Chapter 13
Interleave Division Multiple Access (IDMA)

Yang Hu and Li Ping

13.1 Overview

The capacity of a multiple access channel was studied in [1, 2]. It can generally be achieved by random coding together with other techniques, e.g., power control, linear precoding, and dirty paper coding [3–6]. Random coding does not involve orthogonality among users so it is inherently non-orthogonal. The sub-optimality of orthogonal multiple access (OMA) was investigated in [7, 8]. The gain of non-orthogonal multiple access (NOMA) over OMA was assessed in [9, 10] for both single-input single-output (SISO) and multiple-input multiple-output (MIMO) systems. Recently, NOMA has been promoted for improving system fairness [11, 12].

However, many practical systems still belong to OMA category. This is mainly due to complexity concerns. OMA can work with low-cost single-user detection (SUD), while NOMA may require more complex multi-user detection (MUD). Thus, these two options entail different trade-off between cost and performance.

Historically, the third-generation (3G) direct-sequence code-division multiple access (DS-CDMA) system is non-orthogonal. Normally, only SUD is used in DS-CDMA to avoid high complexity, which is sub-optimal. The spreading operation in DS-CDMA reduces rate, so DS-CDMA is not convenient for high-rate applications. The fourth-generation (4G) orthogonal frequency division multiple access (OFDMA) system returns to OMA. OFDMA allows flexibility for resource allocation over time and frequency, which can bring about noticeable gain.

Y. Hu (✉) · L. Ping
Department of Electronic Engineering, City University of Hong Kong,
Hong Kong, SAR, China
e-mail: yhu228-c@my.cityu.edu.hk
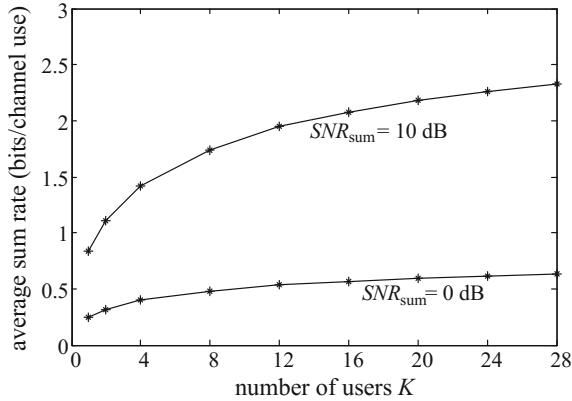
L. Ping
e-mail: eeliping@cityu.edu.hk

Fig. 13.1 Achievable sum-rate of NOMA and OMA under equal energy constraint. Note that the two strategies have exactly the same performance. SNR_{sum} is the sum signal-to-noise ratio (SNR) of all users. Complex channels with both slow- and fast-fading factors are considered. Path loss is based on a hexagon cell with a normalized side length = 1. The minimum normalized distance between users and the base station is 35/289, corresponding to an unnormalized distance of 35 m for an LTE cell with radius 289 m. Path loss factor = 3.76 and lognormal fading deviation = 8 dB. The channel samples are normalized such that the average power gain = 1

Recently, NOMA has been discussed widely for the fifth-generation (5G) [13–16]. A natural question is whether the possible gain of NOMA can justify its higher receiver cost. We examine this question using achievable sum-rate below.

In Fig. 13.1, we compare the achievable sum-rate of NOMA and OMA with equal energy constraint per frame per user in a multi-user SISO system. We can see NOMA improves sum-rate when the number of users K increases, but OMA can achieve the same gain through resource allocation. NOMA has no advantage here since OMA is capacity achieving in this case.

The situation is slightly different if the energy per frame per user can also be freely optimized under the sum energy constraint. For example, consider maximizing sum-log-rate under the proportional fairness criterion [3]. Figure 13.2 illustrates the related numerical results for $K = 2$. The sum-rate curves in Fig. 13.2 show that NOMA is only slightly better than OMA with resource allocation (about 8% gain at $SNR_{sum} = 10$ dB).

The advantage of NOMA over OMA seen in Figs. 13.1 and 13.2 is disappointing compared with many results in the literature. This is mainly for two reasons. First, comparisons with OMA without resource allocation are not fair as resource allocation has already been widely used in LTE. Second, a practical signal-to-noise ratio (SNR) range should be used for comparison. A standard way for this purpose is using the following approximation of the signal-to-noise-plus-interference ratio (SINR) in a cellular system [3]:

$$SINR_{sum} = \frac{P_{sum}}{\beta P_{sum} + \sigma^2}, \qquad (13.1)$$

Fig. 13.2 Achievable sum-rate of NOMA and OMA under sum energy constraint with proportional fairness, $K = 2$. Other system settings are the same as those in Fig. 13.1
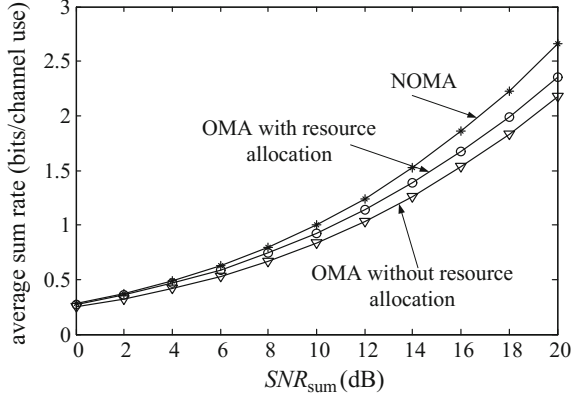
where P_{sum} is the sum received powers of all users in a cell, β a cross-cell interference factor and σ^2 the noise power. As a rule of thumb, a typical value is $\beta = 0.6$. Then, we have $SINR_{sum} \leq 2.2$ dB. Treating interference as noise and allowing a certain range of β, we may consider 0–10 dB as a typical range for SNR_{sum}, as used in Figs. 13.1 and 13.2. Clearly, the gain of NOMA is marginal over such a practical SNR range.

Furthermore, NOMA performance may deteriorate seriously if a practical forward error control (FEC) code is used. To see this, consider a successive interference cancelation (SIC) process with K users in a descendant order of user index k. We employ an FEC code that can achieve (almost) error-free decoding at SNR = Γ. Assume that, when we decode for user k, the signals of all users with indexes $k' > k$ have been successfully decoded and subtracted from the received signal. Let q_k be the received power of user k. Then, user k can achieve error-free decoding provided that

$$SNR_k = \frac{q_k}{\sum_{k'<k} q_{k'} + \sigma^2} \geq \Gamma. \tag{13.2a}$$

For sum-power minimization, q_k can be calculated using the following recursion (with $q_0 = 0$):

$$q_k = \Gamma \times \left(\sum_{k'<k} q_{k'} + \sigma^2 \right). \tag{13.2b}$$

Ideally, the minimum value for Γ can be found from Shannon capacity $R = \log_2(1 + \Gamma)$. For a practical code, however, a larger Γ is required to ensure (almost) error-free decoding. This overhead accumulates during SIC, which can amount to a considerable loss.

Specifically, we consider a practical rate-1/6 channel coding as an example. It achieves bit error rate (BER) $\approx 10^{-5}$ (approximately error-free) at about SNR = -3.5 dB with a relatively short block length. There is a 2.35 dB gap compared with SNR = -5.85 dB calculated from the Shannon formula for Gaussian signaling. Such

Fig. 13.3 Accumulated loss
in the SIC process
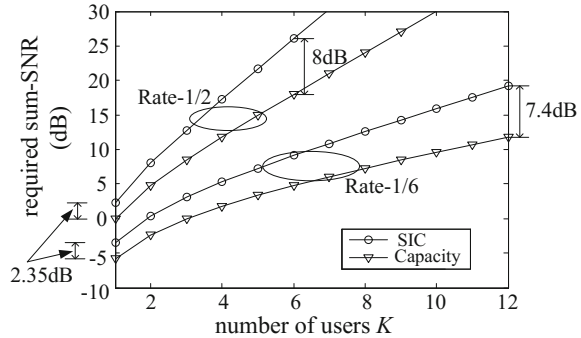
loss accumulates in the SIC process as shown in Fig. 13.3. The accumulated loss is
roughly 7.4 dB for 12 users.

The problem is more serious for higher rate, as seen from Fig. 13.3 for rate-1/2
coding. Assume the same initial single-user gap of 2.35 dB. The accumulated gap
for six users is 8 dB.

Nevertheless, NOMA is still useful. In many situations, it is difficult to establish or-
thogonality due to the lack of centralized control or accurate channel state information
(CSI). Then, we may have to resort to NOMA. In particular, as we will show below,
NOMA based on interleave-division multiple access (IDMA) [17–22] offers robust
and flexible solutions in such environments. IDMA can also recover a considerable
portion of the accumulated loss suffered by SIC as shown in Fig. 13.3. We will use
numerical results to verify these claims. Some of the software used in this chapter are
available at: http://www.ee.cityu.edu.hk/%7Eliping/Research/Simulationpackage/.

