Advances in Base-Station and Mobile-Station Aided Cooperative Wireless Communications


Abstract—Albeit the conception of relaying dates back to the 1970s, in recent years there has been an upsurge of search interest in cooperative wireless communications in both academia and industry. This article presents an easy-reading overview of the pivotal topics in both mobile station (MS) and base station (BS) assisted cooperation in the context of cellular radio systems. Owing to the ever-increasing amount of literature in this particular field, this paper is by no means exhaustive, but intends to serve as a roadmap for this area by assembling a representative sample of recent results and to stimulate further research. The emphasis is initially on relay-base cooperation, relying on network coding, followed by the design of cross-layer cooperative protocols conceived for MS cooperation, as well as on the concept of coalition network element assisted BS cooperation. Then, a range of complexity and backhaul traffic reduction techniques that have been proposed for BS cooperation are reviewed. A more detailed discussion is provided in the context of MS cooperation concerning the pros and cons of dispensing with high-complexity, power-hungry channel estimation. Finally, generalized design guidelines, conceived for cooperative wireless communications, are presented.

I. INTRODUCTION TO COOPERATIVE WIRELESS COMMUNICATIONS

A. Mobile Station Cooperation

Single-Input Single-Output (SISO) communication systems obey the logarithmic Shannon capacity law, whilst Multiple-Input Multiple-Output (MIMO) systems are capable of increasing the achievable throughput linearly, provided that the Input Multiple-Output (MIMO) systems are capable of obeying the logarithmic Shannon capacity law, whilst Multiple-A. Mobile Station Cooperation as well as the associated hardware limitations. Furthermore, owing to the limited separation of the antenna elements, the transmitted signal rarely experiences independent fading; in other words, the corresponding signal replicas collected at the receiver are more likely to be in a deep fade simultaneously, which in turn erodes the achievable diversity gain. The diversity gain may be further compromised by the adverse effects of the large-scale shadow fading [2] at high operating frequencies, where all the MIMO channels tend to fade together rather than independently, imposing further signal correlation amongst the antennas in each other’s vicinity [3]. Apart from the above obstacles in the way of achieving multiple-antenna-aided diversity gains, wireless cellular networks aim to improve the coverage, capacity or the quality of end-user experience (QoE) in inadequately covered areas, such as for example indoor environments and rural areas. The dense deployment of fully-fledged base stations (BSs) constitutes a high-quality solution, albeit this may impose a high infrastructure cost and thus may become economically inviable, especially in low-traffic-density sparsely populated rural areas. In addition to the propagation-loss-induced low-power reception, the mobile stations (MSs) roaming in the cell edge region may also suffer from severe intercell interference.

Hence, to meet the challenging requirements of next-generation wireless networks in terms of coverage, capacity as well as deployment cost, the ingenious relay-aided cooperative transmission technique [4–7] appears to be one of the most promising solutions. The idea of user-cooperation-aided transmissions was originally conceived by simply relying on the fundamental broadcast nature of the wireless medium, which is frequently regarded as a drawback. In a nutshell, in multi-user wireless systems, single-antenna-assisted MSs may cooperatively share their antennas in order to achieve the so-called cooperative diversity as well as a path-loss-reduction based power gain by forming a virtual antenna array (VAA) [8,9] in both uplink (UL) and downlink (DL) transmissions. The concept of user cooperation has been first proposed in [7] for a two-user cooperative CDMA system, where orthogonal codes are employed by the active users in order to avoid multiple access interference. A user who directly sends his/her own information to the destination is regarded as a source node, while the other users who assist in forwarding the information received from the source node are considered as relay nodes. Naturally, the extra tele-traffic between a source MS and a cooperating MS serving as a relay station (RS) demands additional radio resources to be allocated - any of the well-established multiple access schemes can be employed by the users to guarantee their orthogonal interference-free transmission, such as Time Division Multiple Access (TDMA), Frequency Division Multiple Access (FDMA) or Code Division Multiple Access (CDMA) [5].
B. Base Station Cooperation

