along with the Dirty Paper Coding (DPC) technique.

A.1 GP and DPC

The capacity of the single-user memoryless channel with random parameters ("states") has been obtained by Gel'fand and Pinsker (GP) in [27], while assuming that the state sequence is completely known *noncausally* at the transmitter but unknown at the receiver. The corresponding (memoryless) channel is defined via the conditional probability $p(y|x,s)$, where X is the channel input, Y is the channel output, and S is the (input independent) state. The latter is assumed to be i.i.d. over channel uses with distribution $p(s)$. The capacity of this channel is given in the following form of an optimization problem [27] over the conditional distributions of an auxiliary random variable,

U, and the channel input given the interference sequence:

$$\mathcal{C} = \sup_{p(x,u|s)} \{I(U;Y) - I(U;S)\}. \tag{A.1}$$

Furthermore, it can be shown that it is enough to take the supremum over all joint distributions such that the channel input is a *deterministic function* of the auxiliary random variable and the interference [12], i.e.,

$$p(x,u|s) = 1\{x = f(u,s)\}\, p(u|s), \tag{A.2}$$

where $f(u,s)$ is some deterministic function.

A particular case of the GP setting is a channel of the form

$$Y = X + S + Z, \tag{A.3}$$

where $Z \sim \mathbb{N}(0,N)$ is additive Gaussian noise, $S \sim \mathbb{N}(0,Q)$ is additive *Gaussian* interference independent of Z, and the input is constrained by $E[X^2] \leq P$. This channel was considered by Costa in [15], and the setting was coined "writing on dirty paper". Costa's remarkable result was that the GP capacity (Equation (A.1)) for this channel is the same as if the *interference were not present*, i.e., $\frac{1}{2}\log(1 + \frac{P}{N})$. The capacity is obtained letting $p(u|s)$ be a Gaussian distribution $\mathbb{N}(\alpha S, P)$ and $f(u,s) = u - \alpha s$, where α is some constant. Costa showed that the optimum choice for α in this setting is $\alpha = \frac{P}{P+N}$. Thus, the proposed coding technique, commonly referred to as *"dirty paper coding"*, completely eliminates the impact of interference, while being fully known noncausally at the transmitter.

It was later shown in [14] that the same rate can be achieved for arbitrary noise distribution, provided that the interference is Gaussian i.i.d., or for arbitrary interference distribution provided that the noise is Gaussian (possibly colored). This was extended in [23, 147] to arbitrary interference (arbitrary interference statistics, or even arbitrary interference sequences, where the transmitter knows the individual sequence but ignores its statistics), provided that the transmitter and receiver share a common dither signal. The DPC scheme was extended to vector Gaussian interference channels in [139, 140]. The choice of the constant α has also been studied from a variety of standpoints (see, e.g., [25, 26]).

A.2 Application to Broadcast Channels

We now consider the application of DPC to a MIMO Gaussian broadcast channel (GBC). We keep the discussion general and consider a MIMO GBC, in which the transmitter has M antennas and the k-th user, $k = 1, \ldots, K$, has r_k antennas. The $r_k \times 1$ signal received by the k-th user can be written as

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x} + \mathbf{z}_k, \tag{A.4}$$

where $\mathbf{x}$ is the $M \times 1$ transmitted signal, $\mathbf{H}_k \in \mathbb{C}^{r_k \times M}$ denotes the channel transfer matrix from the transmitter to the receiver of user k, and $\mathbf{z_k}$ is the vector of uncorrelated unit-power complex Gaussian noises affecting user k, $k = 1, \ldots, K$. We can rearrange the received signals in a matrix form as

$$\mathbf{y} = \begin{bmatrix} \mathbf{H}_1 \\ \vdots \\ \mathbf{H}_K \end{bmatrix} \mathbf{x} + \begin{bmatrix} \mathbf{z}_1 \\ \vdots \\ \mathbf{z}_K \end{bmatrix} \triangleq \mathbf{H}\mathbf{x} + \mathbf{z}. \tag{A.5}$$

We denote the covariance matrix of the transmitted signal as $\mathbf{\Sigma}_x \triangleq E\left\{\mathbf{x}\mathbf{x}^\dagger\right\}$, and assume that the transmitter is subject to an average power constraint $\operatorname{Tr} \mathbf{\Sigma}_x \leq P$. The focus in the following is first on the case in which the channel transfer matrix $\mathbf{H}$ is fixed throughout the transmission, while assuming full channel state information is available to the transmitter and the receivers. Note that in the MCP scenarios considered in the text (see Equation (2.6)), the transmitter consists of the M BSs, we have $r_k = 1$ and the number of users is KM. The use of K here for the total number of users should not create confusion.

The application of the DPC scheme for the MIMO GBC has been suggested by Caire and Shamai in [12]. Considering first for simplicity the two-user case ($K = 2$), the idea is to use Gaussian coding for encoding the information of user 2, and then encode the information of user 1 using the DPC scheme while treating the (encoded) signal to be transmitted to user 2 as an additive interference signal known noncausally at the encoder. The actual channel input (the transmitted signal) is the superposition of the encoded signals of the two users. At the receiving ends, user 2 decodes the Gaussian code and suffers from

an additional Gaussian interference, due to the signal transmitted to user 1. User 1, on the other hand, due to the DPC scheme employed at the transmitter, effectively experiences no interference and decodes its data as if the signal intended for user 2 were not present. Obviously, the role of the users can be reversed, as the particular encoding order affects the achievable rates of each. This successive encoding scheme can be naturally extended to more than two users, as discussed below.

Consider the general MIMO GBC introduced above. Based on the extension of DPC to vector channels in [139], the DPC achievable region for the MIMO GBC can be defined as (see Equation (A.4))

$$\mathcal{R}_{\mathrm{DPC}}(P, \mathbf{H}_{1,\ldots,K}) = \mathrm{conv}\left\{ \bigcup_{\pi\in\Pi} \bigcup_{\mathbf{\Sigma}_1,\ldots,\mathbf{\Sigma}_K} (R_1^{\mathrm{DPC}}(\pi, \mathbf{\Sigma}_{1,\ldots,K}, \mathbf{H}_{1,\ldots,K}), \ldots, R_K^{\mathrm{DPC}}(\pi, \mathbf{\Sigma}_{1,\ldots,K}, \mathbf{H}_{1,\ldots,K})) \right\}, \tag{A.6}$$

where Π is the set of all permutations on $\{1,\ldots,K\}$, and the permutation $\pi \in \Pi$ corresponds to a reverse encoding order such that user π_{K+1-j} is the j-th user to be encoded. The unions in Equation (A.6) are taken over all possible encoding orders, and over all power allocation matrices $\mathbf{\Sigma}_k \triangleq E\left\{\mathbf{x}_k\mathbf{x}_k^\dagger\right\} \succeq 0\ \forall k$ (with $\mathbf{x}_k$ representing the transmit vector corresponding to user k, and $\mathbf{x} = \sum_{k=1}^K \mathbf{x}_k$), such that $\mathrm{Tr}(\sum_{k=1}^K \mathbf{\Sigma}_k) \le P$. The rates R_{ik}^{DPC} are defined through

$$R_k^{\mathrm{DPC}}(\pi, \mathbf{\Sigma}_{1,\ldots,K}, \mathbf{H}_{1,\ldots,K}) = \log \frac{\det\left(\mathbf{H}_{\pi_k}(\sum_{j=1}^k \mathbf{\Sigma}_{\pi_j})\mathbf{H}_{\pi_k}^\dagger + \mathbf{I}\right)}{\det\left(\mathbf{H}_{\pi_k}(\sum_{j=1}^{k-1} \mathbf{\Sigma}_{\pi_j})\mathbf{H}_{\pi_k}^\dagger + \mathbf{I}\right)} \quad k = 1,\ldots,K. \tag{A.7}$$

The key importance of the DPC region stems from the fact that it determines the entire capacity region of the MIMO GBC, as was proved in [130]. The result therein in fact extends the optimality of the DPC region to more general input constraints than the sum-power constraint, as stated as follows. Let $\mathcal{S}$ be a compact set of positive semidefinite $M \times M$ matrices and consider the MIMO GBC in Equation (A.4) with

input constraint $E\left\{\mathbf{x}\mathbf{x}^{\dagger}\right\} \preceq \mathbf{S}$ for some $\mathbf{S} \in \mathcal{S}$. Then, the capacity region of this channel is given by

$$\mathcal{C}_{\mathrm{BC}}(\mathcal{S}, \mathbf{H}_{1,\ldots,K}) = \bigcup_{\mathbf{S}\in\mathcal{S}} \mathcal{R}_{\mathrm{DPC}}(\mathbf{S}, \mathbf{H}_{1,\ldots,K}). \tag{A.8}$$

So far, the discussion has assumed the channels composing the transfer matrix $\mathbf{H}$ in Equation (A.4) to be fixed. When the channels are time-varying and ergodic, while retaining the assumptions of full channel state information at the transmitter and receivers, the ergodic capacity region of the MIMO GBC is obtained by averaging Equation (A.8) with respect to the channel statistics (while modifying the constraints to be on the average input constraints). It is noted here that this motivates the search for power allocation schemes for the ergodic time varying channel. Some aspects of the problem are investigated for example in [41, 42, 46, 76], and references therein.

B

Uplink and Downlink Duality

Here we provide a brief introduction to the concept of uplink–downlink duality. For generality, we focus on the MIMO GBC introduced in Appendix A. Although the capacity region of the MIMO GBC is known as discussed in Appendix A, some considerable difficulties still remain. More specifically, the DPC achievable region specified in Equation (A.6) has no closed-form solution for its boundary points. Also, since the per-user rates $R_i^{\mathrm{DPC}}(\pi, \mathbf{\Sigma}_{1,\ldots,K}, \mathbf{H}_{1,\ldots,K})$ in Equation (A.7) are neither concave nor convex, the direct numerical analysis of the boundaries of the DPC achievable region is a formidable task. To overcome this problem, it is useful to consider a *dual* MAC, and exploit the duality properties of the MAC and broadcast channel (BC), referred to also as "uplink–downlink duality".