13.2   Basic Principles of IDMA

Following the advent of turbo and low-density parity-check (LDPC) codes [23–
25], iterative detection techniques were developed in late 1990s for equalization in
multipath channels [26] and MUD in DS-CDMA systems [27–29]. It was shown in
[30] that two independently interleaved code sequences can be separated by iterative
detection. This inspired the IDMA scheme in which users are solely separated by
interleavers [17]. Intuitively, a randomly interleaved code results in a different code.
Thus we may also say that different users in IDMA are separated by different codes.
This follows the basic principle of CDMA, except now the set of codes are generated
by a master code followed by different interleavers. We therefore can regard IDMA
as a special case of CDMA. However, IDMA is fundamentally different from DS-
CDMA. The former does not rely on spreading for user separation and so can avoid
the rate loss suffered by the latter.

In the following, we will start with a graphic model originally presented in [31].
We will show that the primary motivation behind IDMA is to break short cycles,

since the latter is detrimental for message-passing detection. The same principles have been successfully used in turbo and LDPC codes.

13.2.1 IDMA Transmitter Principles

Throughout this chapter, we will assume an underlying OFDMA layer that removes inter-symbol interference (ISI). We will focus on uplink multiple access techniques built on this OFDMA layer.

Let K be the user number and $\boldsymbol{c}_k = \{c_k(j), j = 1, 2, \ldots, J\}$ a length-J codeword generated by users k. The transmitted symbols $\{x_k(j), j = 1, 2, \ldots, J\}$ are generated from $\{c_k(j)\}$ after certain operations, such as spreading, scrambling, interleaving, and modulations. At the receiver, the received signals $\{y(j)\}$ are given by

$$y(j) = \sum_{k=1}^{K} h_k \sqrt{e_k} x_k(j) + \eta(j), j = 1, 2, \ldots, J, \tag{13.3}$$

where h_k is the channel coefficient for user k, e_k the transmitted power of user k and $\eta(j)$ an additive white Gaussian noise (AWGN) sample with variance σ^2 per dimension.

Figure 13.4 illustrates the factor graph representation [32] for (13.3) with $K = 2$, $J = 8$, and the same LDPC coding for all users. A circle in Fig. 13.4 represents a variable and a square a constraint. Three types of constraints are involved, namely a square marked with "+" for linear additions in (13.3), a white square for LDPC coding, and a square marked with "×" for modulation.

Optimal detection for the system in (13.3) typically requires prohibitively high complexity. Low-cost message-passing detection, similar to that used for LDPC codes, can be applied instead, as Fig. 13.4 is sparse when $K \ll J$. However, short cylices constitute a problem. To see this, let us call a circle involving m coded
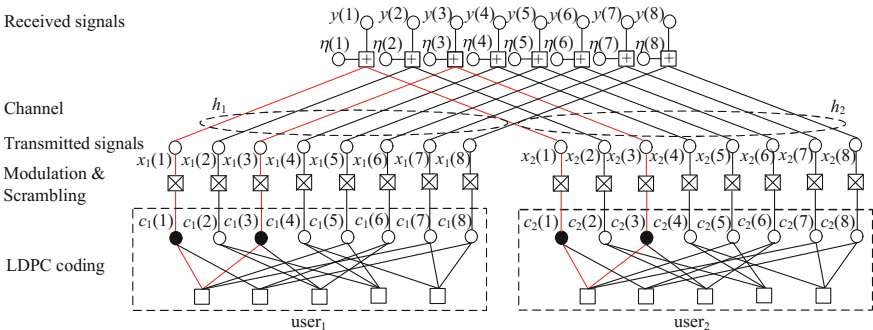

Fig. 13.4 Factor graph of a NOMA LDPC-coded system where $K = 2$ and $J = 8$
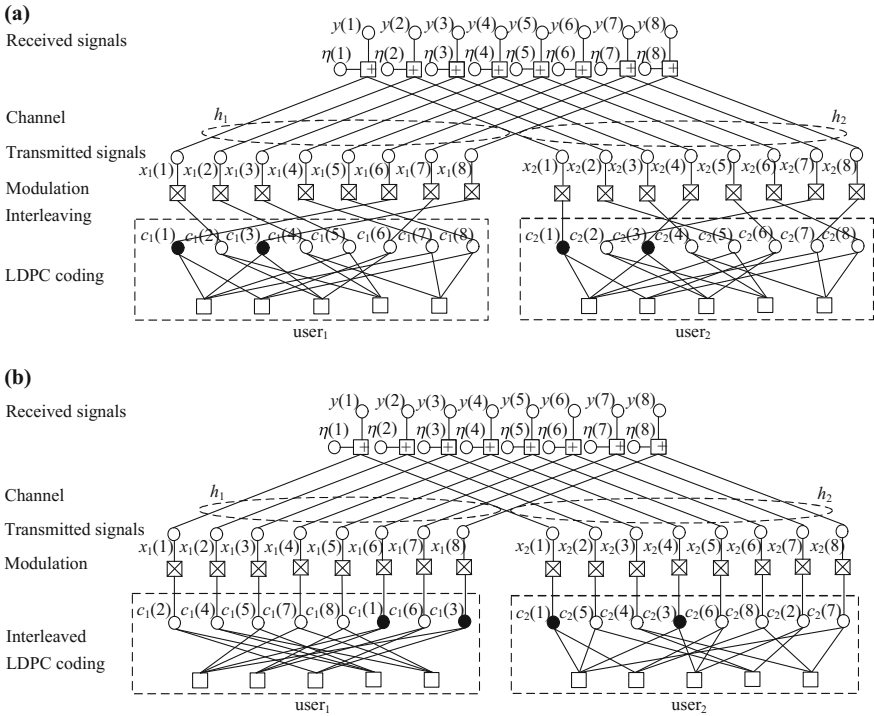
Fig. 13.5 **a** Factor graph of a two-user IDMA system. **b** An equivalence form of (**a**) with re-shuffled $\{c_k(j)\}$. Note that in (**a**) the interleavers for LDPC coding are the same for both users, while in (**b**) they are different

bits as a size-m cycle. An example of a size-4 cycle is shown by the black circles $\{c_1(1), c_1(3), c_2(1), c_2(3)\}$ in Fig. 13.4. There are a large number of such size-4 cycles in Fig. 13.4. Correlation may build up along these short cycles during message-passing detection, which is detrimental to performance [33].

Figure 13.4 can also be used to represent a DS-CDMA system. For example, a spreading operation involving binary sequence of ± 1 can be regarded as a repetition code plus a scrambling operation of sign changes. Repetition coding can be merged with FEC coding. Scrambling can be incorporated in function of the modulation nodes in Fig. 13.4. Note that scrambling does not change the topology of the graph. The problem of short cycles remains the same with or without scrambling. The same conclusion applies to other modulation techniques.

Inspired by the success of turbo and LDPC codes [23, 25], the IDMA scheme proposed in [31] employs user-specific interleaving to reduce short cycles in a statistical sense. This is illustrated by the shuffled edge connections between $\{c_k(j)\}$ and $\{x_k(j)\}$ in Fig. 13.5a. For example, compared with Fig. 13.4, $\{c_1(1), c_1(3), c_2(1), c_2(3)\}$ no longer form a size-4 cycle after interleaving in Fig. 13.5a. This is beneficial for

Fig. 13.6 An MA node. Here, an inbound message $LLR_{\mathrm{DEC}}(x_k(j))$ is from an LDPC decoder. An outbound message $LLR_{\mathrm{ESE}}(x_k(j))$ is generated from the MA node. ESE stands for "elementary signal estimation" and DEC for decoder
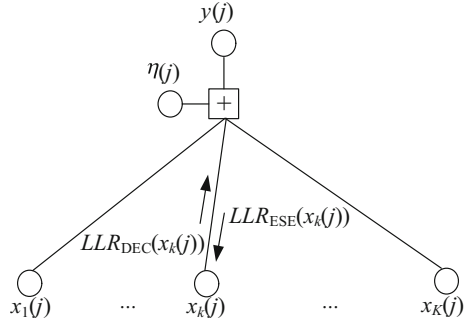
message-passing detection. Similar principles can also be applied to systems involving convolutional or turbo coding.

Two interleavers are used by each user in Fig. 13.5a, one for LDPC encoding and one for multiple access. The former is the same for all users, and the latter is user-specific. We can combine these two interleavers by re-shuffling $\{c_k(j)\}$ for each user, resulting in Fig. 13.5b. The latter interpretation of IDMA is based on [34].

The advantage of Fig. 13.5b is its simpler implementation. If an interleaver is random, its shifted version can be approximately regarded as another independent random interleaver. Thus, Fig. 13.5b can actually be realized by using a shifted version of an LDPC encoder involving a common underlying interleaver. The users, in this case, are separated by the amount of shift after the common encoder structure. Such shifting can be realized with very little cost. We may name such a scheme as code-shift division multiple access (CsDMA).[1] This concept was first discussed in [35].

Note that the use of interleavers in Fig. 13.5 does not incur any rate loss. This is a noticeable advantage of IDMA for high-rate applications.

13.2.2   Operations on a Multiple Access Node

We now consider message-passing detection based on Fig. 13.5. For convenience, we refer to a square marked with "+" in Fig. 13.5 as a multiple access (MA) node. We will only discuss the operations for such an MA node shown in Fig. 13.6. The operations for other nodes follow the standard treatments for an LDPC code [25].