Similar to the cooperating single-antenna aided MSs, the cooperating BSs may also be considered as part of the family of MIMO schemes having distributed antenna elements. Their difference is that in the latter case the MIMO elements are connected by an optical back-bone, instead of a radio channel. The 3rd Generation Partnership Project’s (3GPP) Long Term Evolution (LTE) [10] initiative has attracted substantial interests across the wireless telecommunications industry, including the operators, manufacturers and research institutes. Further enhanced enabling techniques have been submitted to the ITU in the fall of 2009 for their consideration in the very recent 3GPP Releases known as the LTE Advanced (LTE-A) project, where the so-called Cooperative Multi-Point (CoMP) transmissions was formally proposed [11]. There are two different types of CoMP transmissions, namely Single-Cell Processing (SCP) based coordinated transmission and Multi-Cell Processing (MCP) based cooperative transmission, where the former scheme refers to classic Co-Channel Interference (CCI) avoidance techniques based on resource allocation and management, while the latter is constituted by the joint data transmission of multiple cells, mainly aimed at improving the throughput at the cell-edge. In [12], a comprehensive survey of various CCI mitigation techniques was provided.

The MCP based cooperative transmission regime shares the data of all the BSs invoked for jointly processing them [13]. This is typically achieved by assuming the existence of a Central Unit (CU), which connects all the BSs involved via a reliable high-speed optical fibre. However, MCP requires the Channel State Information at all the Distributed Transmitters (CSI-DT). There are two different MCP frameworks designed for sharing the CSI-DT, namely the centralised and decentralised framework [14]. More explicitly, the centralised framework exchanges the CSI of all the BSs involved with the aid of the CU, while the decentralised framework gathers the CSI of all the BSs involved at each individual BS locally.

C. Outline of the Paper

The rest of the paper is organized as follows. The family of cooperative relaying protocols is briefly reviewed in Section II-A followed by a discourse on network-coding-aided processing of multiple source’s information in the context of cooperative networks in Section II-B. The design of cross-layer cooperation aided MS cooperation and the concept of coalition network element (CNE) based relaying involved in BS cooperation are discussed in Sections II-C and II-D, respectively. Then, in Section III potential complexity reduction approaches are reviewed in the context of both BS and MS cooperation. Further specific discussions are dedicated to the design of MS cooperation dispensing with channel estimation in Section IV. Finally, we conclude in Section V by providing a cooperative system design guidelines.

II. RELAY-BASED COOPERATION

A. Cooperative Relaying Protocols and Classification

The underlying idea behind cooperative transmissions can be traced back to the pioneering work on the information theoretic features of the relay channel [4]. Motivated by this contribution, various cooperation strategies and protocols have been proposed. According to the operations carried out at the RS, the relaying protocols may be classified into three categories [1, 3], namely amplify-and-forward (AF), decode-and-forward (DF) and compress-and-forward (CF) relaying. The former two schemes were devised in [8], which have become the most popular ones because of their simplicity and intuitive designs. In the AF scheme, which is also referred to as the analog-repeater-based arrangement [6], the RS simply amplifies and forwards the source node’s ‘overheard’ signal to the intended destination, potentially increasing the system’s overall noise level, since the signal and noise are amplified together. In DF scheme, the RS fully decodes the signal received from the source and provides the destination with a re-encoded signal. Hence, the problem of error propagation may arise, when the RS forwards the erroneously recovered signal, which may deteriorate the detection at the destination and hence the overall system performance. It was recently demonstrated in [8,15] that the fixed DF system dispensing with any error-aware mechanisms at the RS offers no diversity gain over its conventional direct-transmission-based counterpart. Consequently, the selective DF scheme [8, 15] was devised with the aid of error detection codes and/or intelligent RS selection schemes, where the RS may forward the signal if and only if it is correctly decoded. Furthermore, when the signal radiated from the RS is channel encoded to provide extra error protection for the original message, the DF scheme is also known as coded cooperation [16–18]. Recently, the CF-based cooperative scheme also received increasing research attention [19,20], where the RS forwards a quantized or compressed version of the signal received from the source.