The *dual uplink* is defined as a MAC with K users, equipped with r_k $(k = 1, \ldots, K)$ antennas, each transmitting to a receiver equipped with M antennas. The channel transfer matrices between the k-th user and the receiver are given by $\mathbf{H}_k^{\dagger}$, where $\mathbf{H}_k$, $k = 1, \ldots, K$, are the channel transfer matrices of the original BC as defined in Equation (A.5). The

general channel model for this dual uplink is given by

$$\tilde{\mathbf{y}} = \sum_{k=1}^{K} \mathbf{H}_k^{\dagger} \tilde{\mathbf{x}}_k + \tilde{\mathbf{z}}, \tag{B.1}$$

where the notation $\tilde{(\cdot)}$ is used to designate quantities corresponding to the dual uplink. Here $\tilde{\mathbf{z}}$ denotes the $M \times 1$ zero-mean complex Gaussian noise at the dual uplink's receiver ($\tilde{\mathbf{z}} \sim \mathbb{N}_c(\mathbf{0}, \mathbf{I}_{[M \times M]})$), and $\tilde{\mathbf{x}}_k$, $k = 1, \ldots, K$, are the channel inputs due to the K users. Assume the following *individual* covariance matrix constraint $E\left\{\mathbf{x}_k \mathbf{x}_k^{\dagger}\right\} \preceq \tilde{\mathbf{P}}_k$, $k = 1, \ldots, K$. Then, using the fact that the vertices of the MAC's capacity region are achieved by Gaussian codes and successive decoding, the capacity region of the dual uplink can be written as

$$\begin{aligned}
&\tilde{\mathcal{C}}_{\mathrm{MAC}}(\tilde{\mathbf{P}}_{1,\ldots,K}, \mathbf{H}_{1,\ldots,K}^{\dagger}) \\
&\quad = \mathrm{conv}\left\{ \bigcup_{\pi \in \Pi} \{(R_1, \ldots, R_K) : R_k = R_k^{\mathrm{MAC}}(\pi, \tilde{\mathbf{P}}_{1,\ldots,K}, \mathbf{H}_{1,\ldots,K}^{\dagger}) \forall i\} \right\},
\end{aligned} \tag{B.2}$$

where π is the decoding order such that π_j corresponds to the j-th user to be decoded (note the difference with respect to Equation (A.6)), and where

$$\begin{aligned}
&R_i^{\mathrm{MAC}}(\pi, \tilde{\mathbf{P}}_{1,\ldots,K}, \mathbf{H}_{1,\ldots,K}^{\dagger}) \\
&\quad = \log\det\left(\mathbf{I} + \left(\sum_{j=i+1}^{K} \mathbf{H}_{\pi_j}^{\dagger} \tilde{\mathbf{P}}_{\pi_j} \mathbf{H}_{\pi_j} + \mathbf{I} \right)^{-1} \mathbf{H}_{\pi_i}^{\dagger} \tilde{\mathbf{P}}_{\pi_i} \mathbf{H}_{\pi_i} \right).
\end{aligned} \tag{B.3}$$

Consider now a MAC with a *total* power constraint $\sum_{k=1}^{K} \mathrm{Tr}\left(E\left\{\mathbf{x}_k \mathbf{x}_k^{\dagger}\right\}\right) \leq P$ (i.e., a constraint on the sum of the powers of all users). In this case, the capacity region of the MAC is given by [138]

$$\mathcal{C}_{\mathrm{Union}}(P, \mathbf{H}_{1,\ldots,K}^{\dagger}) = \bigcup_{\mathrm{Tr}\left(\sum_{k=1}^{K} \tilde{\mathbf{P}}_k\right) \leq P} \tilde{\mathcal{C}}_{\mathrm{MAC}}(\tilde{\mathbf{P}}_{1,\ldots,K}, \mathbf{H}_{1,\ldots,K}^{\dagger}), \tag{B.4}$$

where the union is over all matrices $\tilde{\mathbf{P}}_k \succeq 0 \ \ \forall k$, such that $\mathrm{Tr}\left(\sum_{k=1}^{K} \tilde{\mathbf{P}}_k\right) \leq P$. The duality between this capacity region for the

MIMO MAC of Equation (A.5) and the DPC achievable region for the BC of Equation (A.5), was established in [127, 128] and can be stated as

$$\mathcal{C}_{\text{Union}}(P, \mathbf{H}_{1,\ldots,K}^{\dagger}) = \mathcal{R}_{\text{DPC}}(P, \mathbf{H}_{1,\ldots,K}), \tag{B.5}$$

where $\mathcal{R}_{\text{DPC}}(P, \mathbf{H}_{1,\ldots,K})$ is given in Equation (A.6).

In addition to the general statement of capacity region duality presented above, a MAC-to-BC transformation is proposed in [127]. The transformation finds, for any set of input covariance matrices, and a given decoding order π of the users in the MAC while assuming successive cancellation at the receiver, a set of BC transmit covariance matrices (one per each of the users) with the same sum power as that of the MAC input covariance matrices, that achieve the same rate vector in the BC with the reverse encoding order π (and vice versa). That is, the transformation guarantees that $R_k^{\text{MAC}}(\pi, \tilde{\mathbf{P}}_{1,\ldots,K}, \mathbf{H}_{1,\ldots,K}^{\dagger}) = R_k^{\text{DPC}}(\pi, \mathbf{\Sigma}_{1,\ldots,K}, \mathbf{H}_{1,\ldots,K})$ (note again the different interpretation of π in both cases). The duality relation of Equation (B.5) enables the characterization of the boundaries of the DPC region by means of the convex dual MAC capacity region. The boundaries of $\mathcal{C}^{\text{Union}}(P, \mathbf{H}_{1,\ldots,K}^{\dagger})$ can be calculated using interior-point methods [9].

The uplink–downlink duality principle turns out to be particularly useful for obtaining the MIMO GBC *sum-rate* capacity. In fact, results for the sum-rate capacity of the MIMO GBC preceded the derivation of the full capacity region. It was already shown in [12, 127, 128, 140] that the DPC strategy achieves the sum-rate capacity of the MIMO GBC, that is,

$$\mathcal{C}_{\text{BC}}(P, \mathbf{H}_{1,\ldots,K}) = \max_{R_{1,\ldots,K} \in \mathcal{R}_{\text{DPC}}(P, \mathbf{H}_{1,\ldots,K})} \sum_{k=1}^{K} R_k. \tag{B.6}$$

The above result, in the most general form, was proved using the Sato upper bound [88] (an approach originally suggested in [12]) in order to give an upper limit on the sum-rate capacity, and then showing that this upper bound is achievable with the DPC scheme. Accordingly, the capacity of the cooperative channel (i.e., allowing the users' receivers to cooperate in order to decode a single message) is used as an upper bound on the sum-rate capacity of the BC. Since in the BC the users

cannot cooperate, the actual sum-rate capacity of the GBC is independent of the correlations between the noise vectors at the receivers of each of the users. The sum-rate capacity is therefore bounded by the corresponding capacity of the cooperative channel with the "least favorable" noise vector $\mathbf{z}$, with structure as in Equation (A.5), where the marginals of the vectors $\{\boldsymbol{z}_k\}$ have identity (or smaller) covariance matrices (as originally defined in Equation (A.5)). We note here that the proof in [140] does not rely on uplink–downlink duality, but rather on a decision-feedback equalization approach where the decision-feedback equalizer is employed as a joint receiver in the cooperative BC (an approach that enabled the proof to hold not only for a sum-power constraint but for general convex input constraints).

Particularizing, however, to the case of single-antenna users, then using the uplink–downlink duality principle, the downlink sum-rate capacity can be compactly formulated by means of the following maximization problem [127, 128]:

$$\mathcal{C}_{\mathrm{BC}}(P) = \sup_{\mathbf{D}\in\mathcal{A}} \log\det(\mathbf{I} + \mathbf{H}^{\dagger}\mathbf{D}\mathbf{H}), \tag{B.7}$$

where $\mathcal{A}$ is the set of $K \times K$ diagonal matrices $\mathbf{D}$ with $\mathrm{Tr}\,\mathbf{D} \leq P$. As can be observed the $\log\det(\cdot)$ expression to be maximized is the expression for the sum-rate capacity of the single-user MIMO channel, with the difference that the maximization is over *diagonal* input covariance matrices, which is due to the fact that no user cooperation can be assumed in the dual uplink.

The sum-rate expressions of Equations (B.6) and (B.7) address the case in which the downlink is subject to a *total* input power constraint. However, with MCP, one has to deal with individual per-antenna power constraints. In [141] a connection is established between the duality of the Gaussian vector MAC and BC, and the Lagrangian duality in minimax optimization (the reader is referred to [137] for an elaborate discussion in this framework). This new minimax duality allows the optimal transmit covariance matrix and the least-favorable noise for the BC to be characterized in terms of the dual variables. Further, it allows BC–MAC duality to be generalized to BCs with arbitrary linear constraints. In particular, it is shown that the *sum-rate capacity* of the Gaussian multi-antenna BC, with individual per-antenna power

constraints $[\mathbf{\Sigma}_x]_{k,k} \leq P_k$, is the same as the sum capacity of a dual MAC with a sum power constraint, and with a diagonal and uncertain noise:

$$\min_{\mathbf{\Lambda}_x} \max_{\mathbf{\Lambda}_z} \log \frac{\det(\mathbf{\Lambda}_z + \mathbf{H}^\dagger \mathbf{\Lambda}_x \mathbf{H})}{\det(\mathbf{\Lambda_z})}, \tag{B.8}$$

where $\mathbf{\Lambda}_x$ and $\mathbf{\Lambda}_z$ are diagonal with $\operatorname{Tr} \mathbf{\Lambda}_x \leq 1$, and $\sum_k P_k [\mathbf{\Lambda}_z]_{k,k} \leq 1$ (note that the per-antenna power constraints are incorporated in the noise constraint in the dual problem). This reduces to Equation (3.20) for the scenario of interest in Section 2. We refer to [148] for further results on uplink-downlink duality.

C

The CEO Problem

The Chief Executive Officer (CEO) problem refers to the distributed source coding setting of Figure C.1. The formulation consists of a single random source X which is measured by many "agents" subject to some noise. The agents need to forward their observations to the CEO. The CEO wishes to use the observations in order to estimate the source X. The agents communicate to the CEO through unidirectional links from each of them to the CEO. These links have finite capacities. The objective is to minimize the average distortion of the estimation at the CEO, given the finite capacities of the links between the agents and the CEO. Note that, when the links are of infinite capacity, the problem reduces to a simple estimation problem.

For a more formal description of the problem, let X^n be an i.i.d. source of n symbols, and let the received signals at M agents be $\{Y_i^n\}_{i=1}^M$. The received signals are generated in an i.i.d. fashion according to the conditional probability $P_{Y_i|X}(y_i|x)$. All agents are connected to the CEO with links with finite capacities of $\{C_i\}_{i=1}^M$ bits per source symbol, respectively. The estimate $\hat{X}^n$ at the receiver is produced based on the messages received from the agents. The objective is to minimize the average distortion between X^n and $\hat{X}^n$ for some distortion metric.