Denote by DEC k the decoder for user k. We define two messages: an inbound message $LLR_{\mathrm{DEC}}(x_k(j))$ and an outbound message $LLR_{\mathrm{ESE}}(x_k(j))$ that are, respectively log-likelihood ratios (LLRs), generated by DEC k and elementary signal estimation

---

[1]The underlying code in CsDMA should be properly interleaved. An LDPC code naturally meets this requirement. Without interleaving, however, the correlation among the consecutive bits in a convolutional code may cause a problem in CsDMA. This problem can be easily avoided by shuffling the coded sequence.

(ESE) operations at the jth MA node in Fig. 13.6. The discussions for $LLR_{\text{DEC}}(x_k(j))$ follow the standard LDPC decoding principles [25] and so will be omitted. We will focus on $LLR_{\text{ESE}}(x_k(j))$ below since its computation is not part of a standard LDPC decoder.

For simplicity, let us first assume binary phase-shift keying (BPSK) modulation $x_k(j) = \pm 1$ for all k. We define $LLR_{\text{ESE}}(x_k(j))$ by the following LLR for $x_k(j)$:

$$LLR_{\text{ESE}}(x_k(j)) = \log \frac{\Pr(y(j)|x_k(j) = +1)}{\Pr(y(j)|x_k(j) = -1)}, \tag{13.4}$$

where $y(j)$ and $x_k(j)$ are defined in (13.3). Assume that $\eta(j)$ in (13.3) is Gaussian with mean $\mu(j) = \text{E}(\eta(j))$ and variance $v = \text{Var}(\eta(j))$. (For simplicity, we will assume that v is not a function of j. We will explain the rationale for this assumption later.) For a single-user system with $K = 1$ in (13.3), the conditional probabilities in (13.4) are given by

$$\Pr(y(j)|x_1(j) = \pm 1) = \frac{\exp\left(-\frac{(y(j)-(\mu(j)\pm h_1\sqrt{e_1}))^2}{2v}\right)}{\sqrt{2\pi v}}, \tag{13.5}$$

so

$$LLR_{\text{ESE}}(x_1(j)) = 2h_1\sqrt{e_1}\frac{y(j) - \mu(j)}{v}. \tag{13.6}$$

For $K > 1$, the problem is much more complicated. We need to consider all possible combinations of $\{x_k(j)\}$. The exact result is the maximum likelihood (ML) estimator [36] below:

$$LLR_{\text{ESE}}(x_k(j)) = \log \frac{\sum_i \Pr\left(y(j)|x_k(j) = +1, X_{\sim k}^i(j)\right)\Pr\left(X_{\sim k}^i(j)\right)}{\sum_i \Pr\left(y(j)|x_k(j) = -1, X_{\sim k}^i(j)\right)\Pr\left(X_{\sim k}^i(j)\right)}, \tag{13.7}$$

where $X_{\sim k}^i(j)$ is one among all 2^{K-1} possibilities of the set $\{x_1(j), x_2(j), \ldots, x_{k-1}(j), x_{k+1}(j), \ldots, x_K(j)\}$ (since $x_{k'}(j) \in \{-1, +1\}, \forall k'$). In (13.7), $\Pr\left(y(j)|x_k(j) = \pm 1, X_{\sim k}^i(j)\right)$ can be computed similarly to (13.5) and $\Pr\left(X_{\sim k}^i(j)\right)$ can be computed from messages $\{LLR_{\text{DEC}}(x_k(j))\}$. Following [25], we define $LLR_{\text{DEC}}(x_k(j))$ by an LLR:

$$LLR_{\text{DEC}}(x_k(j)) = \log \frac{\Pr(x_k(j) = +1)}{\Pr(x_k(j) = -1)}. \tag{13.8}$$

We can obtain $\Pr(x_k(j) = \pm 1)$ by solving (13.8) together with $\Pr(x_k(j) = +1) + \Pr(x_k(j) = -1) = 1$. Then,

$$\Pr\left(X_{\sim k}^i(j)\right) = \prod_{k'=1, k'\neq k} \Pr(x_{k'}(j)). \tag{13.9}$$

The complexity of ML is $O(2^K)$ for BPSK, which increases exponentially with K. For a higher order modulation with an M-point constellation, the complexity of ML is $O(M^K)$. This can be a serious problem in practice.

Gaussian approximation (GA) is a low-cost alternative. We rewrite (13.3) as

$$y(j) = h_k\sqrt{e_k}x_k(j) + \zeta_k(j), \tag{13.10a}$$

where

$$\zeta_k(j) = y(j) - h_k\sqrt{e_k}x_k(j) = \sum_{k'=1,k'\neq k}^{K} h_{k'}\sqrt{e_{k'}}x_{k'}(j) + \eta(j) \tag{13.10b}$$

is the distortion (including interference-plus-noise) with respect to user k. From the central limit theorem, we apply GA to $\zeta_k(j)$ in (13.10b) and assume $\zeta_k(j) \sim N(\mu_k(j), \text{Var}(\zeta_k(j)))$. Now we can treat (13.10a) as a single-user system. For simplicity, we assume a real channel. (We will discuss a complex channel in Sect. 13.5.2.) Then, we have

$$\Pr(y(j)|x_k(j) = \pm 1) = \frac{\exp\left(-\frac{(y(j)-(\mu_k(j)\pm h_k\sqrt{e_k}))^2}{2\text{Var}(\zeta_k(j))}\right)}{\sqrt{2\pi\,\text{Var}(\zeta_k(j))}}. \tag{13.11}$$

Substituting (13.11) into (13.4) and evaluating $\mu_k(j)$ via (13.10b), we have the following ESE operations for the jth MA node in Fig. 13.6. (We will discuss the generation of $\text{Var}(\zeta_k(j))$ in (13.12c) later in (13.13).)

ESE operations

$$\text{(i) } \text{E}(x_k(j)) = \Pr(x_k(j) = +1) - \Pr(x_k(j) = -1), \tag{13.12a}$$

$$\text{(ii) } \mu_k(j) = \sum_{k'=1}^{K} h_{k'}\sqrt{e_{k'}}\text{E}(x_{k'}(j)) - h_k\sqrt{e_k}\text{E}(x_k(j)), \tag{13.12b}$$

$$\text{(iii) } LLR_{\text{ESE}}(x_k(j)) = 2h_k\sqrt{e_k}\frac{y(j) - \mu_k(j)}{\text{Var}(\zeta_k(j))}. \tag{13.12c}$$

The following are some details related to the ESE operations.

- Initially, there is no decoder feedback, and we can set $\text{E}(x_k(j)) = 0$ in (13.12a) for $\forall k, j$.
- GA is approximate. However, we observed a very good performance based on GA.
- The summation in (13.12b) can be shared by all users. The cost per information bit per user is independent of the number of users K.
- The principles of GA for higher-order modulations (such as quadrature phase-shift keying (QPSK)) and complex channels can be derived similarly. Related discussions will be shown in Sect. 13.5.2.

We now discuss the evaluation of $\mathrm{Var}(\zeta_k(j))$ involved in (13.12c). From (13.10b),

$$\mathrm{Var}(\zeta_k(j)) = \sum_{k'=1}^{K} |h_{k'}|^2 e_{k'} \mathrm{Var}(x_{k'}(j)) - |h_k|^2 e_k \mathrm{Var}(x_k(j)) + \sigma^2. \qquad (13.13a)$$

We observed that the system performance is not sensitive to $\mathrm{Var}(\zeta_k(j))$. Therefore, we take the following approximation

$$\mathrm{Var}(\zeta_k(j)) \approx \sum_{k'=1}^{K} |h_{k'}|^2 e_{k'} v_{k'} - |h_k|^2 e_k v_k + \sigma^2 \qquad (13.13b)$$

based on the following assumption

$$\mathrm{Var}(x_k(j)) = 1 - (\mathrm{E}(x_k(j)))^2 \approx v_k, \forall j. \qquad (13.13c)$$

To evaluate v_k in (13.13c), we can simply compute a few samples of $\mathrm{Var}(x_k(j))$ and take their average. In practice, computation for each $\mathrm{Var}(x_k(j))$ can be implemented using a look-up table. We observed that the required number of samples is small, so the related cost is negligible.

From the above discussions, the total cost for the ESE operations is (approximately) four additions and two multiplications per chip per iteration. (Some operations, such as $h_k \sqrt{e_k}$ and $2h_k \sqrt{e_k}/\mathrm{Var}(\zeta_k(j))$, can be precalculated and need not be repeated for every j. The related cost is negligible.)

13.2.3   Overall IDMA Receiver

Now return to an overall IDMA system in Fig. 13.5. We divide the receiver into an ESE module and K DEC modules. The operations in the ESE module is based on (13.12). Consider two types of schedules below.

- *Serial schedule*: In each iteration, operations are carried out as follows:
  ESE for user 1, DEC for user 1, ESE for user 2, DEC for user 2, . . .
  As an example, ESE for user 1 means running (13.12) for every j with k fixed to 1. The LLRs generated in (13.12) are then fed to DEC for user 1. Then, $\{\mu_1(j)\}$ and v_1 are updated based on the DEC outputs and the process continues to user 2.
- *Parallel schedule*: In each iteration, operations are carried out as follows:
  ESE for all users in parallel, then
  DEC for all users.
  In the above, in each iteration, (13.12) is run through every pair of j and k. After the ESE operations, the outputs are fed to the K local DECs. Then, the K DECs are run in parallel. Afterward, $\{\mu_k(j)\}$ and v_k are updated simultaneously for all k.