On the other hand, based on the time slots required to complete a full cycle of UL and DL transmissions, the family of cooperative relaying systems may be divided into another four subgroups, namely the traditional four-phase mechanisms, the network-coding-aided three-phase and two-phase schemes, as well as the successive relaying strategy, as portrayed in Fig. 1. As demonstrated by Fig. 1(a), although the four-
phase cooperative scheme, which is also referred to as one-way relaying, may achieve an enhanced transmit diversity gain, attain path-loss reductions, while retaining complete orthogonality between the broadcast and relaying phases, the system’s effective throughput is halved in comparison to the conventional direct-transmission scheme owing to the half-duplex communications of practical transceivers. Thus, it is hard to formulate an immediate judgement on whether the benefits of MS cooperation justify the cost incurred in the interest of increasing the achievable transmission efficiency. For example, recent research disseminated in [21] has revealed that the AF-based cooperative system may suffer from a significant capacity loss in comparison to the conventional direct-transmission system. Hence, the three-phase [22, 23] and two-phase [24, 25] bidirectional relaying schemes of Figs. 1(b) and 1(c) have been proposed in order to recover the effective throughput erosion, where advanced network coding techniques [26] are employed at the RS to generate and transmit a combined signal stream encapsulating both the DL and UL signals during the relaying phase. As shown by Figs. 1(b) and 1(c), the two-phase scheme requires less time slots to complete a full cycle of UL and DL transmissions, than its three-phase counterpart, albeit this is achieved at the expense of a typically worse decoding performance at the RS imposed by the mutual interference between the UL and DL signals in phase 1 of Fig. 1(c). Recently, the successive relaying technique of Fig. 1(d) has been devised in [27], which needs an additional RS for the sake of recovering the half-duplex-relaying-induced multiplexing loss. The successive relaying that is carried out by the pair of parallel RSs allows the source to transmit continuously, while still achieving second-order diversity and maintaining almost the same slot efficiency as the direct-transmission system, provided that the number of communications phases is sufficiently high. This technique was then further developed in [28] by assigning orthogonal CDMA sequences to the potentially interfering links. Hence second-order diversity was achieved at the cost of assigning two spreading codes to the cooperating users. Furthermore, space-time coding techniques [29–31] constitute another spatially-efficient approach applicable to cooperative systems, leading to the concept of distributed space-time coding schemes [32, 33]. For example, each cooperating RS can transmit a column of an orthogonal space-time code matrix during the relaying phase, as detailed in [1].

B. Efficient Processing of Source Information: Multisource Network Coding

Cooperative communications attracted substantial research interests in recent years [7, 34–36], spanning from the classic single-source single-relay scenario [37] to the generalised Multiple Source Multiple Relay (MSMR) scenario [38]. When considering the MSMR network topology, a fundamental issue is the efficient processing of numerous source information streams during their relaying [39].

The processing of multiple sources may be treated analogously to the classic multiplexing problem, which may be based either on an orthogonal or on a non-orthogonal Code Division Multiplexing (CDM) approach [40]. Specifically, the information-theoretically attractive superposition modulation aided multiple source cooperation scenario was considered in the context of two sources in [42] and for multiple sources in [43]. On the other hand, the relay may generate the ’XOR’ed information of the multiple source streams in the context of both the original bit-based Classic Network Coding (CNC) scheme [44, 45] and in the modified waveform-based Physical-layer Network Coding (PNC) arrangement [46, 47]. It is worth noting that the concept of both CDM and of CNC may be considered as a modulation technique, where the former scheme is implemented using arithmetic additions in the complex-valued domain, while the latter scheme is realised using modulo additions over the finite Galois field.

On the other hand, a coding-related interpretation may also be conceived for both CNC and PNC, because both techniques impose a certain encoding constraint, which is reminiscent of channel coding. Since the decoding (demapping) of CNC (PNC) for a large number of source information streams is non-trivial, the CNC and PNC concept is predominantly used in cooperative scenarios, when the number of source information streams is small. This specific scenario is encountered in two-way communications [48, 49] or for transmission over twin-source multiple access relay channels. To take a further step forward, the so-called joint channel and network coding [49] or multiplexed coding [50, 51] concept was proposed in order to provide an additional channel coding gain by imposing carefully designed redundancy, where the sources’ information streams are treated as a single amalgamated stream, before it is channel encoded.

Meanwhile, extensive research efforts have also been dedicated to Multiple Source Cooperation (MSC) [52, 53], which constitutes a specific instantiation of the MSMR scenario, where the relays are also active sources. A high throughput MSC framework was proposed in [43] and was extended to a multiplexed coding regime with the aid of a Low Density Generator Matrix (LDGM) based design [54]. Apart from the sophisticated joint channel and network coding schemes proposed in [43, 54] for MSC that rely on a channel code, the performance of the pure CNC scheme has not been explored in the context of MSC. Hence, in [55] a range of multiple source processing techniques were considered, ranging from the basic CDM concept to the CNC technique, where the soft decoding of CNC carried out with the aid of factor graphs was conceived, which is capable of reliable operation even in the presence of unreliable network information streams. Importantly, a novel Variable-rate Network Coding (VNC) regime was also proposed [55] that is capable of operating near the achievable capacity without necessitating a sophisticated

1Realistic transceivers cannot transmit and receive simultaneously, because at a typical transmit power of say 0dBm and receiver-sensitivity of $-100$dBm the transmit-power leakage imposed by the slightest power-amplifier non-linearity would leak into the receiver’s Automatic Gain Control (AGC) circuit and would saturate it. Hence, the saturated AGC would become desensitized against low-power received signals.