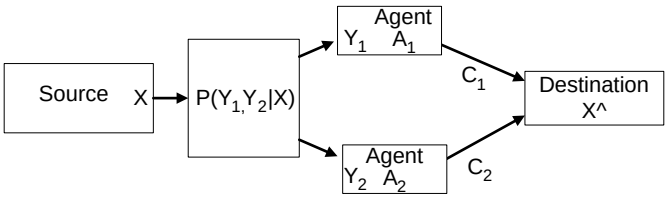

Fig. C.1 The CEO setting. A source X is observed by two agents through a memoryless channel. The agents process their measurements and forward messages to the destination, which is referred to as the CEO. The CEO reproduces the estimate $\hat{X}$. Note that the agents cannot cooperate in sending their messages to the CEO.

The CEO problem was introduced by Berger et al. [6], where both upper and lower bounds on the minimum sum-rate required to obtain a given distortion level were derived. In particular, [6] showed that, when the number of agents is taken to infinity, i.e., $M \to \infty$, the distortion is nonzero even if the sum-rate is larger than the entropy of X, that is, $\sum_{i=1}^{M} C_i > H(X)$. This conclusion demonstrates the loss due to the fact that the physically separated agents are not able to cooperate before forwarding their messages to the CEO. In fact, if the agents could pool together their observations, as $M \to \infty$, they could first estimate X to an arbitrary degree of accuracy (e.g., for additive noise channels $P_{Y_i|X}(y_i|x)$), and then convey the latter to the CEO with sum-rate equal to $H(X)$.

The compression scheme proposed in [6] employs random coding, in which the signal Y_i^n received by the i-th agent is compressed using a given codebook U_i^n. This compression step allows Y_i^n to be represented as U_i^n within some average distortion. This step is sufficient to achieve optimality (in terms of minimum required rate for a given distortion level) in the case of a single agent. But, in the case of more than one agent, the compressed codewords from the agents are typically dependent, due to the correlation of the received signals Y_i^n. Therefore, additional compression gains can be acquired by leveraging such correlation. A standard way to do this is via partitioning or binning of the compression codebook U_i^n. The basic idea is that each agent will inform the CEO only about the partition (or bin) in which the compressed codeword U_i^n is located. The CEO will then recover all the codewords U_i^n for $i = 1, \ldots, M$ via joint decoding by leveraging the signal

correlations. This binning operation leads to optimal performance in some scenarios, such as lossless compression [109] and lossy compression with side information [133]. It is, however, not necessarily optimal for the CEO problem. Note that the CEO estimates the source letter-by-letter using classical estimation theory for i.i.d. sources, based on the codewords U_i^n for $i = 1, \ldots, M$.

Reference [6] mainly considered discrete-alphabet signals. Further work, such as [80, 129], extended several derivations done in [6] to continuous-alphabet signals with quadratic distortion measure and Gaussian statistics. A full solution for this setting is given in [81] and [83]. The CEO problem has also been treated in many other papers from different standpoints. For example, reference [146] derived a lattice based approach to the CEO problem and [71, 129] presented practical approaches for implementing CEO systems.

D

Some Results from Random Matrix Theory

In this appendix, we provide a brief introduction to some of the results of random matrix theory that have found application to the study of cellular systems, as reviewed in this monograph. The treatment is based on [121], to which we refer for further discussion.

D.1 Preliminaries

Most of the literature that studies various aspects of cellular systems eventually deals with memoryless linear vector channels of the form

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \tag{D.1}$$

where $\mathbf{x}$ is the K-dimensional input vector, $\mathbf{y}$ is the N-dimensional output vector, the N-dimensional vector $\mathbf{n}$ models an additive zero-mean circularly symmetric Gaussian noise, and $\mathbf{H}$ is the $N \times K$ random channel transfer matrix. All the above quantities are in general complex-valued. The above model applies to a wide variety of system settings incorporating in particular the impacts of fading, multi-antenna, and multicell systems, as discussed in this monograph.

In addition to aspects such as input constraints, the information available at either transmitter(s) or receiver(s) and the level of

cooperation between them, system performance in the different settings that can be described by Equation (D.1) heavily depends on the statistical properties of the entries of $\mathbf{H}$. In particular, as discussed in Section 3 (Section 3.1.2), a key tool for system performance analysis in such cases is the characterization of the singular value distribution of $\mathbf{H}$, or of related quantities.

To see that, let $\mathsf{F}^n_{\mathbf{A}}$ denote the *empirical* cumulative distribution function of the eigenvalues of an $n \times n$ Hermitian matrix $\mathbf{A}$ (also referred to as the "spectrum", or "empirical distribution" [121]). It is defined as

$$\mathsf{F}^n_{\mathbf{A}}(x) = \frac{1}{n}\sum_{i=1}^{n} 1\{\lambda_i(\mathbf{A}) \leq x\}, \tag{D.2}$$

where $\lambda_1(\mathbf{A}), \ldots, \lambda_n(\mathbf{A})$ are the eigenvalues of $\mathbf{A}$ and $1\{\cdot\}$ is the indicator function. This definition was used in Equation (3.11), but here we emphasize its dependence on the dimension of the matrix. If $\mathsf{F}^n_{\mathbf{A}}$ converges as $n \to \infty$, then the corresponding limit (also referred to as the *asymptotic* spectrum, or *asymptotic* empirical distribution) is denoted by $\mathsf{F}_{\mathbf{A}}$ [121].

Assume now that the entries of the input vector $\mathbf{x}$ are i.i.d. and Gaussian, then the normalized input–output mutual information of (D.1) conditioned on the channel matrix $\mathbf{H}$ is

$$\begin{aligned}\frac{1}{N} I(\mathbf{x};\mathbf{y}|\mathbf{H}) &= \frac{1}{N}\log\det(\mathbf{I} + \mathsf{SNR}\,\mathbf{H}\mathbf{H}^\dagger) \\ &= \frac{1}{N}\sum_{i=1}^{N}\log(1 + \mathsf{SNR}\,\lambda_i(\mathbf{H}\mathbf{H}^\dagger)) \\ &= \int_0^\infty \log(1 + \mathsf{SNR}\,x)\cdot F^N_{\mathbf{H}\mathbf{H}^\dagger},\end{aligned} \tag{D.3}$$

where SNR denotes the transmit SNR,

$$\mathsf{SNR} = \frac{NE\{\|\mathbf{x}\|^2\}}{KE\{\|\mathbf{n}\|^2\}}, \tag{D.4}$$

and $\lambda_i(\mathbf{H}\mathbf{H}^\dagger)$ denotes the i-th squared singular value of $\mathbf{H}$. Further assuming that the channel is stationary and ergodic, and that its realizations are known at the receiver, the expectation of Equation (D.3)

with respect to $\mathbf{H}$ gives the ergodic channel capacity (*per degree of freedom*). Alternatively, treating the mutual information in Equation (D.3) as a random variable, its distribution specifies the outage capacity [8].

Another performance measure of practical interest is the MMSE achieved by *linear* processing of the received signals, which also determines the maximum achievable output SINR. For an arbitrary $N \times K$ matrix $\mathbf{H}$, recall that the nonzero singular values of $\mathbf{H}$ and $\mathbf{H}^\dagger$ are identical, and hence [121]

$$N\mathsf{F}^N_{\mathbf{H}\mathbf{H}^\dagger} - N\mathscr{U}(x) = K\mathsf{F}^K_{\mathbf{H}^\dagger\mathbf{H}} - K\mathscr{U}(x), \tag{D.5}$$

where $\mathscr{U}(x)$ is the unit-step function ($\mathscr{U}(x) = 0,\ x \le 0\ ;\ \mathscr{U}(x) = 1,\ x > 0$). It then follows that for i.i.d. inputs, the arithmetic mean of the MMSE (which is a function of $\mathbf{H}$) is given by [124]

$$\begin{aligned}
\frac{1}{K}\min_{\mathbf{M}\in\mathbb{C}^{K\times N}} E\left\{\|\mathbf{x}-\mathbf{M}\mathbf{y}\|^2\right\} &= \frac{1}{K}\mathrm{Tr}((\mathbf{I}+\mathsf{SNR}\mathbf{H}^\dagger\mathbf{H})^{-1}) \\
&= \frac{1}{K}\sum_{i=1}^{K}\frac{1}{1+\mathsf{SNR}\,\lambda_i(\mathbf{H}^\dagger\mathbf{H})} \\
&= \int_0^\infty \frac{1}{1+\mathsf{SNR}\,x}\cdot F^K_{\mathbf{H}^\dagger\mathbf{H}} \\
&= \frac{N}{K}\int_0^\infty \frac{1}{1+\mathsf{SNR}\,x}\cdot F^N_{\mathbf{H}\mathbf{H}^\dagger} - \frac{N-K}{K},
\end{aligned} \tag{D.6}$$

where the expectation is with respect to $\mathbf{x}$ and $\mathbf{n}$, and the last equality follows from Equation (D.5). The two performance measures are coupled through the following relation (see an elaboration in [31]):

$$\mathsf{SNR}\frac{\partial}{\partial\mathsf{SNR}}\log\det(\mathbf{I}+\mathsf{SNR}\mathbf{H}\mathbf{H}^\dagger) = K - \mathrm{Tr}((\mathbf{I}+\mathsf{SNR}\mathbf{H}^\dagger\mathbf{H})^{-1}). \tag{D.7}$$

The above examples let us conclude that the empirical distribution of the *squared* singular values of the channel matrix $\mathbf{H}$ determines both the capacity and the MMSE with linear processing for the channel in Equation (D.1). In the simple case in which the entries of $\mathbf{H}$ are i.i.d. and Gaussian, the expected empirical distribution of

the squared singular values can be explicitly expressed. In many other cases of interest useful results and insights can be obtained by focusing on the asymptotic regime, in which the number of rows and columns of $\mathbf{H}$ both go to infinity, while their ratio goes to some arbitrary constant. In this regime, the empirical (squared) singular value distribution converges (for certain cases of interest) to a limiting deterministic distribution. Furthermore, the convergence of the asymptotic distribution is insensitive to the actual probability distribution function (p.d.f.) of the random entries of $\mathbf{H}$, the results exhibit an ergodic behavior, and the convergence to the asymptotic limit is fast. Analyses based on asymptotic random matrix theory can therefore be considered to provide good predictions of the performance of practical *finite dimensional* systems, a fact that has motivated the extensive use of these tools in recent years.