Fig. 13.7 Performance of IDMA with GA in AWGN channels
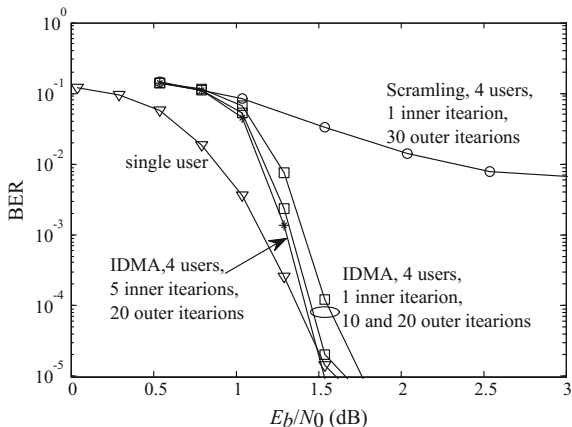
Note that in the parallel schedule, the value of the summation in (13.12b), that is generated using the results of the previous iteration, remains the same for all users in one iteration. In the serial schedule, this summation is updated user-by-user. We noticed that serial scheduling converges slightly faster than the parallel one.

A question arises whether it is helpful if multiple inner DEC iterations are carried out between two consecutive ESE iterations. For example, with the serial schedule, we can run multiple LDPC decoding iterations for one user before going to the next user. We observed that such inner iterations are generally unnecessary. For fixed overall cost, better performance is achieved without inner iterations. However, slipping some inner or outer iterations may lead to reduced complexity.

Intuitively, we can treat an IDMA system in Fig. 13.5 as a generalized code system on a graph (where an MA node is just for a special type of constraint). The inner-iteration method means more iterations on some parts of the graph for LDPC coding constraints. Such uneven message-passing process does not help in general.

For comparison of overall cost, let us consider a K-user TDMA system in which LDPC decoding is individually carried out for each user. The only difference between SUD for such a TDMA system and MUD for IDMA is the ESE operations. As we have seen above, the cost of the extra ESE operations is quite moderate. Thus, an IDMA receiver involving (13.12) has only moderately higher complexity than SUD for corresponding TDMA systems with the same number of users.

Figure 13.7 shows an example of a four-user IDMA system in AWGN channels. A rate-1/2 LTE turbo code with 1200 information bits per user is used, followed by a rate-1/8 repetition code and QPSK modulation. We can see that iterative detection nearly converges with about ten outer iterations between ESE and DECs with one inner-iteration. Here, one inner-iteration means running each component decoder once in a turbo decoder. We can also run multiple iterations in each turbo decoder within each outer iteration. The result of five inner iterations is shown in Fig. 13.7. We can see that multiple inner iterations can only offer marginal improvement on performance, even though at considerably higher overall cost.

Figure 13.7 also shows the result with user-specific scrambling by spreading each coded bit with a random binary sequence of $+1$ and -1 before modulation. No user-specific interleaving is used in this case. It is seen that interleaving offers better performance than scrambling. This is due to the short cycle problem as explained earlier.

13.2.4   Performance Evaluation Through SNR Evolution

We now outline an SNR evolution technique [17] for tracking IDMA performance. Similar techniques have been successfully applied to turbo and LDPC codes [25] and more recently to AMP algorithms [37]. This analysis method also provides the basis of power allocation for IDMA performance optimization in the next section.

Figure 13.8a is a so-called protograph [38] representation of Fig. 13.5, in which each circle represents a vector and each edge represents a vector connection. Relatively thick lines are used in Fig. 13.8a to distinguish it from Fig. 13.5. The messages $LLR_{\text{DEC},k}$ and $LLR_{\text{ESE},k}$ in Fig. 13.8a are LLR sequences generated by the coding and MA constraints, respectively.

We will use the following SNR-variance relationship to characterize the behavior of the system in Fig. 13.8a. Recall (13.10a): $y_k(j) = h_k \sqrt{e_k} x_k(j) + \zeta_k(j)$. We define the average SNR for user k as

$$SNR_k \equiv \frac{\text{E}\left(|h_k \sqrt{e_k} x_k(j)|^2\right)}{\text{E}(\text{Var}(\zeta_k(j)))} = \frac{|h_k|^2 e_k}{\text{E}(\text{Var}(\zeta_k(j)))}. \tag{13.14}$$

From (13.13), we have

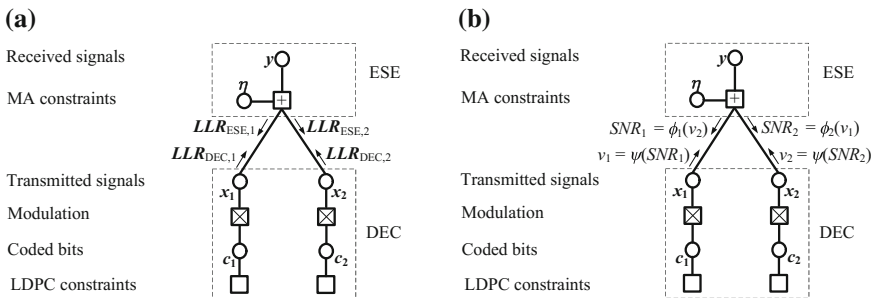$$\text{E}(\text{Var}(\zeta_k(j))) = \sum_{k'=1, k' \neq k}^{K} |h_{k'}|^2 e_{k'} v_{k'} + \sigma^2. \tag{13.15}$$

(a)   **(b)**

Fig. 13.8 **a** Protograph representation of Fig. 13.5. **b** Evolution characterization for (**a**)

Since $\boldsymbol{LLR}_{\text{DEC},k}$ is the output of DEC k with input $\boldsymbol{LLR}_{\text{ESE},k}$ characterized by SNR_k, we can write v_k as a function $v_k = \psi(SNR_k)$. Here for simplicity, we assume that all users employ the same LDPC code and therefore have the same $\psi(\cdot)$. (Note that the re-shuffling operation in Fig. 13.5b has no effect on $\psi(\cdot)$.) Combining this with (13.14) and (13.15), we have the following recursion to characterize an iterative IDMA detector:

$$SNR_k^{(t)} = \frac{|h_k|^2 e_k}{\sum_{k'=1,k'\neq k}^{K} |h_{k'}|^2 e_{k'} v_{k'}^{(t)} + \sigma^2} \equiv \phi_k\left(v_1^{(t)}, \ldots v_{k-1}^{(t)}, v_{k+1}^{(t)}, \ldots v_K^{(t)}\right),$$

(13.16a)

$$v_k^{(t)} = \psi\left(SNR_k^{(t-1)}\right),$$

(13.16b)

where t is an iteration index. The initialization is $v_k^{(0)} = 1$, $\forall k$, implying no information from DECs. In general, there is no closed form expression for $\psi(\cdot)$, but it can be obtained by simulating a single-user APP decoder in an AWGN channel with specified SNRs. Using $\{SNR_k^{(T)}\}$ in the final iteration in (13.16), we can estimate the BER by a function

$$BER_k = g\left(SNR_k^{(T)}\right),$$

(13.17)

where $g(\cdot)$ can be obtained through simulation of DECs [17].

13.2.5 Superposition Coded Modulation (SCM)

The above IDMA scheme involves multiple signal streams from different users. We may simply allocate these signal streams to a single-user. Such scheme is referred to as superposition coded modulation (SCM) [39, 40].

We define a standard QPSK constellation as $S_{\text{QPSK}} = \{00 \to (+1, +1), 01 \to (+1, -1), 10 \to (-1, +1), 11 \to (-1, -1)\}$. Figure 13.9a is a 16-ary scheme formed by superimposing S_{QPSK} and a scaled version of S_{QPSK} with a scaling factor of 2 and a 45° phase shift [39]. Figure 13.9b is a 64-ary scheme formed by superimposing S_{QPSK} and two scaled versions of S_{QPSK} with, respectively, scaling factors of 1.18 and 1.10 plus 60° and 120° phase shifts. From the central limit theorem, the SCM signaling is more Gaussian-like when the number of streams is large. This can offer the so-called shaping gain as analyzed in [41, 42]. It has been proved that, among all possible signaling methods, an SCM constellation achieves the minimum mean squared error (MMSE) bound; that is, it minimizes the function $\psi(\cdot)$ in (13.16b) for a fixed underlying binary decoder. The details are discussed in [39].
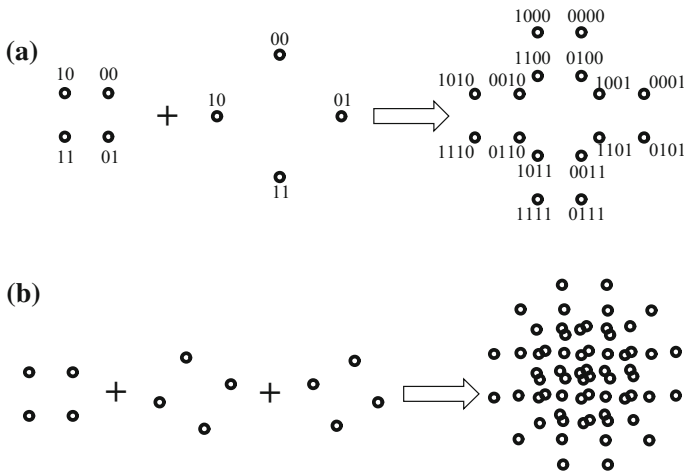
Fig. 13.9 a A 16-ary SCM signaling by superimposing two streams of QPSK constellations. **b** A 64-ary SCM signaling by superimposing three streams of QPSK constellations

13.3 Power Control for IDMA

At a relatively low sum-rate, such as less than 1 in a complex channel, IDMA with a GA receiver works well with equal received power. At a higher sum-rate, unequal power control is required. The situation is similar to power control for SIC in (13.2), except that iterative detection makes the problem more complicated.