2Superposition modulation overlays several transmitted signals and hence results in a near-Guassian-distributed signal. Therefore, it approaches the Continuous Input Continuous Output Memoryless Channel’s (CCMC) capacity [41].
joint channel and network code design. Finally, the linkage of classic modulation and the new concept of network-coded modulation was established in [56].

C. Cooperative Relaying in MS Cooperation: Cross-Layer Cooperative Protocol Design

The benefits of cooperative communications may be eroded by the conventional higher layer protocols, which were designed for classic non-cooperative systems. Hence, it is important to design appropriate Medium Access Control (MAC) protocols for supporting cooperative physical layer techniques. Most recent cooperative MAC protocols were designed for maximizing the throughput and for reducing the outage probability [57–63]. Often the energy efficiency was hence traded off against these benefits. Additionally, some contributions minimized the energy consumption by developing energy-efficient cooperative MAC protocols, but these often remained oblivious of the associated throughput performance [64–67]. By contrast, both Zhao et al. [68] and Shirazi et al. [69] designed meritorious algorithms for improving the achievable throughput, while reducing the energy consumption imposed. However, the above-mentioned cooperative MAC protocols were developed on the basis of the common assumption that the relays agree to altruistically forward the data frames of the source. This unconditional altruistic behavior is unrealistic to expect for the mobile terminals.

In order to consider the either selfish or ‘win-win’ behavior of the mobile relays, Stanojev et al. [70] proposed an auction-based cooperative Automatic Repeat reQuest (ARQ) scheme relying on a so-called spectrum-leasing paradigm. However, the attainable energy efficiency was not quantified in this cooperative ARQ scheme. As a further advance, Mukherjee et al. [71] developed an auction-theoretic cooperative partner selection scheme for striking a tradeoff between the attainable throughput and energy efficiency. However, the potentially corrupted data received from the direct transmission link was not actively exploited with the aid of frame combining, when the destination attempted to retrieve the source data frame. Furthermore, no particular transmission frame structure and signaling procedures were designed in [70, 71].

Against the above background, a cooperative MAC-layer protocol was proposed in [72] for a network supporting the source with the aid of relays for the sake of minimizing the total energy consumption and for improving the source’s throughput, while simultaneously conveying the relay’s own traffic. The proposed cooperative MAC-layer protocol benefits from auction-style single relay selection for striking a tradeoff between the achievable throughput and energy efficiency for both the source and relay in a practical network scenario, where the proposed idea was implemented using a signaling procedure that is compliant with the 802.11 legacy protocol. More particularly, superposition coding [73] is invoked at the relay for encoding both the source’s and relay’s data and beneficially amalgamates the direct and relayed components using frame combining.

D. Cooperative Relaying in BS Cooperation: Coalition Network Elements

Naturally, the presence of imperfect and outdated CSI at the cooperative BS transmitters as well as the limited backhaul throughput will erode the efficiency of this MCP-aided mitigation technique in theory. A straightforward solution to eliminate the effects of malfunctioning MCPs is to employ Automatic Repeat reQuest (ARQ) type retransmissions from the cooperating BSs. By contrast, the joint potential of BS cooperation and relaying was explored in [74] with the goal of mitigating the effects of the CCI, where the BS cooperatively transmits to the cell-edge MSs in the first hop and the so-called remote coalition Network Element (CNE) is responsible for the second-hop transmission, provided of course that the latter are available. To elaborate a little further, the CNE carries traffic for the primary BSs to the critical cell-edge area in the unutilised frequency bands of the primary network, where the availability of these free channels is explicitly signalled to the BSs, rather than being sensed. Hence, this approach is reminiscent of the cooperative cognitive philosophy [75]. In contrast to the conventional relaying, the CNE will reserve part of its resources assigned by the BSs for its own use and leave the rest of it for cooperative transmission to the cell-edge MSs. Hence, the CNE is capable of acting as a fall-back solution in support of the primary BS cooperative transmission, when for example one of the BSs malfunctions due to impairments, such as CSI estimation errors, CSI quantisation errors and CSI feedback errors imposed by channel errors and latency. As a result, the cell-edge MSs will benefit from additional spatial diversity upon combining the pair of independent copies received from both the BSs and CNE activated in the two-hop scenario. The specific improvement attained will depend on how ‘greedy’ or altruistic the CNE is and how many idle channels are available in the primary network.