In the following subsections, a representative sample of useful random matrix theory results is briefly reviewed, focusing on those results mostly relevant to the problems investigated in this monograph. The review in this appendix is based heavily on [121], which includes an outstanding comprehensive summary of the state-of-the-art in this field. This appendix is provided mainly to make the current monograph more self-contained. The reader is referred to [121] and references therein for proofs and a more detailed description of the results. A related reference is also [16].

D.2 Transforms

In many wireless communications problems of interest the limiting spectrum of the random matrix of concern cannot be obtained in an explicit form, but rather in terms of certain transforms of its distribution. Fortunately, these transforms can be related to system performance measures of interest, and therefore provide useful analytical tools. In the following, some key transforms are briefly reviewed.

D.2.1 Stieltjes Transform

The Stieltjes transform of a cumulative distribution function $F_X(\cdot)$ of a real-valued random variable X is defined by the following function of

complex arguments:

$$\mathcal{S}_X(z) = E\left\{\frac{1}{X - z}\right\} = \int_{-\infty}^{\infty} \frac{1}{\lambda - z}\, dF_X(\lambda). \tag{D.8}$$

It is customary to restrict the domain of $\mathcal{S}_X(z)$ to arguments having positive imaginary parts. Also, by Equation (D.8) the signs of the imaginary parts of z and $\mathcal{S}_X(z)$ coincide. The Stieltjes transform has the following inversion formula:

$$f_X(\lambda) = \lim_{\omega \to 0^+} \frac{1}{\pi} \mathrm{Im}\left[\mathcal{S}_{\mathrm{X}}(\lambda + \mathrm{j}\omega)\right] \tag{D.9}$$

where $f_X(\lambda)$ is the p.d.f. of the random variable X.

D.2.2 η-Transform

The η-transform of a nonnegative random variable X is

$$\eta_X(\gamma) = E\left\{\frac{1}{1 + \gamma X}\right\} \tag{D.10}$$

where γ is a nonnegative real number (thus $\eta_X(\gamma) \in (0, 1]$). The η-transform has an engineering interpretation. In particular, in many cases of interest, such as multiuser systems, the η-transform gives the ratio of the expected SINR at the output of a linear MMSE (multiuser) receiver, and the single-user SNR, corresponding to the notion of *multiuser efficiency* in multiuser detection [124]. The η-transform is also related to the Stieltjes transform through

$$\eta_X(\gamma) = \frac{\mathcal{S}_X\left(-\frac{1}{\gamma}\right)}{\gamma}, \tag{D.11}$$

which either requires analytic continuation or the inclusion of the negative real line in the domain of $\mathcal{S}_X(\cdot)$ [121].

D.2.3 Shannon Transform

The Shannon transform of a nonnegative random variable X is defined as

$$\mathcal{V}_X(\gamma) = E\left\{\log(1 + \gamma X)\right\}, \tag{D.12}$$

where γ is a nonnegative real number. The Shannon transform is related to the Stieltjes and η-transform through [121]

$$\begin{aligned}\frac{\gamma}{\log e}\frac{\partial}{\partial\gamma}\mathcal{V}_X(\gamma) &= 1-\frac{1}{\gamma}\mathcal{S}_X\left(-\frac{1}{\gamma}\right)\\ &= 1-\eta_X(\gamma).\end{aligned} \tag{D.13}$$

The Shannon transform of the empirical distribution of the eigenvalues of $\mathbf{HH}^\dagger$ gives the ergodic capacity of the Gaussian linear vector channel in Equation (D.1) for i.i.d. inputs (see Equation (D.3)).

D.3 Asymptotic Results

The following subsections include some of the main results on the limiting empirical eigenvalue distribution of random matrices of particular relevance to this monograph.

D.3.1 The Marčenko–Pastur Law and Its Generalizations

Theorem D.1 [75, 121, Theorem 2.35]. Consider an $N\times K$ matrix $\mathbf{H}$ whose entries are independent zero-mean complex (or real) random variables with variances $\frac{1}{N}$ and fourth moments of order $O(\frac{1}{N^2})$. As $K,N\to\infty$ with $\frac{K}{N}\to\beta$, the empirical distribution of $\mathbf{H}^\dagger\mathbf{H}$ converges almost surely to a nonrandom limiting distribution having density function

$$\mathsf{f}_\beta(x)=\left(1-\frac{1}{\beta}\right)^+\delta(x)+\frac{\sqrt{(x-a)^+(b-x)^+}}{2\pi\beta x} \tag{D.14}$$

where $(z)^+=\max(0,z)$ and

$$a=(1-\sqrt{\beta})^2,\quad b=(1+\sqrt{\beta})^2\quad .$$

The distribution in Equation (D.14) gives the Marčenko–Pastur law with ratio index β. It is noted that the zero-mean condition for the entries of $\mathbf{H}$ can be relaxed to entries having identical means, and the condition on their fourth moments can be relaxed to a Lindeberg-type

condition ([4, Theorem 2.8]):

$$\frac{1}{\delta^2 K}\sum_{i,j} E\left\{|[\mathbf{H}]_{i,j}|^2\, 1\{|[\mathbf{H}]_{i,j}| \geq \delta\}\right\} \to 0 \quad \forall \delta > 0. \tag{D.15}$$

The limiting empirical distribution of the eigenvalues of $\mathbf{H}\mathbf{H}^\dagger$ converges almost surely to a nonrandom limit whose density function is

$$\begin{aligned}\tilde{\mathsf{f}}_\beta(x) &= (1-\beta)\delta(x) + \beta \mathsf{f}_\beta(x) \\ &= (1-\beta)^+\delta(x) + \frac{\sqrt{(x-a)^+(b-x)^+}}{2\pi x}.\end{aligned} \tag{D.16}$$

Theorem D.2 [75, 101, 121, Theorem 2.38]. Let $\mathbf{H}$ be an $N \times K$ matrix whose entries are i.i.d. complex random variables with zero means and variances $\frac{1}{N}$. Further, let $\mathbf{T}$ be a $K \times K$ real diagonal random matrix whose empirical eigenvalue distribution converges almost surely to the distribution of a random variable T, and let $\mathbf{W}_0$ be an $N \times N$ Hermitian complex random matrix with empirical eigenvalue distribution converging almost surely to a nonrandom distribution whose Stieltjes transform is $\mathcal{S}_0$. Then, if $\mathbf{H}$, $\mathbf{T}$, and $\mathbf{W}_0$ are independent, the empirical eigenvalue distribution of

$$\mathbf{W} = \mathbf{W}_0 + \mathbf{H}\mathbf{T}\mathbf{H}^\dagger \tag{D.17}$$

converges almost surely, as $K, N \to \infty$ with $\frac{K}{N} \to \beta$, to a nonrandom distribution whose Stieltjes transform $\mathcal{S}(\cdot)$ satisfies

$$\mathcal{S}(z) = \mathcal{S}_0\left(z - \beta E\left\{\frac{\mathsf{T}}{1+\mathsf{T}\mathcal{S}(z)}\right\}\right). \tag{D.18}$$

Particularizing to the case in which $\mathbf{W}_0 = \mathbf{0}$, while relaxing the conditions to a Hermitian matrix $\mathbf{T}$, the following results for the η-transform and Shannon transform of the limiting empirical eigenvalue distribution can be derived based on [100].

Theorem D.3 [121, Theorem 2.39]. Let $\mathbf{H}$ be an $N \times K$ matrix whose entries are i.i.d. complex random variables with variances $\frac{1}{N}$. Further, let $\mathbf{T}$ be a $K \times K$ Hermitian nonnegative random matrix,

independent of $\mathbf{H}$, whose empirical eigenvalue distribution converges almost surely to a nonrandom limit. Then, the empirical eigenvalue distribution of $\mathbf{HTH}^\dagger$ converges almost surely, as $K, N \to \infty$ with $\frac{K}{N} \to \beta$, to a distribution whose η-transform $\eta_{\mathbf{HTH}^\dagger}(\cdot)$ satisfies

$$1 = \eta + \beta(1 - \eta_{\mathbf{T}}(\gamma\eta)), \tag{D.19}$$

while its Shannon transform satisfies

$$\mathcal{V}_{\mathbf{HTH}^\dagger}(\gamma) = \beta\mathcal{V}_{\mathbf{T}}(\eta\gamma) + \log\frac{1}{\eta} + (\eta - 1)\log\mathrm{e}, \tag{D.20}$$

where η stands for the corresponding value of $\eta_{\mathbf{HTH}^\dagger}(\cdot)$.

The condition of i.i.d. entries can be relaxed to independent entries with common means and variances $\frac{1}{N}$ satisfying the Lindeberg-type condition (Equation (D.15)).

Returning to Theorem D.2, while focusing on the particular case of $\mathbf{T} = \mathbf{I}$, then $\eta_{\mathbf{T}}(\gamma) = \frac{1}{1+\gamma}$ and Equation (D.19) becomes

$$\eta = 1 - \beta + \frac{\beta}{1 + \gamma\eta}, \tag{D.21}$$

yielding an explicit solution, which is the η-transform of the Marčenko–Pastur distribution $\tilde{\mathsf{f}}_\beta$ (Equation (D.16)):

$$\eta(\gamma) = 1 - \frac{\mathcal{F}(\gamma, \beta)}{4\gamma}. \tag{D.22}$$

Here $\mathcal{F}(x, z)$ is the function introduced in [126]:

$$\mathcal{F}(x, z) = \left(\sqrt{x(1 + \sqrt{z})^2 + 1} - \sqrt{x(1 - \sqrt{z})^2 + 1}\right)^2. \tag{D.23}$$

The corresponding Shannon transform in this case is explicitly given by

$$\begin{aligned}\mathcal{V}_{\mathbf{HH}^\dagger} \;=\; & \beta\log\left(1 + \gamma - \frac{1}{4}\mathcal{F}(\gamma, \beta)\right) \\ & + \log\left(1 + \gamma\beta - \frac{1}{4}\mathcal{F}(\gamma\beta)\right) - \frac{\mathcal{F}(\gamma, \beta)}{4\gamma}\log\mathrm{e}.\end{aligned} \tag{D.24}$$

D.3.2 The Girko Law

In some wireless communications problems of interest, the entries of the underlying channel transfer matrix are no longer i.i.d. random variables (in particular, they do not have identical means and variances), as required for the basic Marčenko–Pastur law to hold. For such settings a useful result is the following theorem by Girko [29], which also forms the basis for several wireless communications related random matrix theory results of recent years. The following formulation is due to [121], and the reader is referred to [45] for an alternative representation of the original result of [29].