13.3.1 Transmitted and Received Power Minimization

Denoting the received powers by

$$q_k = |h_k|^2 e_k. \tag{13.18}$$

We combine (13.16a) and (13.16b) into a compact form as

$$SNR_k^{(t)} = \frac{q_k}{\sum_{k'=1, k' \neq k}^{K} q_{k'} \psi \left(SNR_{k'}^{(t-1)} \right) + \sigma^2}, \forall k. \tag{13.19}$$

Let $\sum e_k$ and $\sum q_k$ be, respectively, sum transmitted power and sum received power. We can minimize them, respectively. The latter is simpler since the channel gains $\{|h_k|^2\}$ are not involved. The following Remark establishes a connection between these two problems [7, 9, 43].

Remark 1 Assume that $\{q_k^*\}$ is a minimizer for $\sum q_k$. Then, $\{e_k^* = q_k^*/|h_k|^2\}$ is a minimizer for $\sum e_k$ provided that $\{q_k^*\}$ and $\{|h_k|^2\}$ have the same order, i.e., $q_k^* \leq q_{k'}^*$ if $|h_k|^2 \leq |h_{k'}|^2$.

Based on Remark 1, we can first find the minimizer $\{q_k^*\}$ for $\sum q_k$. We then re-label $\{q_k^*\}$ such that it has the same order as $\{|h_k|^2\}$. Then, the minimizer for $\sum e_k$ can be obtained as $\{e_k^* = q_k^*/|h_k|^2\}$.

Incidentally, Remark 1 implies that a user with a higher channel gain should be assigned a higher transmitted power and vice versa. Next, we focus on minimizing $\{q_k\}$.

13.3.2   Feasible Profile

We now impose an SNR requirement Γ after T iterations. This SNR requirement can be equivalently translated into a BER requirement through (13.17). We write the received power optimization problem as follows.

$$\text{minimize} \quad \sum q_k, \tag{13.20a}$$

$$\text{subject to} \quad SNR_k^{(t)} = \frac{q_k}{\sum_{k'=1, k' \neq k}^{K} q_{k'} \psi \left(SNR_{k'}^{(t-1)} \right) + \sigma^2}, \forall k, \tag{13.20b}$$

$$SNR_k^{(T)} \geq \Gamma, \forall k. \tag{13.20c}$$

The problem in (13.20) is non-convex. We will outline two searching techniques for this problem. For convenience, we will call $\{q_k\}$ a feasible profile if it ensures the constraints in (13.20b) and (13.20c).

Incidentally, it is interesting to compare (13.2a) and (13.20b). In (13.20b), $q_{k'} \psi \left(SNR_{k'}^{(t-1)} \right)$ represents the residual interference from user k' after soft cancelation. Such terms disappear in (13.2a) for decoded users due to the error-free assumption and hard cancelation.

13.3.3   Greedy Search

We first set $T = 1$, i.e., only one iteration. Assume that approximate error-free decoding can be achieved at a sufficiently large SNR in the single-user case. We can construct an initial feasible profile $\mathcal{Q} = \{q_k\}$ according to (13.2). The sum-power for such a $\mathcal{Q}$ is typically large.

We next consider a general T. Starting from the above initial $\mathcal{Q}$, we search for a minimum value for each q_k individually to achieve (13.20c), while keeping other elements in $\mathcal{Q}$ unchanged. This involves a one-dimensional search, so its complexity

is affordable. Let the search result be q_k^*. We then update $q_k \leftarrow q_k - \epsilon(q_k - q_k^*)$ in $\mathscr{Q}$, where ϵ is a damping factor (e.g., $\epsilon = 0.5$). We repeat the above process for all k iteratively. We observed reasonably good performance of this simple method for a relatively small K.

13.3.4 Approximate Linear Programming Method

Inspired by the linear program technique for LDPC code design [25], we can use the approximate technique below for a large K. The key idea is to transform the problem of finding the power for each user into finding the number of users on different given power levels, which makes the problem convex [44, 45].

Let us quantize the received power into $M + 1$ discrete values: $\{q(m), m = 0, 1, \ldots, M\}$ with $q(m-1) < q(m)$. The received powers of all users are selected from $\{q(m)\}$. We partition K users into $M + 1$ groups according to their power levels. Let $\lambda(m)$ be the number of users assigned with power level $q(m)$ and $z(m)$ be the total power of these $\lambda(m)$ users. As such,

$$\sum_m \lambda(m) = K, \tag{13.21a}$$

$$z(m) = \lambda(m)q(m) \tag{13.21b}$$

and the sum received power

$$\sum_k q_k = \sum_m \lambda(m)q(m) = \sum_m z(m). \tag{13.21c}$$

Denote by $SNR(m)$ the SNR for the users in the mth group with power $q(m)$. Define

$$I = \sum_m z(m)\psi(SNR(m)) + \sigma^2, \tag{13.22}$$

which is the total interference power (including noise) after soft cancelation. When K is large, (13.20b) can be approximated as

$$SNR(m)^{(t)} = \frac{q(m)}{I^{(t)} - q(m)\psi\left(SNR(m)^{(t-1)}\right)} \approx \frac{q(m)}{I^{(t)}}, \tag{13.23}$$

where $I^{(t)}$ denotes the value of I at the tth iteration. Using (13.22) and (13.23), we have the update rule

$$I^{(t)} = \sum_m z(m)\psi\left(\frac{q(m)}{I^{(t-1)}}\right) + \sigma^2. \tag{13.24}$$

Fig. 13.10 IDMA with EPC
and UPC in AWGN
channels. Rate-1/3 turbo
coding followed by rate-1/2
repetition coding is used for
each user. Information length
of each user is 1200. QPSK
modulation. Sum-rate = 1, 2,
and 4 for $K = 3$, 6, and 12,
respectively
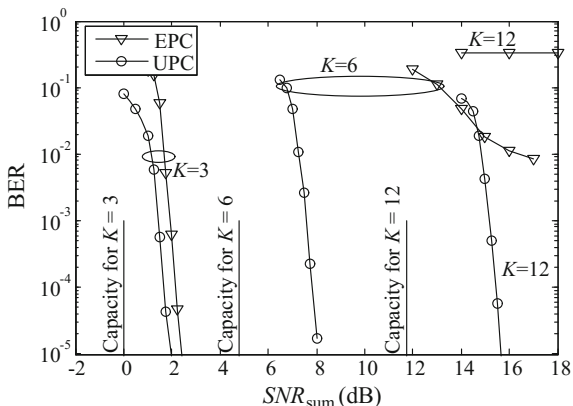
Equation (13.24) characterizes the evolution of the total interference at each iteration.
If iterative detection converges, $I^{(t)}$ should be lower than $I^{(t-1)}$. Equivalently, we can
write the convergence condition as

$$\sum_m z(m)\psi\left(\frac{q(m)}{I}\right) + \sigma^2 \leq (1-\delta)I, \quad I_{\min} \leq I \leq I_{\max}. \tag{13.25}$$

where $0 < \delta < 1$ is a decay factor that controls the convergence speed. $I_{\max}$ and $I_{\min}$
specify the total interference at the beginning and end of the iterative detection. In
summary, we re-formulate the optimization problem as,

$$\text{minimize} \quad \sum z(m), \tag{13.26a}$$

$$\text{subject to} \quad \sum_m z(m)\psi\left(\frac{q(m)}{I}\right) + \sigma^2 \leq (1-\delta)I, \quad I_{\min} \leq I \leq I_{\max}, \tag{13.26b}$$

$$z(m) \geq 0, \forall m. \tag{13.26c}$$

The above optimization problem is linear with respect to $\{z(m)\}$. Hence, it can be
resolved by linear programming. More details can be found in [44, 45].

Figure 13.10 is an example to illustrate the necessity of unequal power control
(UPC). We can see that UPC improves the system performance for all the three cases
of $K = 3$, 6, and 12 compared with equal power control (EPC). Particularly, when
$K = 12$, the IDMA system does not work at all with EPC, but works well with UPC.
The gap between IDMA with UPC and capacity is about 3.7 dB for $K = 12$.

Compared with the 12-user example in Fig. 13.3, we can see that IDMA recovers
about 3.7 dB relative to the loss incurred by SIC. This is impressive, but there is
definitely room for improvement. The gap towards capacity can potentially be further
narrowed using more sophisticated techniques, such as curve-matching-based degree

sequence optimization [46, 47], spatial coupling [48–57], or more sophisticated modulations [58].

A criterion called overloading, which represents the user capacity in a NOMA system, is used to assess the system performance in the recent literature. Figure 13.10 demonstrates that IDMA can offer very high overloading with centralized power control. In the next section, we will discuss IDMA techniques to achieve high throughput without centralized power control.