III. COMPLEXITY REDUCTION IN COOPERATIVE NETWORKS

A. Complexity Reduction in BS Cooperation: Reducing CSI and Data Exchanges

To provide the required CSI, the quantised version of each user’s CSI estimated at the MS’s DL receiver may be fed back to the BS transmitters using a finite-delay, limited-rate feedback link assuming a Frequency Division Duplex (FDD) system [76]. Hence the resultant CSI-DT may suffer from both quantisation noise as well as feedback errors. This undesirable phenomenon dominates the achievable MCP performance, when various linear precoding techniques are employed. The family of DL precoding techniques may be invoked at the BSs for eliminating the effects of CCI at the BS transmitter for all MSs, hence potentially facilitating the employment of ‘low-complexity’ single-user MS receivers. The optimal Dirty Paper Coding (DPC) aided precoding technique [77] imposes a high computational complexity, thus it is less attractive than other low-complexity linear precoding techniques. In the context of MCP, linear BS precoding techniques may be implemented in either a joint or distributed fashion. Linear joint DL precoding techniques generally determine the precoding matrix for all the
BSs involved. By contrast, distributed linear precoding techniques optimise the DL precoding matrix of each individual BS locally.

Although individual reports on the attainable MCP performance of linear precoding techniques may be found in the literature, they are based on different system configurations associated with different assumptions. In [78], a comparative study of the various joint and distributed linear precoding techniques was provided for both centralised and decentralised CSI-DT scenarios in the presence of potential CSI feedback errors. As a further step, since most of the backhaul-limited MCP research was concentrated on either reducing the required CSI-DT or (dynamically) determining the number of actively cooperating BSs, the challenges of MCP relying on reduced data - rather than reduced CSI feedback - exchange have not been explored in the open literature. Hence, for the sake of further reducing the burdens imposed on practical limited-rate backhaul design, in [79], a range of reduced-complexity MCP structures employing distributed linear precoding was proposed relying on a reduced amount of data exchange, where the different BSs have to carry out different amounts of processing and information exchange. The performance of various reduced-complexity MCP structures was investigated in terms of their achievable throughput without encountering an outage rate, which demonstrated the attractive throughput improvements over the conventional SCP scheme and their different geographic rate profile distributions. The delay performance of the best-supported MS and worst-supported MS of various reduced-complexity MCP structures was also investigated, which demonstrated the capability of supporting different Quality of Service (QoS) requirements.

B. Complexity Reduction in MS Cooperation: Dispensing with Channel Estimation

In practice, the employment of channel estimation for all mobile-to-mobile links in MS-cooperation-based systems may become unrealistic, since it may impose both an excessive complexity and a high pilot overhead, especially when the number of cooperating MSs is high and/or when the channel conditions fluctuate relatively rapidly in mobile environments. Moreover, it is particularly challenging for the BS to accurately estimate the source-relay channel using pilots in the context of AF-based cooperative systems, since the pilots may be further contaminated by noise amplification. Furthermore, a significant performance erosion may be imposed by inaccurate CSI as demonstrated in [80, 81] in the context of cooperative systems. Therefore, differentially encoded signaling combined with low-complexity non-coherent detection and thus bypassing the complex yet potentially inaccurate channel estimation process at the receiver becomes an attractive design alternative, leading to differential modulation assisted cooperative communications [3, 82–87]. Thus, a simple receiver robust may be implemented for the MSs, which is robust against the phase ambiguities induced by rapid fading, while dispensing with complex timing recovery and channel estimation for the mobile-to-mobile links. Naturally, in the light of the distributed space-time coding principles, the differential space-time coding regime can also be implemented in a distributed manner for user-cooperation aided systems [88–90].

IV. OPEN ISSUES ON MS COOPERATION DISPENSING WITH CHANNEL ESTIMATION

In view of the benefits of bypassing the potentially excessive-complexity and yet inaccurate channel estimation, the family of differential modulation schemes combined with non-coherent detection is advocated in this treatise as a viable candidate to be employed for MS-cooperation-based systems. The conception of MS cooperation dispensing with channel estimation naturally leads to a number of new challenges, among others the design of robust non-coherent detectors, appropriate cooperating cluster formation, resource allocation, multiuser/multistream interference management as well as adaptive rate control, some of which will be detailed in the ensuing sections.