Consider an $N \times K$ random matrix $\mathbf{H}$ whose entries have variances

$$\mathrm{Var}[\mathbf{H}_{i,j}] = \frac{[\mathbf{P}]_{i,j}}{N}, \tag{D.25}$$

with $\mathbf{P}$ an $N \times K$ deterministic matrix having uniformly bounded entries. For each N let

$$v^N : [0,1) \times [0,1) \to \mathbb{R}$$

be the *variance profile* function given by

$$v^N(x,y) = [\mathbf{P}]_{i,j}, \quad \frac{i-1}{N} \le x < \frac{i}{N}, \quad \frac{j-1}{K} \le y < \frac{j}{K}. \tag{D.26}$$

Whenever $v^N(x,y)$ converges uniformly to a limiting bounded measurable function $v(x,y)$, we define this limit as the *asymptotic variance profile* of $\mathbf{H}$.

Theorem D.4 [121, Theorem 2.50]. Let $\mathbf{H}$ be an $N \times K$ random matrix whose entries are independent zero-mean complex random variables (arbitrarily distributed) satisfying the Lindeberg condition (D.15), and with variances

$$E\left\{ |[\mathbf{H}]_{i,j}|^2 \right\} = \frac{[\mathbf{P}]_{i,j}}{N}, \tag{D.27}$$

where $\mathbf{P}$ is an $N \times K$ deterministic matrix with uniformly bounded entries, from which the asymptotic variance profile of $\mathbf{H}$, denoted $v(x,y)$, can be obtained as defined above. As $K, N \to \infty$ with $\frac{K}{N} \to \beta$,

the empirical eigenvalue distribution of $\mathbf{HH}^\dagger$ converges almost surely to a limiting distribution whose η-transform is

$$\eta_{\mathbf{HH}^\dagger}(\gamma) = E\left\{\Gamma_{\mathbf{HH}^\dagger}(\mathsf{X},\gamma)\right\}, \tag{D.28}$$

with $\Gamma_{\mathbf{HH}^\dagger}(x,\gamma)$ satisfying the equations

$$\Gamma_{\mathbf{HH}^\dagger}(x,\gamma) = \frac{1}{1+\beta\gamma E\left\{v(x,\mathsf{Y})\Upsilon_{\mathbf{HH}^\dagger}(\mathsf{Y},\gamma)\right\}} \quad \text{and} \tag{D.29}$$

$$\Upsilon_{\mathbf{HH}^\dagger}(y,\gamma) = \frac{1}{1+\gamma E\left\{v(\mathsf{X},y)\Gamma_{\mathbf{HH}^\dagger}(\mathsf{X},\gamma)\right\}}, \tag{D.30}$$

where X and Y are independent random variables uniform on $[0,1]$.

The Shannon transform of the asymptotic spectrum of $\mathbf{HH}^\dagger$ is given by [121]

$$\begin{aligned}\mathcal{V}_{\mathbf{HH}^\dagger}(\gamma) = {} & \beta E\left\{\log(1+\gamma E\left\{v(\mathsf{X},\mathsf{Y})\Gamma_{\mathbf{HH}^\dagger}(\mathsf{X},\gamma)\,|\,\mathsf{Y}\right\})\right\} \\ & + E\left\{\log(1+\gamma\beta E\left\{v(\mathsf{X},\mathsf{Y})\Upsilon_{\mathbf{HH}^\dagger}(\mathsf{Y},\gamma)\,|\,\mathsf{X}\right\})\right\} \\ & - \gamma\beta E\left\{v(\mathsf{X},\mathsf{Y})\Gamma_{\mathbf{HH}^\dagger}(\mathsf{X},\gamma)\Upsilon_{\mathbf{HH}^\dagger}(\mathsf{Y},\gamma)\right\}\log\mathrm{e}\end{aligned} \tag{D.31}$$

with $\Gamma_{\mathbf{HH}^\dagger}(\cdot,\cdot)$ and $\Upsilon_{\mathbf{HH}^\dagger}(\cdot,\cdot)$ defined as in Equations (D.29) and (D.30), respectively.

D.3.3 Asymptotic Results for Band Matrices

Consider the case in which the matrix $\mathbf{H}$ is square $N \times N$ and can be expressed as

$$\mathbf{H} = \mathbf{G} \circ \tilde{\mathbf{H}}, \tag{D.32}$$

where $\mathbf{A} \circ \mathbf{B}$ denotes the Hadamard product of the matrices $\mathbf{A}$ and $\mathbf{B}$, $\tilde{\mathbf{H}}$ consists of i.i.d. standard complex Gaussian random variables, and $[\mathbf{G}]_{i,j} = \sigma_{i,j}$, where $\sigma_{i,j}^2$ is the variance of the (i,j)-th entry in $\mathbf{H}$. The matrix $\mathbf{G}$ can be referred to, in the context of wireless MIMO channels, as the *channel pattern mask* [66].

The particular case in which $\mathbf{G}$ is a *circulant matrix* given by

$$[\mathbf{G}]_{i,j} = \begin{cases} 1 & \text{if } j = (i + l \mod N), \\ & \text{where } l = \begin{cases} -\dfrac{D-1}{2}, \ldots, \dfrac{D-1}{2} & \text{if } D \text{ is odd}, \\ -\dfrac{D}{2} + 1, \ldots, \dfrac{D}{2} & \text{if } D \text{ is even}, \end{cases} \\ 0 & \text{otherwise}, \end{cases} \tag{D.33}$$

is considered in [66] and used to represent various practical physical channel scattering scenarios (the channel model is referred to therein as a *D-connected channel*). Other cases of interest are the ones in which $\mathbf{G}$ is a Toeplitz matrix. In particular, the cases in which the nonzero row elements of $\mathbf{G}$ are $\{\alpha, 1, \alpha\}$, for some constant $\alpha \in (0, 1]$, corresponding to the Wyner model, or $\{1, 1\}$, have been the basis for several analyses of multicell systems, some of which are reviewed in the main body of this monograph (see, e.g., [78, 111, 132]).

For the case in which the band size (D in Equation (D.33)) grows with the matrix dimensions at an appropriate rate, an explicit result on the limiting spectral distribution is given in [66, Theorem 4]. In contrast, the derivation of explicit results on the limiting spectral distributions of band matrices, when the band size does not grow with the matrix dimensions, poses considerable difficulties (see, e.g., the discussion in [116]).

One of the rare exceptions in which explicit limiting results can be obtained for finite-band matrices is reported in [78]. The result therein focuses on the case in which the channel transfer matrix $\mathbf{H}$ is equal to

$$\mathbf{H} = \begin{pmatrix} \alpha_{2,1} & \alpha_{1,1} & 0 & \cdots & \\ 0 & \alpha_{2,2} & \alpha_{2,1} & 0 & \cdots \\ & & \ddots & & \\ & & & \alpha_{2,N} & \alpha_{1,N} \end{pmatrix}, \tag{D.34}$$

where the entries $\{\alpha_{i,j}\}$ are i.i.d. zero-mean circularly symmetric complex Gaussian random variables with unit variances (this model corresponds to case in which the power mask matrix $\mathbf{G}$ of

Equation (D.33) is Toeplitz, with basic nonzero row elements $\{1, 1\}$). Let $\mathbf{A}$ be defined as the covariance matrix

$$\mathbf{A} = \mathbf{I} + \gamma \mathbf{H}\mathbf{H}^{\dagger}, \tag{D.35}$$

where γ is a real nonnegative constant. Applying the Cholesky decomposition, the matrix $\mathbf{A}$ is decomposed as $\mathbf{A} = \mathbf{LDU}$, where $\mathbf{L}$ and $\mathbf{U}$ are lower and upper triangular matrices, respectively, with unit diagonal entries, and $\mathbf{D}$ is a diagonal matrix. From this decomposition, the k-th diagonal element of $\mathbf{D}$ is given by

$$\begin{aligned} d_k &= 1 + \gamma(|\alpha_{1,k}|^2 + |\alpha_{2,k}|^2) - \frac{\gamma^2 |\alpha_{1,k-1}|^2 |\alpha_{2,k}|^2}{d_{k-1}} \\ &= 1 + \gamma |\alpha_{1,k}|^2 + \gamma |\alpha_{2,k}|^2 \left(1 - \gamma \frac{|\alpha_{1,k-1}|^2}{d_{k-1}}\right). \end{aligned} \tag{D.36}$$

The main result in [78] (in the context of random matrix theory) can now be summarized as follows.

Theorem D.5. As $N \to \infty$, the diagonal elements $\{d_k\}$ of the Cholesky decomposition matrix $\mathbf{D}$ may be viewed as a first order discrete-time continuous space Markov chain, with an ergodic stationary distribution given by

$$f_d(x) = \frac{\log(x)\, \mathrm{e}^{-\frac{x}{\gamma}}}{\mathrm{Ei}\left(\frac{1}{\gamma}\right)\gamma}, \qquad x \geq 1 \ , \tag{D.37}$$

where $\mathrm{Ei}(x) = \int_1^{\infty} \frac{\exp(-xt)}{t}\, dt$ is the exponential integral function.

For the case in which the power mask matrix $\mathbf{G}$ is Toeplitz, with nonzero row elements $\{\alpha, 1, \alpha\}$, corresponding to the Wyner model, the *power moments* of the limiting eigenvalue distribution of $\mathbf{H}\mathbf{H}^{\dagger}$ can be calculated for any finite order, as shown in [111]. When the entries of $\mathbf{H}$ are zero-mean and Gaussian, the empirical eigenvalue distribution of $\mathbf{H}\mathbf{H}^{\dagger}$ was shown in [111] to converge weakly to a unique limiting distribution. Some generalizations and additional characterization of the Shannon-transform of the limiting distribution can be found in [62]. However, explicitly determining the limiting distribution itself, as well as any of its useful transforms, is still an open problem [116].