13.4   Random Access via IDMA

13.4.1   Limitations of Conventional Systems

A conventional uplink system with centralized control involves a connection setup procedure before data transmission. The overhead incurred by this procedure is not serious for services with long-lasting connections since it can be amortized across the connection duration. One of the main tasks envisaged for the next 5G cellular systems is to support machine-type communication (MTC) which is characterized by short and sporadic data communication. In this case, the cost of establishing centralized control can be substantial.

Random access is decentralized, in which each user makes an individual decision to transmit data packets. This avoids the overhead of connection setup, but the packets from different users may collide. In a conventional random-access scheme, such as ALOHA [59], colliding packets are discarded, which reduces throughput. For this reason, such techniques cannot satisfy the demand of high spectral efficiency in 5G cellular systems.

In a fading channel, the received powers of different users may form a feasible profile defined in Sect. 13.3.2, even without centralized control. This is captured in the multi-packet reception (MPR) model [60, 61]. Most existing MPR techniques rely on channel fading to form feasible profiles. Such a passive approach achieves limited throughput gain. In what follows, we will discuss an active approach based on the power control technique. The main idea is to optimize the probability of forming a feasible profile through decentralized power control at the transmitters.

13.4.2   Random IDMA with Decentralized Power Control

13.4.2.1   Problem Formulation

Recall from Sect. 13.3.2, a feasible received power profile $\{q_k\}$ can be formed by centralized control. Without centralized control, however, it is difficult to guarantee

this. A randomized power control (RPC) technique [62–67] is discussed below to handle this difficulty.

The principles of RPC are as follows. Let $\{Q^{(l)}, l = 1, 2, \ldots, L\}$ (L is the maximum level index) be a set of pre-defined power levels and $\{P^{(l)}, l = 1, 2, \ldots, L\}$ a set of related probabilities. Upon a packet arrival, each user randomly draws a power level $Q^{(l)}$ with probability $P^{(l)}$ and uses it to transmit. Different users act individually and so their transmissions may collide. However, as long as their received powers $\{q_k\}$ form a feasible profile, their signals can still be recovered. Our aim is to optimize the probability that $\{q_k\}$ form a feasible profile.

Mathematically, $\{q_k\}$ defined above are the realizations of an underlying random variable. The random variable is characterized by a probability mass distribution $\{P^{(l)}\}$. Each q_k is independently drawn by a user from the support $\{Q^{(l)}\}$. Once the distribution is given, there is no need for centralized control.

13.4.2.2 Type-2 Collisions

The problem formulated above turns out to be difficult when K is large. So far, we have no general solution. We will discuss a sub-optimal technique below.

We say that a collision is of type-M if it involves M active users. Let $Q^{(0)} = 0$ be an element in $\{Q^{(l)}\}$. In RPC, a user will not transmit if its selected power is $Q^{(0)}$. The related $P^{(0)}$ is equivalent to the back-off probability in 802.11 Wi-fi systems. Intuitively, collisions are dominated by type-2 ones when $P^{(0)}$ is sufficiently large. Therefore, we will focus on type-2 collisions.

We consider a type-2 collision involving user i and user j. We define the union of all possible feasible profiles of the received power pair $\{q_i, q_j\}$ as a feasible region. The collision is resolvable if $\{q_i, q_j\}$ falls in this region. Figure 13.11a shows an example of the feasible region for SIC with ideal coding and decoding. The area marked by "A" in Fig. 13.11a is formed by all possible $\{q_i, q_j\}$ that meets the following conditions (see (13.2)):

$$SNR_j = \frac{q_j}{q_i + \sigma^2} \geq \Gamma, \tag{13.27a}$$

$$SNR_i = \frac{q_i}{\sigma^2} \geq \Gamma. \tag{13.27b}$$

The area marked by "B" in Fig. 13.11a is formed similarly by changing decoding order. The value of Γ here is determined by the Shannon capacity $R = \log_2(1 + \Gamma)$. Any power pair in the feasible region is resolvable by SIC.[2]

Figure 13.11b is an example of an IDMA system involving two LDPC-coded users with coding rate 0.5 per user. The receiver can achieve BER $\leq 10^{-5}$ in the feasible region, which is regarded as approximately error-free. The border of this feasible region is obtained using simulation.

---

[2]Figure 13.11a is for $R < 1$. If $R \geq 1$, the feasible region is divided into two disjoint sub-regions A and B symmetric to the 45° line $q_i = q_j$ [63].
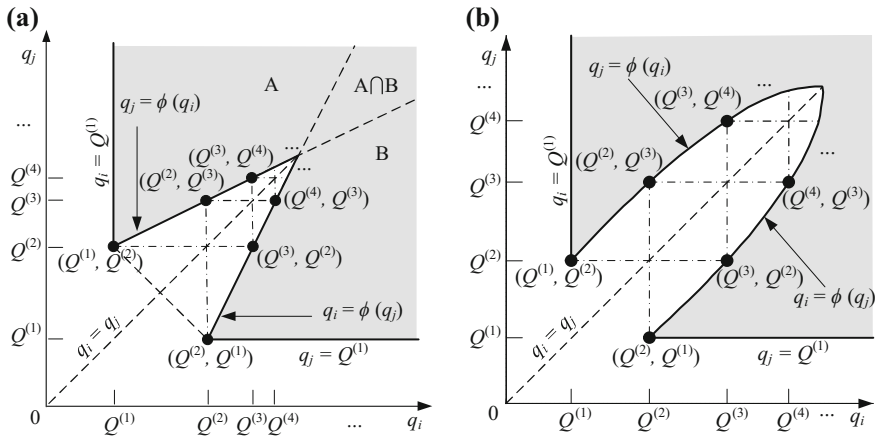
Fig. 13.11 Feasible regions for **a** a two-user ideally coded system with SIC, and **b** a two-user LDPC-coded IDMA system

Each of the two feasible regions in Fig. 13.11 is bounded by four curves $q_i = Q^{(1)}$, $q_j = \phi(q_i)$, $q_j = Q^{(1)}$ and $q_i = \phi(q_j)$. Here, the function $\phi(\cdot)$ is determined by (13.27a) taking mark of equality (for Fig. 13.11a) or by simulation (for Fig. 13.11b). $Q^{(1)}$ is the minimum power for successful single-user transmission. We construct the set $\{Q^{(l)}\}$ as follows:

$$Q^{(l)} = \begin{cases} 0, & l = 0, \\ Q^{(1)}, & l = 1, \\ \phi(Q^{(l-1)}), & l > 1. \end{cases} \tag{13.28}$$

For $\{q_k\}$ randomly selected from $\{Q^{(l)}\}$, we have the following situations:

Case 1: All $\{q_k\}$ are zeros. In this case, throughput is zero.

Case 2: Only one element in $\{q_k\}$ is nonzero. The transmission of the only active user will be successful.

Case 3: Exactly two elements in $\{q_k\}$, say q_i and q_j, are nonzero. This is a type-2 collision. It can be shown that $\{q_i, q_j\}$ falls in the feasible region (so collision is resolvable) provided that $q_i \neq q_j$.

Case 4: More than two elements in $\{q_k\}$ are nonzero. For simplicity, such events are regarded as unresolvable, which is a pessimistic assumption.

Based on the above cases, we can find an optimized probability set $\{P^{(l)}\}$ for a K user system. For convenience, we assume that the packets of all users arrive independently, following the Bernoulli process with parameter λ. The system throughput is then given by

$$T = T_1 + T_2. \tag{13.29}$$

In (13.29), T_1 is the throughput related to transmissions without collision:

$$T_1 = \sum_{k=1}^{K} C_K^k \lambda^k (1-\lambda)^{(K-k)} C_k^1 (1-P^{(0)})(P^{(0)})^{k-1}$$

$$= K\lambda(1-P^{(0)})(1-\lambda+\lambda P^{(0)})^{K-1}, \tag{13.30}$$

where $C_K^k \lambda^k (1-\lambda)^{(K-k)}$ is the probability of k users among total K users having packets to transmit and $C_k^1(1-P^{(0)})(P^{(0)})^{k-1}$ the probability of only one user among these k users transmitting with nonzero power. Also in (13.29), T_2 is the throughout related to type-2 collisions. From case 3 above, a type-2 collision is unresolvable only if the two transmitting users are using the same received powers. Therefore, we have

$$T_2 = 2 \sum_{k=2}^{K} C_K^k \lambda^k (1-\lambda)^{(K-k)} C_k^2 \left((1-P^{(0)})^2 - \sum_{l>0} (P^{(l)})^2 \right) (P^{(0)})^{k-2}$$

$$= K(K-1)\lambda^2 (1-\lambda+\lambda P^{(0)})^{K-2} \left((1-P^{(0)})^2 - \sum_{l>0} (P^{(l)})^2 \right), \tag{13.31}$$

where $(1-P^{(0)})^2 - \sum_{l>0} (P^{(l)})^2$ is the probability that two active users transmit with different received powers.