1) The Need for Robust and Flexible Non-Coherent Detectors: The low-complexity conventional differential detector (CDD) [91] employed at the receiver may extract the data by simply calculating the phase difference between consecutive time samples, provided that the rate of the CIR fluctuation is sufficiently low. However, this low-complexity processing is facilitated at the cost of the potential formation of a high-Doppler-induced error-floor. Specifically, when the channel linking the cooperating MSs becomes more time-selective in high-velocity mobile environments, the slow-channel-fluctuation prerequisite imposed by the CDD no longer holds. Hence, a potentially significant performance degradation is expected for CDD-aided differentially encoded transmissions, which implies that the cooperative diversity gains achieved by the CDD-aided cooperative system may also erode, as shown in Fig. 2, where an uncoded differentially encoded cooperative system in [92]. The MSDSD has been specifically designed for a differentially encoded non-coherently detected cooperative system’s Bit Error Rate (BER) performance is exemplified. Hence, we will propose flexible solutions for striking a balance between the performance achieved and the complexity imposed in typical dynamic wireless environments.

(a) Combating Channel Fluctuations: In pursuit of an improved resilience against the high-Doppler-induced performance degradation, one may resort to the employment of multiple-symbol differential detection (MSDD) [93, 94], which jointly detects \( N_{\text{wind}} \) number of symbols, hence exploiting the correlation between the phase distortion experienced by the consecutively transmitted differential phase shift keying (DPSK) symbols. The complexity of the MSDD, which increases exponentially with the detection window size \( N_{\text{wind}} \), may be substantially mitigated with the aid of the sphere detection (SD) mechanism, yielding the so-called multiple-symbol differential sphere detection (MSDSD) [95]. Recently, the MSDD has been specifically designed for a differentially encoded non-coherently detected cooperative system in [92]. Observe in Fig. 2 that the high-Doppler-induced error floor was essentially eliminated with the aid of the MSDSD employed at both the MS and BS.

(b) Enhancing the Iterative Gains Attained by Turbo Receivers: As another benefit in addition to the robustness against
the high-velocity mobility-induced performance degradation, it is worthwhile noting that the MSDSD is also capable of increasing the iterative gain attained by the turbo receiver in the context of channel-coded systems. This is because the generation of soft-information by the MSDSD for the bits within the same detection window benefits from exploiting each other’s improved-confidence reliability information provided by the channel decoder. As a result, the enhanced iterative gain attained by the MSDSD-aided turbo receiver for each direct transmission link may be translated to an increased error-free transmission rate for MS-cooperation-based systems, as exemplified in Fig. 3.

(c) Trade off between performance and complexity: Since the channel conditions of each mobile-to-mobile and mobile-to-BS link typically fluctuate due to both the mobility of the MSs themselves as well as owing to that of their surrounding objects, meeting stringent QoE requirements in hostile wireless environments may become unrealistic for the low-complexity but inflexible CDD. Subsuming the CDD as its special case when the detection window size is \( N_{\text{wind}} = 2 \), the MSDSD is capable of striking a flexible compromise between the achievable performance and the imposed complexity, when adaptively choosing an appropriate detection window size according to the time-varying channel conditions and/or to the prevalent QoE requirements. For example, an adaptive window scheme was proposed in [96] for the single-relay-assisted cooperative system in order to achieve a near-capacity performance at a moderate complexity. In the light of the above discourse, the MSDSD constitutes a promising candidate for employment in the differentially encoded MS-cooperation-based systems.

(d) The Design of High-Order Differentially Encoded Modulation: In pursuit of high bandwidth efficiency, differential amplitude and phase shift keying (DAPSK) was devised [97–99] using constellations of multiple concentric rings. However, this non-constant-modulus constellation precludes the direct application of the SD technique for the complexity reduction of MSDSD assisted DAPSK systems. Until very recently the conception of an efficient MSDD for DAPSK-aided systems has been an open problem. This open problem was then closed by the proposal of an iterative amplitude/phase (A/P) detection framework for MSDSD-aided DAPSK systems in [100]. The iterative information exchange between the decoupled A/P detection stages was specifically tailored for mitigating any potential performance penalty imposed by the separate - rather than joint - A/P detection stages. For the sake of achieving a further complexity reduction, the SD mechanism can be incorporated in the computationally demanding phase detection stage, which contributes the majority of the total complexity imposed.