Abbreviations and Acronyms

BC: Broadcast Channel
BS: Base Station
CP: Central Processor
CSI: Channel State Information
DPC: Dirty Paper Coding
GBC: Gaussian Broadcast Channel
MAC: Multiple Access Channel
MCP: Multi-Cell Processing
MIMO: Multiple-Input Multiple-Output
MMSE: Minimum Mean Square Error
MS: Mobile Station
RS: Relay Station
SCP: Single-Cell Processing
SINR: Signal-to-Interference-plus-Noise Ratio
SNR: Signal-to-Noise Ratio
TDMA: Time Division Multiple Access
WB: Wide-Band

Acknowledgments

The work of O. Simeone has been partly supported by the U.S. National Science Foundation under Grant No. 0914899. The work of H.V. Poor has been partly supported by the U.S. National Science Foundation under Grant No. 0905398. The work of S. Shamai has been supported by the Israel Science Foundation.

References

[1] 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*, p. 484, Chicago, IL, USA, June–July 2004.

[2] G. W. Anderson and O. Zeitouni, "A CLT for a band matrix model," *Probability Theory and Related Fields*, vol. 134, pp. 283–338, 2005.

[3] J. G. Andrews, A. Ghosh, J. Zhang, and R. Muhamed, *Fundamentals of LTE*. Boston, MA, USA: Prentice-Hall, 2010.

[4] Z. D. Bai, "Methodologies in spectral analysis of large dimensional random matrices, a review," *Statistica Sinica*, vol. 9, no. 3, pp. 611–677, 1999.

[5] I. Bergel, D. Yellin, and S. Shamai, "Linear precoding bounds for the Wyner cellular channel model with limited cooperation," in *Proceedings of the IEEE International Workshop on Signal Processing Advances in Wireless Communications*, pp. 206–210, San Francisco, CA, USA, June 2011.

[6] T. Berger, Z. Zhang, and H. Viswanathan, "The CEO problem," *IEEE Transactions on Information Theory*, vol. 42, no. 3, pp. 887–902, May 1996.

[7] R. Bhagavatula and R. W. Heath, "Adaptive limited feedback for sum-rate maximizing beamforming in cooperative multicell systems," *IEEE Transactions on Signal Processing*, vol. 59, no. 2, pp. 800–811, February 2011.

[8] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: Information-theoretic and communications aspects," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2619–2692, October 1998.

[9] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge University Press, 2004.

[10] L. Brandenburg and A. Wyner, "Capacity of the Gaussian channel with memory: The multivariate case," *Bell System Technical Journal*, vol. 53, pp. 745–779, May 1974.

[11] G. Caire, N. Jindal, and S. Shamai, "On the required accuracy of transmitter channel state information in multiple antenna broadcast channels," in *Proceedings of the 41st Asilomar Conference on Signals, Systems and Computers*, pp. 287–291, Pacific Grove, CA, USA, November 2007.

[12] G. Caire and S. Shamai (Shitz), "On the achievable throughput of multiantenna Gaussian broadcast channels," *IEEE Transactions on Information Theory*, vol. 49, no. 7, pp. 1691–1706, July 2003.

[13] R. S. Cheng and S. Verdú, "Gaussian multiaccess channels with ISI: Capacity region and multiuser water-filling," *IEEE Transactions on Information Theory*, vol. 39, no. 3, pp. 773–785, May 1993.

[14] A. Cohen and A. Lapidoth, "The Gaussian watermarking game," *IEEE Transactions on Information Theory*, vol. 48, no. 6, pp. 1639–1667, June 2002.

[15] M. H. M. Costa, "Writing on dirty paper," *IEEE Transactions on Information Theory*, vol. IT-29, no. 3, pp. 439–441, May 1983.

[16] R. Couillet and M. Debbah, *Random Matrix Methods for Wireless Communications*. Cambridge, UK: Cambridge University Press, 2011.

[17] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Hoboken, NJ, USA: John Wiley and Sons, Inc., 1991.

[18] R. Dabora and S. Servetto, "On the role of estimate-and-forward with time sharing in cooperative communication," *IEEE Transactions on Information Theory*, vol. 54, no. 10, pp. 4409–4431, October 2008.

[19] R. A. DiFazio and P. J. Pietraski, "The bandwidth crunch: Can wireless technology meet the skyrocketing demand for mobile data?," in *Proceedings of the IEEE Systems, Applications and Technology Conference*, pp. 1–6, New York City, NY, USA, May 2011.

[20] S. C. Draper, B. J. Frey, and F. R. Kschischang, "Iterative decoding of a broadcast message," in *Proceedings of the 41st Annual Allerton Conference on Communications, Control and Computing*, Monticello, IL, USA, 2003.

[21] M. Duarte, C. Dick, and A. Sabharwal, "Experiment-driven characterization of full-duplex wireless systems," *submitted [arXiv:1107.1276v1]*, 2011.

[22] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge, UK: Cambridge University Press, 2012.

[23] U. Erez, S. Shamai (Shitz), and R. Zamir, "Capacity and lattice strategies for canceling known interference," *IEEE Transactions on Information Theory*, vol. 51, no. 11, pp. 3820–3833, November 2005.

[24] U. Erez and S. ten Brink, "A close-to-capacity dirty paper coding scheme," *IEEE Transactions on Information Theory*, vol. 51, no. 10, pp. 3417–3432, October 2005.

[25] U. Erez and R. Zamir, "Achieving $\frac{1}{2}\log(1 + \text{SNR})$ on the AWGN channel with lattice encoding and decoding," *IEEE Transactions on Information Theory*, vol. 50, no. 10, pp. 2293–2314, October 2004.

[26] G. D. Forney Jr., "On the role of MMSE estimation in approaching the information-theoretic limits of linear Gaussian channels: Shannon meets Wiener," in *Proceedings of the 41st Annual Allerton Conference on*

Communication, Control and Computing, pp. 430–439, Monticello, IL, USA, October 2003.

[27] S. Gelfand and M. Pinsker, "Coding for channels with random parameters," *Problems of Control and Information Theory*, vol. 9, no. 1, pp. 19–31, 1980.

[28] D. Gesbert, S. Hanly, H. Huang, S. Shamai Shitz, O. Simeone, and W. Yu, "Multi-cell MIMO cooperative networks: A new look at interference," *IEEE Journal on Selected Areas in Communications*, vol. 28, no. 9, pp. 1380–1408, December 2010.

[29] V. L. Girko, *Theory of Random Determinants*. Boston, MA, USA: Kluwer, 1990.

[30] R. M. Gray, "Toeplitz and circulant matrices: A review," *Trends in Communications and Information Theory*, vol. 2, no. 3, Hanover, MA, USA: Now Publishers Inc., December 2006.

[31] D. Guo, S. Shamai (Shitz), and S. Verdú, "Mutual information and minimum mean-square error in Gaussian channels," *IEEE Transactions on Information Theory*, vol. 51, no. 4, pp. 1261–1282, April 2005.

[32] T. S. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Transactions on Information Theory*, vol. 27, no. 1, pp. 49–60, January 1981.

[33] S. V. Hanly and P. A. Whiting, "Information-theoretic capacity of multi-receiver networks," *Telecommunications Systems*, vol. 1, pp. 1–42, 1993.

[34] Z. Hasan, H. Boostanimehr, and V. K. Bhargava, "Green cellular networks: A survey, some research issues and challenges," *IEEE Communications Surveys and Tutorials*, vol. 13, no. 4, pp. 524–540, Fourth Quarter 2011.

[35] 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, April 2003.

[36] S.-N. Hong and G. Caire, "Quantized compute and forward: A low-complexity architecture for distributed antenna systems," in *Proceedings of the IEEE Information Theory Workshop*, pp. 420–424, Paraty, Brazil, October 2011.

[37] S.-N. Hong and G. Caire, "Reverse compute and forward: A low-complexity architecture for downlink distributed antenna systems," *ArXiv e-prints*, `http://adsabs.harvard.edu/abs/2012arXiv1202.0854H`, February 2012.

[38] J. Hoydis, M. Kobayashi, and M. Debbah, "Optimal channel training in uplink network mimo systems," *IEEE Transactions on Signal Processing*, vol. 59, no. 6, pp. 2824–2833, June 2011.

[39] H. Huang, C. B. Papadias, and S. Venkatesan, *MIMO Communication for Cellular Networks*. New York, USA: Springer, 2012.

[40] S. A. Jafar, "Interference alignment: A new look at signal dimensions in a communication network," *Foundations and Trends in Communications and Information Theory*, vol. 7, no. 1, Hanover MA, USA: Now Publishers Inc., 2011.

[41] N. Jindal, "Multi-user communication systems: Capacity, duality, and cooperation," PhD Thesis, Stanford University, Stanford, CA, July 2004.

[42] N. Jindal, W. Rhee, S. Vishwanath, S. A. Jafar, and A. Goldsmith, "Sum power iterative water-filling for multi-antenna Gaussian broadcast channels,"

IEEE Transactions on Information Theory, vol. 51, no. 4, pp. 1570–1580, April 2005.

[43] S. Jing, D. N. C. Tse, J. Hou, J. Soriaga, J. E. Smee, and R. Padovani, "Downlink macro-diversity in cellular networks," in *Proceedings of the IEEE International Symposium on Information Theory*, pp. 1–5, Nice, France, June 2007.

[44] G. Keshet, Y. Steinberg, and N. Merhav, "Channel coding in the presence of side information," *Foundations and Trends in Communications and Information Theory*, vol. 4, no. 6, pp. 445–586, Hanover MA, USA: Now Publishers Inc., 2008.

[45] Kiran and D. N. C. Tse, "Effective interference and effective bandwidth of linear multiuser receivers in asynchronous CDMA systems," *IEEE Transactions on Information Theory*, vol. 46, no. 4, pp. 1426–1447, July 2000.

[46] M. Kobayashi and G. Caire, "An iterative water-filling algorithm for maximum weighted sum-rate of Gaussian MIMO-BC," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 8, pp. 1640–1646, August 2006.

[47] G. Kramer, "Outer bounds on the capacity of Gaussian interference channels," *IEEE Transactions on Information Theory*, vol. 50, no. 3, pp. 581–586, March 2004.

[48] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Transactions on Information Theory*, vol. 51, no. 9, pp. 3037–3063, September 2005.

[49] M. G. Krein and A. A. Nudelman, *The Markov Moment Problem and External Problems*, vol. 50. Providence, RI, USA: American Mathematical Society, 1977.

[50] L. Lai and H. E. Gamal, "On cooperation in energy efficient wireless networks: The role of altruistic nodes," *IEEE Transactions on Wireless Communications*, vol. 7, no. 5, pp. 1868–1878, May 2008.

[51] N. Laneman, D. Tse, and G. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Transactions on Information Theory*, vol. 50, no. 12, pp. 3062–3080, December 2004.