We further consider an average transmitted power constraint $\bar{q}$ for each user. In AWGN channels with unit channel power gain, the constraint is given by

$$\sum_{l \geq 0} Q^{(l)} P^{(l)} \leq \bar{q}. \tag{13.32}$$

Under such power constraint, we can search for $\{P^{(l)}\}$ that maximize the throughput T in (13.29). It can be verified that the problem is convex if $P^{(0)}$ is fixed. The treatments for fading channels are somewhat more complicated. The details can be found in [63].

Figure 13.12 shows two examples, one for an ideally coded system and the other for an LDPC-coded IDMA system [63]. Conventional ALOHA is included as a reference. We can see that the RPC-based scheme can offer noticeably throughput gain compared with ALOHA.

It is proved in [63] that the $\{Q^{(l)}\}$ in (13.28) forms an optimal support for the decentralized power control when $K = 2$. It is sub-optimal for $K > 2$, but it can still provide excellent performance gain, as seen in Fig. 13.12.

As a short summary, we can treat collisions as NOMA cases. A conventional scheme, such as ALOHA, treats such NOMA cases as failures. The discussions in this section aim at optimizing the probability of successful detection in such NOMA cases.
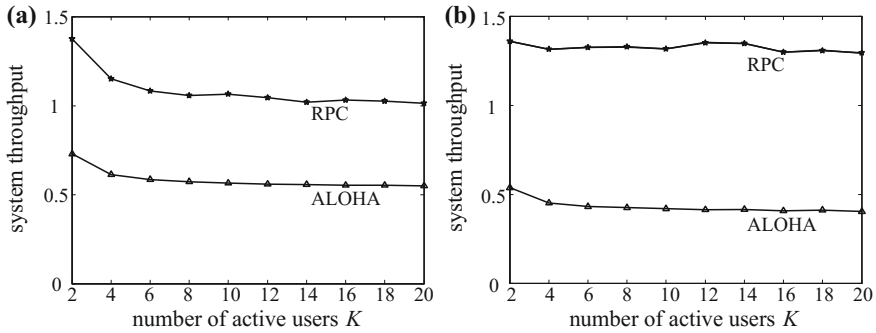
Fig. 13.12 Performance comparison of RPC and ALOHA in **a** an ideally coded system and **b** an LDPC-coded IDMA system. Rayleigh fading channel with averaged power gain 1 for all users. Same power constraint for RPC and ALOHA

13.5 IDMA in MIMO Systems

13.5.1 Multi-User Gain in MIMO Systems

MIMO is a wireless technology employing multiple transmit and receive antennas [68–74]. Multi-user gain refers to the advantage of allowing a large number of users to transmit simultaneously over the same time and same frequency in MIMO [10, 75]. This is illustrated in Fig. 13.13 by the potential sum-rate capacity gain for a single-cell system. Perfect CSI is assumed in Fig. 13.13. The curves apply to both up- and downlinks following the duality principle [3, 76, 77]. We can see that multi-user gain is very attractive. The potential gain is in the order of tens of times. Diversifying power over more users, i.e., increasing K, is a very effective way to increase sum-rate.

Fig. 13.13 Achievable sum-rate of ZF under perfect CSI. The number of antennas at the base station is 64. Single antenna assumed is for each user. Other system parameters are the same as those in Fig. 13.1
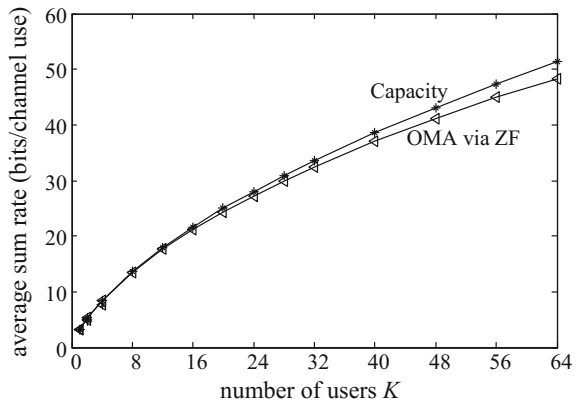
Figure 13.13 also includes the performance of zero-forcing (ZF) with proper power allocation. ZF is an OMA technique [75]. Different users are divided into different orthogonal subspaces in ZF, which avoids interference among users. It is seen from Fig. 13.13 that, with accurate CSI, ZF can offer very good multi-user gain. Since the gap between ZF and capacity is small, any further gain by NOMA is limited. In this case, OMA via ZF can be preferred for its low-cost SUD receiver.

However, in practice, we usually do not have reliable CSI to establish spatial orthogonality initially. ZF performance deteriorates seriously when CSI is not accurate. In the following, we will see that NOMA via IDMA offers a solution to the problem.

13.5.2 Iterative Maximum Ratio Combining (I-MRC)

We first extend the GA-based detection technique in Sect. 13.2.2 to MIMO. Consider a multi-user uplink system model with N_{BS} antennas at the base station. For simplicity, we assume a single antenna at each user. The IDMA principle discussed in Sect. 13.2.1 can be directly used here. We assume perfect CSI first and will return to the CSI estimation problem later.

The received signal at time j is written as

$$\boldsymbol{y}(j) = \sum_{k=1}^{K} \boldsymbol{h}_k \sqrt{e_k} x_k(j) + \boldsymbol{\eta}(j), \tag{13.33}$$

where $\boldsymbol{y}(j)$ is an $N_{\mathrm{BS}} \times 1$ signal vector received at base station antennas, h_k an $N_{\mathrm{BS}} \times 1$ complex channel coefficient vector, e_k the transmitted power of user k, $x_k(j)$ a symbol transmitted from user k, and $\boldsymbol{\eta}(j)$ an $N_{\mathrm{BS}} \times 1$ vector of complex AWGN with mean 0 and variance $\sigma^2 = N_0/2$ per dimension.

Maximum ratio combining (MRC) is a common strategy for MIMO systems. An MRC estimator is defined in a symbol-by-symbol form as

$$\hat{x}_k(j) = \boldsymbol{h}_k^H \boldsymbol{y}(j). \tag{13.34}$$

Substituting (13.33) into (13.34),

$$\hat{x}_k(j) = \lambda_k x_k(j) + \xi_k(j), \tag{13.35a}$$

where $\lambda_k \equiv \|\boldsymbol{h}_k\|^2 \sqrt{e_k} = \boldsymbol{h}_k^H \boldsymbol{h}_k \sqrt{e_k}$ is a scalar and

$$\xi_k(j) \equiv \sum_{k'=1,k'\neq k}^{K} \boldsymbol{h}_k^H \boldsymbol{h}_{k'} \sqrt{e_{k'}} x_{k'}(j) + \boldsymbol{h}_k^H \boldsymbol{\eta}(j) \tag{13.35b}$$

is an interference (from $\{x_{k'}(j), k' \neq k\}$ to $x_k(j)$) plus noise term. MRC does not involve matrix inversion and so has low cost. However, interference is a problem for MRC, especially when K is large. Iterative GA technique can alleviate this problem. Similar as that in Sect. 13.2.2, we approximate $\xi_k(j)$ in (13.35b) by a Gaussian random variable. We assume that the real and imaginary parts of $x_k(j)$ carry two bits of information in the QPSK modulation. Similar to (13.12), the real part of $x_k(j)$ can be estimated as

$$LLR_{\text{ESE}}\left(\text{Re}(x_k(j))\right) = \frac{2\lambda_k}{\text{Var}\left(\text{Re}\left(\xi_k(j)\right)\right)}\text{Re}\left(\hat{x}_k - \text{E}\left(\xi_k(j)\right)\right). \qquad (13.36)$$

The mean and variance in (13.36) are updated as

$$\text{E}\left(\xi_k(j)\right) = \boldsymbol{h}_k^H\left(\sum_{k'=1}^{K}\boldsymbol{h}_{k'}\sqrt{e_{k'}}\text{E}(x_{k'}(j)) - \boldsymbol{h}_k\sqrt{e_k}\text{E}(x_k(j))\right), \qquad (13.37a)$$

$$\text{Var}\left(\text{R}e\left(\xi_k(j)\right)\right)$$
$$= \sum_{k'=1}^{K}\text{Var}\left(\text{Re}\left(\boldsymbol{h}_k^H\boldsymbol{h}_{k'}\sqrt{e_{k'}}x_{k'}(j)\right)\right) - \text{V}ar\left(\text{Re}\left(\|\boldsymbol{h}_k\|^2\sqrt{e_k}x_k(j)\right)\right) + \|\boldsymbol{h}_k\|^2\sigma^2.$$
$$(13.37b)$$

Some detailed computation techniques for (13.37) can be found in [75]. The imaginary part of $x_k(j)$ can be estimated similarly. We call the above process iterative MRC (I-MRC).

Figure 13.14 illustrates the effectiveness of I-MRC [75]. We consider three different settings:

 (i) $K = 1$ and sum-rate $R_{\text{sum}} = 5$ with five signal streams (each stream with rate 1) assigned to the sole user using the SCM principle discussed in Sect. 13.2.5,
 (ii) $K = 8$ and $R_{\text{sum}} = 16$ with two streams per user, and
(iii) $K = 8$ and $R_{\text{sum}} = 24$ with three streams per user.

For $K = 1$, all the signal streams see the same channel so there is no spatial diversity among them, which results in poor performance. Increasing K from 1 to 8 results in drastically enhanced rate or reduced power or both in Fig. 13.14. Figure 13.14 is a compelling evidence for multi-user gain: allowing more concurrent transmitting users is more efficient than increasing single-user rate.