2) Resource Optimization for Differentially Modulated Mobile Station Cooperation: Although it is well-recognized that a full spatial diversity may be achieved for MS-cooperation-based systems [7, 8], the achievable end-to-end performance may significantly depend both on the specific choice of the cooperative protocols employed and/or on the cooperative resource allocation. Hence, the design of flexible cooperative protocols, the appropriate cooperating cluster formation strategies, as well as the conception of matching cooperative resource allocation procedures become necessary in order to further enhance the attainable performance and to maximize the overall system capacity.

(a) Power-Related Resource Allocation: The transmit power sharing and allocation amongst the cooperating MSs plays a crucial role in the performance enhancement of MS-cooperation-based systems. Hence, this topic has attracted im-
mense attention from the entire research community. Since the average power assigned to the mobile-to-mobile and mobile-to-BS links is essentially related to the roaming MS’s location, the cooperating cluster formation may also be regarded as a power-related resource allocation technique. Various optimization criteria have been adopted for the power-related resource allocation strategy, such as for example the minimum BER/SER optimization strategy [82, 86, 101] and the minimum outage probability based policy [84, 86], etc. The comparative study of the differential AF- and differential DF-aided cooperative systems designed in [101] indicated these two relaying mechanisms tend to exhibit complementary characteristics, reflected for example by their distinct optimum cooperative resource allocations. Hence, for the sake of exploiting the complementarity of these distinct relaying schemes, a flexible hybrid cooperative regime may be conceived, where different schemes may be activated in diverse scenarios [101, 102]. More specifically, as shown in Fig. 4, in contrast to the conventional MS-cooperation-based system employing a single cooperative mechanism, the cooperating MSs roaming in different areas between the source MS and the BS may be activated and the relaying schemes employed by each activated MS may be adaptively selected. The beneficial application examples of hybrid cooperative relaying schemes designed in [101, 102] were demonstrated to be capable of significantly enhancing the achievable BER and/or outage probability performance of the cooperative system, while maintaining a moderate complexity, thus indicating the need for developing new, flexible hybrid cooperative protocols.

(b) Time-Resource/Code-Rate Optimization: Since the majority of TDMA-based cooperative system optimization efforts have been focused on power allocation and RS selection [8, 101, 104], the time slot-duration resource allocation (TRA) between the source and RS has remained an open problem until recently. To resolve this open design issue, the TRA problem was investigated in [105] in order to maximize the so-called effective capacity in a two-source single-relay-aided system. The optimum TRA policy was then deduced in [103] for the sake of maximizing the differentially encoded cooperative system’s capacity. These contributions become useful in the design of near-capacity channel coding/decoding schemes conceived for cooperative systems [96], since the code rate employed by the source and RS is directly related to their allocated transmission slot-duration, which may in fact be adaptively selected according to the proposed TRA scheme. Fig. 5 demonstrates that a significant capacity gain can be achieved with the aid of the TRA scheme [96] for the single-relay-aided MS-cooperation-based system. Note that the increasing value of the optimal TRA factor $\alpha$ inferred from Fig. 5 indicates that longer time-slots should be allocated to the source MS when the SNR is high, which implies assigning lower-rate channel codes to the source than to the relay.

3) Multiple-Access Interference Management without CSI: When aiming for sharing a given frequency/time-slot with the aid of Spatial Division Multiple Access (SDMA) by several users, the users or data-streams are classically differentiated with the aid of their unique CIRs. However, dispensing with channel estimation in differentially modulated user-cooperation-based systems imposes another challenging problem, namely that of managing the multiple-access interference (MAI) at the BS in spatial domain without CSI. One possible solution is to estimate the MAI and cancel it with the aid of adaptive receiver for the desired user. For example, the adaptive minimum mean square error (MMSE) criterion [106] using the least mean square (LMS) or the recursive least squares (RLS) algorithm could be used. Alternatively, the more recently proposed maximum signal-to-interference-plus-noise ratio (MSINR) based differential interference suppression (DIS) scheme of [107] may be employed. For the former the coefficients of the interference suppression filter are adapted in order to minimize the MSE between the transmitted signal and the filter’s output signal, while for the latter the MAI-suppression filter coefficients are adjusted to maximize the SINR at its output. As demonstrated in [107], the DIS scheme is additionally capable of mitigating the effects of carrier phase variations. Although they do differ in their concept, the
MSINR solution subsumes its MMSE-based counterpart as a special case [108].