[52] A. Lapidoth, N. Levy, S. Shamai (Shitz), and M. Wigger, "Cognitive Wyner networks with clustering decoding," submitted to *IEEE Transactions on Information Theory*, arXiv:1203.3659.

[53] A. Lapidoth, S. Shamai (Shitz), and M. Wigger, "A linear interference network with local side-information," in *Proceedings of the 2007 IEEE International Symposium on Information Theory*, Nice, France, June 2007.

[54] A. Lapidoth, S. Shamai (Shitz), and M. A. Wigger, "On the capacity of fading MIMO broadcast channels with imperfect transmitter side-information," in *Proceedings of the 43rd Annual Allerton Conference on Communication, Control, and Computing*, Monticello, IL, USA, September 2005. (Invited paper).

[55] L. Le and E. Hossain, "Multihop cellular networks: Potential gains, research challenges, and a resource allocation framework," *IEEE Communications Magazine*, vol. 45, no. 9, pp. 66–73, September 2007.

[56] N. Letzepis, "Gaussian cellular multiple access channels," PhD. Thesis, Institute for Telecommunications Research, University of South Australia, Mawson Lakes, Australia, December 2006.
[57] N. Levy, "Information theoretic aspects of cellular communication models with joint multiple-cell processing," Ph.D. Thesis, Technion, Haifa, Israel, July 2009.
[58] N. Levy, S. Shamai, M. Wigger, and A. Lapidoth, "A cognitive network with clustered decoding," in *Proceedings of the IEEE Symposium on Information Theory*, Seoul, South Korea, June–July 2009.
[59] N. Levy and S. Shamai (Shitz), "Clustered local decoding for Wyner-type cellular models," *IEEE Transactions on Information Theory*, vol. 55, no. 11, pp. 4967–4985, November 2009.
[60] N. Levy and S. Shamai (Shitz), "Information theoretic aspects of users' activity in a Wyner-like cellular model," *IEEE Transactions on Information Theory*, vol. 56, no. 5, pp. 2241–2248, May 2010.
[61] N. Levy, O. Zeitouni, and S. Shamai (Shitz), "Central limit theorem and large deviations of the fading Wyner cellular model via product of random matrices theory," *Problems of Information Transmission*, vol. 45, no. 1, pp. 5–22, 2009.
[62] N. Levy, O. Zeitouni, and S. Shamai (Shitz), "On information rates of the fading Wyner cellular model via the Thouless formula for the strip," *IEEE Transactions on Information Theory*, vol. 56, no. 11, pp. 5495–5514, November 2010.
[63] Y. Liang and A. Goldsmith, "Symmetric rate capacity of cellular systems with cooperative base stations," in *Proceedings of the IEEE Global Communications Conference*, San Francisco, CA, USA, November–December 2006.
[64] Lightradio: White paper 1 White paper, Alcatel-Lucent, 2011.
[65] S. H. Lim, Y.-H. Kim, A. El Gamal, and S.-Y. Chung, "Noisy network coding," *IEEE Transactions on Information Theory*, vol. 57, no. 5, pp. 3132–3152, May 2011.
[66] K. Liu, V. Raghavan, and A. M. Sayeed, "Capacity scaling and spectral efficiency in wide-band correlated MIMO channels," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2504–2526, October 2003.
[67] Y. Liu and E. Erkip, "On the sum capacity of K-user cascade Gaussian Z-interference channel," in *Proceedings of the IEEE International Symposium on Information Theory*, Saint Petersburg, Russia, July 2011.
[68] A. Lozano, R. W. Heath, Jr., and J. G. Andrews, "Fundamental limits of cooperation," *ArXiv e-prints*, March 2012.
[69] A. Lozano, A. M. Tulino, and S. Verdú, "High-SNR power offset in multi-antenna communication," in *Proceedings of the IEEE International Symposium on Information Theory*, Chicago, IL, USA, June–July 2004.
[70] M. Maddah-Ali and D. Tse, "Completely stale transmitter channel state information is still very useful," in *Proceedings of the 48th Annual Allerton Conference on Communication, Control, and Computing*, pp. 1188–1195, Monticello, IL, USA, September–October 2010.
[71] G. Maierbacher and J. Barros, "Low-complexity coding for the CEO problem with many encoders," in *Proceedings of the Symposium on Information Theory in the Benelux*, Brussels, Belgium, May 2005.

[72] H. Maleki, S. A. Jafar, and S. Shamai, "Retrospective interference alignment," *submitted*, abs/1009.3593, 2010.

[73] I. Marict', B. Bostjancic, and A. Goldsmith, "Resource allocation for constrained backhaul in picocell networks," in *Proceedings of the Information Theory and Applications Workshop*, pp. 1–6, La Jolla, CA, USA, February 2011.

[74] P. Marsch and G. Fettweis, "On uplink network MIMO under a constrained backhaul and imperfect channel knowledge," in *Proceedings of the IEEE International Conference on Communications*, pp. 1–6, Dresden, Germany, June 2009.

[75] V. A. Marčenko and L. A. Pastur, "Distributions of eigenvalues for some sets of random matrices," *Math USSR-Sbornik*, vol. 1, pp. 457–483, 1967.

[76] M. Mohseni, R. Zhang, and J. M. Cioffi, "Optimized transmission for fading multiple-access and broadcast channels with multiple antennas," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 8, pp. 1627–1639, August 2006.

[77] R. Narasimhan, "Individual outage rate regions for fading multiple access channels," in *Proceedings of the IEEE International Symposium on Information Theory*, pp. 1571–2660, Nice, France, June 2007.

[78] A. Narula, "Information theoretic analysis of multiple-antenna transmission diversity," PhD Thesis, Massachusetts Institute of Technology (MIT), Cambridge, MA, USA, June 1997.

[79] B. Nazer, A. Sanderovich, M. Gastpar, and S. Shamai, "Structured superposition for backhaul constrained cellular uplink," in *Proceedings of the IEEE Symposium on Information Theory*, pp. 1530–1534, Seoul, South Korea, June–July 2009.

[80] Y. Oohama, "The rate-distortion function for the quadratic Gaussian CEO problem," *IEEE Transactions on Information Theory*, vol. 44, no. 3, pp. 1057–1070, May 1998.

[81] Y. Oohama, "Rate-distortion theory for Gaussian multiterminal source coding systems with several side informations at the decoder," *IEEE Transactions on Information Theory*, vol. 51, no. 7, pp. 2577–2593, July 2005.

[82] L. H. Ozarow, S. Shamai (Shitz), and A. D. Wyner, "Information theoretic considerations for cellular mobile radio," *IEEE Transactions on Vehicular Technology*, vol. 43, no. 2, pp. 359–378, May 1994.

[83] V. Prabhakaran, D. Tse, and K. Ramachandran, "Rate region of the quadratic Gaussian CEO problem," in *Proceedings of the International Symposium on Information Theory*, p. 119, Chicago, IL, USA, June–July 2004.

[84] S. Ramprashad and G. Caire, "Cellular vs. network MIMO: A comparison including the channel state information overhead," in *Proceedings of the 20th IEEE International Symposium on Personal, Indoor and Mobile Radio*, pp. 878–884, Tokyo, Japan, September 2009.

[85] 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, July 2008.

[86] A. Sanderovich, O. Somekh, H. V. Poor, and S. Shamai (Shitz), "Uplink macro diversity of limited backhaul cellular network," *IEEE Transactions on Information Theory*, vol. 55, no. 8, pp. 3457–3478, August 2009.

[87] I. Sason, "On achievable rate regions for the Gaussian interference channel," *IEEE Transactions on Information Theory*, vol. 50, no. 6, pp. 1345–1356, June 2004.

[88] H. Sato, "An outer bound on the capacity region of braodcast channels," *IEEE Transactions on Information Theory*, vol. 24, no. 3, pp. 374–377, May 1978.

[89] M. Sauer, A. Kobyakov, and J. George, "Radio over fiber for picocellular network architectures," *Journal of Lightwave Technology*, vol. 25, no. 11, pp. 3301–3320, November 2007.

[90] S. Shamai and R. Laroia, "The intersymbol interference channel: Lower bounds on capacity and channel precoding loss," *IEEE Transactions on Information Theory*, vol. 42, no. 5, pp. 1388–1404, September 1996.

[91] S. Shamai and M. Wigger, "Rate-limited transmitter-cooperation in Wyner's asymmetric interference network," in *Proceedings of the IEEE International Symposium on Information Theory Proceedings*, pp. 425–429, Saint Petersburg, Russia, July–August 2011.

[92] S. Shamai (Shitz), O. Somekh, O. Simeone, A. Sanderovich, B. M. Zaidel, and H. V. Poor, "Cooperative multi-cell networks: Impact of limited-capacity backhaul and inter-users links," in *Proceedings of the Joint Workshop on Communications and Coding*, Dürnstein, Austria, October 2007.

[93] S. Shamai (Shitz), O. Somekh, and B. M. Zaidel, "Multi-cell communications: An information theoretic perspective," in *Proceedings of the Joint Workshop on Communications and Coding*, Donnini, Florence, Italy, October 2004.

[94] S. Shamai (Shitz) and S. Verdú, "The impact of frequency-flat fading on the spectral efficiency of CDMA," *IEEE Transactions on Information Theory*, vol. 47, no. 4, pp. 1302–1327, May 2001.

[95] S. Shamai (Shitz) and A. D. Wyner, "Information-theoretic considerations for symmetric, cellular, multiple-access fading channels — Parts I & II," *IEEE Transactions on Information Theory*, vol. 43, no. 6, pp. 1877–1911, November 1997.

[96] 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*, pp. 1745–1749, Rhodes, Greece, May 2001.

[97] X. Shang, G. Kramer, and B. Chen, "A new outer bound and the noisy-interference sum rate capacity for Gaussian interference channels," *IEEE Transactions on Information Theory*, vol. 55, no. 2, pp. 689–699, February 2009.

[98] M. Sharif and B. Hassibi, "On the capacity of MIMO broadcast channel with partial side information," *IEEE Transactions on Information Theory*, vol. 51, no. 2, pp. 506–522, February 2005.

[99] O. Shental, N. Shental, S. Shamai (Shitz), I. Kanter, A. J. Weiss, and Y. Weiss, "Discrete-input two-dimensional Gaussian channels with memory: Estimation and information rates via graphical models and statistical mechanics," *IEEE Transactions on Information Theory*, vol. 54, no. 4, pp. 1500–1513, April 2008.

[100] J. W. Silverstein, "Strong convergence of the empirical distribution of eigenvalues of large dimensional random matrices," *Journal of Multivariate Analysis*, vol. 55, pp. 331–339, 1995.

[101] J. W. Silverstein and Z. D. Bai, "On the empirical distribution of eigenvalues of a class of large dimensional random matrices," *Journal of Multivariate Analysis*, vol. 54, pp. 175–192, 1995.

[102] O. Simeone, O. Somekh, Y. Bar-Ness, H. V. Poor, and S. Shamai (Shitz), "Capacity of linear two-hop mesh networks with rate splitting, decode-and-forward relaying and cooperation," in *Proceedings of the 45th Annual Allerton Conference on Communication, Control and Computing*, Monticello, IL, USA, September 2007.

[103] O. Simeone, O. Somekh, Y. Bar-Ness, and U. Spagnolini, "Uplink throughput of TDMA cellular systems with multicell processing and amplify-and-forward cooperation between mobiles," *IEEE Transactions on Wireless Communications*, vol. 6, no. 8, pp. 2942–2951, August 2007.

[104] O. Simeone, O. Somekh, Y. Bar-Ness, and U. Spagnolini, "Throughput of low-power cellular systems with collaborative base stations and relaying," *IEEE Transactions on Information Theory*, vol. 54, no. 1, pp. 459–467, January 2008.

[105] O. Simeone, O. Somekh, G. Kramer, H. V. Poor, and S. Shamai, "Throughput of cellular systems with conferencing mobiles and cooperative base stations," *EURASIP Journal on Wireless Communications and Networking*, vol. 2008, pp. 27:1–27:14, January 2008.

[106] O. Simeone, O. Somekh, G. Kramer, H. V. Poor, and S. Shamai, "Uplink sum-rate analysis of a multicell system with feedback," in *Proceedings of the 46th Annual Allerton Conference on Communication, Control, and Computing*, pp. 1030–1036, Monticello, IL, USA, September 2008.

[107] O. Simeone, O. Somekh, H. V. Poor, and S. Shamai, "Downlink multicell processing with limited-backhaul capacity," *EURASIP Journal on Advances in Signal Processing*, vol. 2009, pp. 3:1–3:10, February 2009.

[108] O. Simeone, O. Somekh, H. V. Poor, and S. Shamai, "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, January 2009.

[109] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," vol. IT-19, pp. 471–480, July 1973.

[110] D. Soldani and S. Dixit, "Wireless relays for broadband access [radio communications series]," *IEEE Communications Magazine*, vol. 46, no. 3, pp. 58–66, March 2008.

[111] O. Somekh and S. Shamai (Shitz), "Shannon-theoretic approach to a Gaussian cellular multi-access channel with fading," *IEEE Transactions on Information Theory*, vol. 46, no. 4, pp. 1401–1425, July 2000.

[112] O. Somekh, O. Simeone, Y. Bar-Ness, A. Haimovich, and S. Shamai, "Cooperative multicell zero-forcing beamforming in cellular downlink channels," *IEEE Transactions on Information Theory*, vol. 55, no. 7, pp. 3206–3219, July 2009.

[113] O. Somekh, O. Simeone, Y. Bar-Ness, A. M. Haimovich, U. Spagnolini, and S. Shamai (Shitz), *Distributed Antenna Systems: Open Architecture for Future*

Wireless Communications, chapter: An information theoretic view of distributed antenna processing in cellular systems. Boca Raton, FL, USA: Auerbach Publications, CRC Press, May 2007.

[114] O. Somekh, O. Simeone, H. V. Poor, and S. Shamai, "Cellular systems with non-regenerative relaying and cooperative base stations," *IEEE Transactions on Wireless Communications*, vol. 9, no. 8, pp. 2654–2663, August 2010.

[115] O. Somekh, O. Simeone, H. V. Poor, and S. Shamai (Shitz), "Throughput of dynamic cellular uplink with dynamic user activity and cooperative base-stations," in *Proceedings of the IEEE Information Theory Workshop*, Taormina, Sicily, October 2009.

[116] O. Somekh, O. Simeone, B. Zaidel, H. V. Poor, and S. Shamai, "On the spectrum of large random hermitian finite-band matrices," in *Proceedings of the Information Theory and Applications Workshop*, pp. 476–479, La Jolla, CA, USA, January–February 2008.

[117] O. Somekh, B. Zaidel, and S. Shamai, "Spectral efficiency of joint multiple cell-site processors for randomly spread DS-CDMA systems," *IEEE Transactions on Information Theory*, vol. 53, no. 7, pp. 2625–2637, July 2007.

[118] 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, December 2007.

[119] D. N. C. Tse, P. Viswanath, and L. Zheng, "Diversity-multiplexing tradeoff in multiple-access channels," *IEEE Transactions on Information Theory*, vol. 50, pp. 1859–1874, 2004.

[120] A. M. Tulino, H. Huh, and G. Caire, "Network MIMO with linear zero-forcing beamforming: Large system analysis, impact of channel estimation and reduced-complexity scheduling," submitted to *IEEE Transactions on Information Theory*, [Online]. Available: http://arxiv.org/abs/ 1012.3198, December 2010.

[121] A. M. Tulino and S. Verdú, "Random matrix theory and wireless communications," *Trends in Communications and Information Theory*, vol. 1, no. 1, Hanover, MA, USA: Now Publishers Inc., 2004.

[122] S. Venkatesan, A. Lozano, and R. Valenzuela, "Network MIMO: Overcoming intercell interference in indoor wireless systems," in *Proceedings of the 41st Asilomar Conference on Signals, Systems and Computers*, pp. 83–87, Pacific Grove, CA, USA, November 2007.

[123] S. Verdu, "The capacity region of the symbol-asynchronous Gaussian multiple-access channel," *IEEE Transactions on Information Theory*, vol. 35, no. 4, pp. 733–751, July 1989.

[124] S. Verdú, *Multiuser Detection.* Cambridge, UK: Cambridge University Press, 1998.

[125] S. Verdú, "Spectral efficiency in the wideband regime," *IEEE Transactions on Information Theory*, vol. 48, no. 6, pp. 1329–1343, June 2002.

[126] S. Verdú and S. Shamai (Shitz), "Spectral efficiency of CDMA with random spreading," *IEEE Transactions on Information Theory*, vol. 45, no. 2, pp. 622–640, March 1999.

[127] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2658–2669, October 2003.

[128] P. Viswanath and D. N. C. Tse, "Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality," *IEEE Transactions on Information Theory*, vol. 49, no. 8, pp. 1912–1921, August 2003.

[129] H. Viswanathan and T. Berger, "The quadratic Gaussian CEO problem," *IEEE Transactions on Information Theory*, vol. 43, no. 5, pp. 1549–1559, September 1997.

[130] H. Weingarten, Y. Steinberg, and S. Shamai (Shitz), "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Transactions on Information Theory*, vol. 52, no. 9, pp. 3936–3964, September 2006.

[131] F. M. J. Willems, "Information theoretical results for the discrete memoryless multiple access channel," PhD Thesis, Katholieke Universiteit, Leuven, Belgium, December 1982.

[132] 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, November 1994.

[133] A. D. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Transactions on Information Theory*, vol. 22, no. 1, pp. 1–10, January 1976.

[134] J. Xu, J. G. Andrews, and S. A. Jafar, "Broadcast channels with delayed finite-rate feedback: Predict or observe?," *CoRR*, abs/1105.3686, 2011.

[135] J. Xu, J. Zhang, and J. G. Andrews, "On the accuracy of the Wyner model in cellular networks," *IEEE Transactions on Wireless Communications*, vol. 10, no. 9, pp. 3098–3109, September 2011.

[136] T. Yoo and A. Goldsmith, "Optimality of zero-forcing beam forming with multiuser diversity," in *Proceedings of the IEEE International Conference on Communications*, Seoul, South Korea, May 2005.

[137] W. Yu, "Uplink–downlink duality via minimax duality," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 361–374, February 2006.

[138] W. Yu, "Competition and cooperation in multiuser communication environments," PhD Thesis, Stanford University, Stanford, CA, USA, June 2002.

[139] W. Yu and J. M. Cioffi, "Trellis precoding for the broadcast channel," in *Proceedings of IEEE Global Telecommunications Conference*, San Antonio, TX, USA, November 2001.

[140] W. Yu and J. M. Cioffi, "Sum capacity of Gaussian vector broadcast channels," *IEEE Transactions on Information Theory*, vol. 50, no. 9, pp. 1875–1892, September 2004.

[141] W. Yu and T. Lan, "Minimax duality of Gaussian vector broadcast channels," in *Proceedings of the 2004 IEEE International Symposium on Information Theory*, p. 177, Chicago, IL, USA, June–July 2004.

[142] W. Yu and T. Lan, "Transmitter optimization for the multi-antenna downlink with per-antenna power constraints," *IEEE Transactions on Signal Processing*, vol. 55, no. 6, pp. 2646–2660, June 2007.

[143] B. M. Zaidel, S. Shamai (Shitz), and S. Verdú, "Multi-cell uplink spectral efficiency of randomly spread DS-CDMA in Rayleigh fading channels," in *Proceedings of the International Symposium on Communication Techniques and Applications*, pp. 499–504, Ambleside, UK, July 2001.

[144] R. Zakhour and S. V. Hanly, "Base station cooperation on the downlink: Large system analysis," *IEEE Transactions on Information Theory*, vol. 58, no. 4, pp. 2079–2106, April 2012.

[145] R. Zamir, "Lattices are everywhere," in *Proceedings of the Information Theory and Applications Workshop*, pp. 392–421, La Jolla, CA, USA, February 2009.

[146] R. Zamir and T. Berger, "Multiterminal source coding with high resolution," *IEEE Transactions on Information Theory*, vol. 45, no. 1, pp. 106–117, January 1999.

[147] R. Zamir, S. Shamai (Shitz), and U. Erez, "Nested linear/lattice codes for structured multiterminal binning," *IEEE Transactions on Information Theory*, vol. 48, no. 6, pp. 1250–1276, June 2002.

[148] L. Zhang, R. Zhang, Y.-C. Liang, Y. Xin, and H. V. Poor, "On Gaussian MIMO BC-MAC duality with multiple transmit covariance constraints," *IEEE Transactions on Information Theory*, vol. 58, no. 4, pp. 2064–2078, April 2012.

[149] L. Zhou and W. Yu, "Approximate capacity region of the k-user cyclic Gaussian interference channel," in *Proceedings of the IEEE International Symposium on Information Theory*, Saint Petersburg, Russia, July 2011.