Inspired by the block least-squares algorithm of [106], which was originally designed for standard MMSE criterion-based coefficient adaptation, a new adaptive multiple-symbol DIS (MS-DIS) scheme has been proposed recently in [109]. This solution is based on the multiple-symbol differential SDMA (MS-DSDMA) system model, which was designed for the sake of reducing the filter adaptation overheads and, even more importantly, for facilitating the employment of the low-complexity yet powerful MSDSD of [95]. Meanwhile, as a benefit of employing the MSDSD [95], extra coding gains may be gleaned for differentially encoded systems by exploiting the correlation between the phase distortions experienced by the consecutively transmitted symbols. In order to further increase the achievable differential detector’s performance in the context of our adaptive MS-DIS scheme, a new channel-code-aided three-stage turbo DIS receiver has been proposed in [109], which facilitates a beneficial information exchange amongst the concatenated adaptive MS-DIS filter bank, the MSDSD and the channel decoder.

4) In Pursuit of Near-Capacity Operation: Inspired by the idea of distributed turbo codes [17] proposed for “distributed MIMO” systems, a novel Irregular Distributed Differential (IrDD) coding scheme has been conceived in [96] for the differential DF-aided cooperative system, in order to achieve a near-capacity performance. Specifically, the near-capacity design of the transceiver employed in [96] at the MS and BS was reduced to an (EXtrinsic Information Transfer) EXIT curve matching problem, which served as the fundamental method invoked for approaching the cooperative network’s capacity for the single-relay-aided user-cooperation-based system\(^3\). It was also demonstrated that the joint source-and-relay mode design procedure of the single-relay-aided cooperative system can be decoupled into two separate EXIT curve matching problems. Although it was demonstrated in [96] that the IrDD-aided user-cooperation-based system was indeed capable of performing close to the system’s non-coherent DCMC capacity, the system had to be re-designed in an offline manner, if the system’s operating SNR was changed in order to maintain a near-capacity performance. Therefore, in pursuit of maintaining high-bandwidth-efficiency communication in dynamically fluctuating wireless environments, the design of a joint adaptive modulation and coding rate control assisted user-cooperation-based system dispensing with CSI estimation is necessary, which remains an open problem at the time of writing.

V. CONCLUSIONS AND DESIGN GUIDELINES

In Fig 6 we classified the subject of cooperative communications into BS cooperation and MS cooperation which are presented by the overlapping ellipses. The intersection of these two sets highlights the key issues that should be taken into account when designing cooperative communication systems, which are related to the resource limitations and general communication system design objectives. The key design problems are highlighted for both BS and MS cooperation within the respective ellipses. The scattered keywords around these two design ellipses allude to the available advanced enabling techniques, ranging from the related transceiver design issues to air-interface techniques and to high-layer protocols. In addition to the above qualitative portrayal of the associated problems, below we list a range of important design guidelines based on our original research:

- In order to design a cooperative system, one may first identify the most pertinent Quality of Service (QoS) metrics as well as other constraints according to the application at hand. For example, delay-sensitive or delay-tolerant as well as bandwidth- or power-limited applications require different designs.
- From a physical layer point of view, we may amalgamate the best possible transceiver components, such as near-capacity channel coding, iterative detection and appropriate multiple access / random access schemes etc. A range of influential design factors must be considered, including but not limited to the level of interference, the presence or absence of channel knowledge, the tolerable computational complexity, transceiver’s robustness, etc.
- In order to facilitate cross-layer design, a holistic view of the upper layers’ behaviour should be jointly considered, bearing in mind for example the queuing model, the routing model and the TCP model etc. In general, this may lead to a multi-objective optimisation problem, which may be solved with the aid of a semi-analytical approach. Last but not least, since the associated non-linear dynamic control problems typically rely on feedback, the stability of the cross layer design should always be tested so as to avoid any potential instability.

REFERENCES


\(^3\) The near-capacity EXIT-chart-based designs detailed in [3] rely on exploiting that the area between two iterative decoder components is proportional to the SNR-discrepancy with respect to capacity. Hence, the components have to be designed to have the lowest possible area between them, which is achieved by matching their EXIT curve.
Fig. 6. Designing issues of BS and MS cooperation systems