13.5.3 Data-Aided Channel Estimation (DACE)

We now consider the CSI acquisition problem. Many factors may affect CSI accuracy in MIMO. In particular, the correlation among the pilots used by different users can lead to the pilot contamination problem [78].
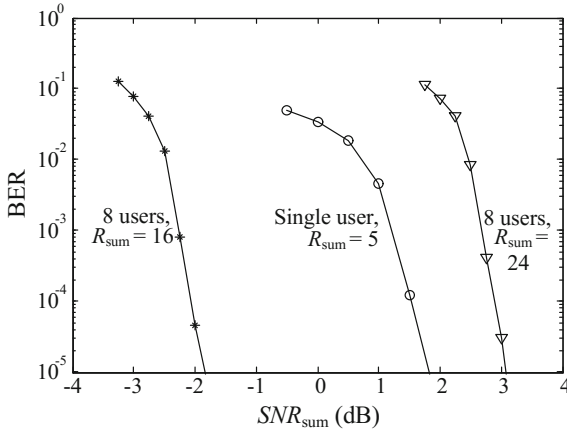
Fig. 13.14 Multi-user gain for $K = 8$ with I-MRC. Rayleigh fading. $N_{BS} = 64$. Equal transmitted power is assumed for different users. Power control is used for the streams assigned to the same user. The power allocation levels are obtained through heuristic search. Rate-1/2 turbo coding and information length $= 1200$ for each stream. QPSK modulation. A codeword is transmitted over ten resource blocks. Each resource block contains 180 symbols experiencing the same fading conditions

IDMA with data-aided channel estimation (DACE) [17, 79–83] technique can be used to improve CSI accuracy. The basic principle of DACE is as follows. Recall that a key difference between pilot and data is that the former is known at the receiver, while the latter is not. Therefore, if a data symbol is known, it can be used as a pilot. Furthermore, partial information of a data symbol, such as its mathematical mean, can also be used to refine the channel estimates. Such partial information is readily available in an IDMA receiver (as given in (13.8)).

DACE can be used jointly with I-MRC, which involves iterations of the following two operations [75]:

(a) using both pilots and partially decoded data information to refine CSI, and
(b) using improved channel estimates to refine data estimation by I-MRC.

The advantages of DACE are twofolds: (i) With DACE, the estimated data is gradually used for channel estimation. Pilot energy can be greatly reduced since only very coarse CSI is required initially. (ii) Data sequences are typically much longer than pilots and correlation is low among them. Therefore, DACE is robust against the pilot contamination problem. Such problem is typically caused by the correlation among the pilot sequences re-used in neighboring cells. Without DACE, longer pilot sequences will be required to reduce such correlation. Thus, DACE also reduces the time overhead related to pilots.

I-MRC and DACE can be naturally combined in an overall iterative process. After MRC and decoding operations in each iteration, partially detected data are used to refine channel estimates that are in turn used for MRC and decoding in the next iteration. This is referred to as I-MRC-DACE. Figure 13.15 compares BER
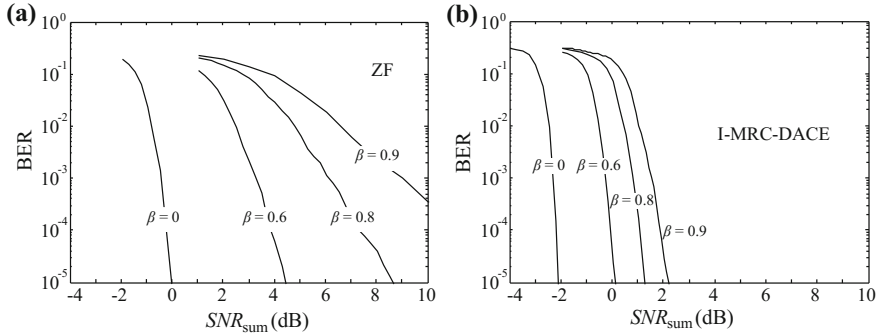
Fig. 13.15 Performance comparison of **a** ZF and **b** I-MRC-DACE with different β values. Rayleigh fading. $N_{BS} = 64$ and $K = 16$. Rate-1/3 turbo coding and information length $= 1312$ for each user. QPSK modulation. Each codeword is divided into 12 sections, and each section is transmitted over a resource block (including 16 pilot symbols). Different users in a cell are assigned different orthogonal pilots. These pilots are repeated for users in different cells. The pilot and data symbols have the same average power

performance for ZF and I-MRC-DACE, in which β is the cross-cell interference factor defined in (13.1). A larger β indicates a more serious pilot contamination problem due to more severe interference among the pilots. From Fig. 13.15, we can see that I-MRC-DACE noticeably outperforms ZF. The difference becomes very significant when β is large (e.g., $\beta \geq 0.6$).

IDMA is a natural choice for I-MRC-DACE since it is beneficial for iterative detection. Note that Fig. 13.5 can also be used to characterize IDMA in MIMO systems, if each scalar $y(j)$ in Fig. 13.5 is replaced by its vector counterpart $\mathbf{y}(j)$ in (13.33). The discussions on short cycles in Sect. 13.2.1 are still applicable after such replacement.

IDMA also allows a superimposed pilot scheme that can reduce the power overhead and rate loss. The related discussions can be found in [83–86].

13.6 Prospective Applications of IDMA in 5G Systems

Various approaches have been proposed recently for 5G radio link under LTE, including IDMA [87], RSMA [88], IGMA [89], PDMA [90] and SCMA [91, 92]. In the following, we will show that these schemes all share, explicitly or implicitly, the basic principle of IDMA.

We first represent these different schemes using a unified protograph framework. Assume that N resource blocks (RBs) defined in LTE are available for transmission. We label the observations from these RBs by $\{\mathbf{y}^{(1)}, \mathbf{y}^{(2)}, \ldots, \mathbf{y}^{(N)}\}$.

Figure 13.16 shows a scheme in which each user transmits on all available RBs as illustrated for two system settings: (a) three users over two RBs, and (b) six users
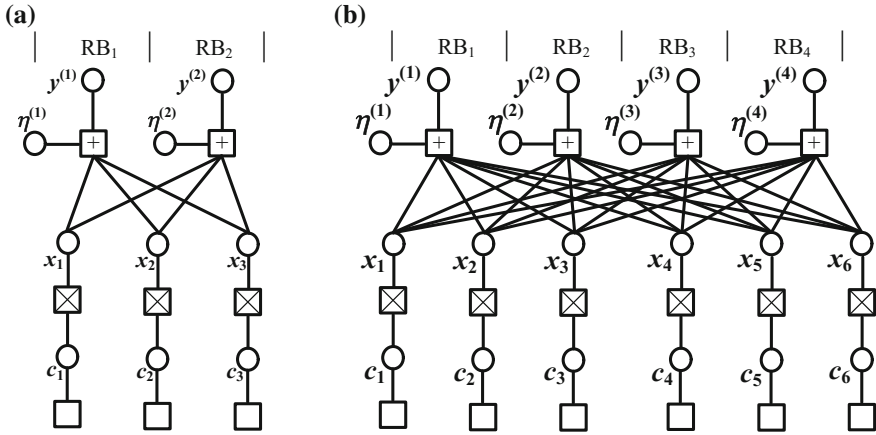
Fig. 13.16 Protograph representations of DS-CDMA and IDMA with **a** three users over two RBs, and **b** six users over four RBs

over four RBs. This can be realized by transmitting replicas of each x_k over multiple RBs. Alternatively, we may use a low-rate code to generate each x_k. Each x_k can be segmented into several blocks, with each block transmitted over an RB. The latter approach can potentially provide higher coding gain [93]. We may also use different modulations for the bits on different RBs, as for SCMA [92].

Incidentally, both DS-CDMA and IDMA can be represented using the protographs in Fig. 13.16. They are distinguished by the absence or presence of user-specific interleaving within each RB. RSMA [88] is a DS-CDMA scheme. However, user-specific interleaving is stated as an option for RSMA. If this option is used, it is equivalent to IDMA. The advantage of this option can be seen in Fig. 13.7.

Alternatively, each user can transmit over only some of the available RBs. This is referred to as sparse coding in [91]. Figure 13.17 shows an example for sparse coding. IGMA, PDMA, and SCMA all involve such treatment. Note that symbol-level interleaving as in Fig. 13.5a is not explicitly seen in Figs. 13.16 and 13.17. If such underlying interleaving is not used, size-4 cycles can be a problem in Fig. 13.16. Sparse coding in Fig. 13.17 avoids this problem. Clearly, sparse coding leads to user-specific edge connections between users and RBs. It has the same effect as symbol-level interleaving; they both reduce short cycles.

With sparse coding, each user does not fully occupy all RBs. This may cause problem for decentralized grant-free [13] or random-access applications, where each user determines its activity individually. In these cases, the number of active users, denoted by K_{active}, is a random variable. When K_{active} is small, sparse coding may lead to inefficient use of the available RBs and so low power efficiency. This implies poor scalability of user numbers. On the other hand, an IDMA system in Fig. 13.16 based on symbol-level interleaving does not have this problem, since all available RBs are fully used for any value of K_{active}.

Fig. 13.17 Protograph
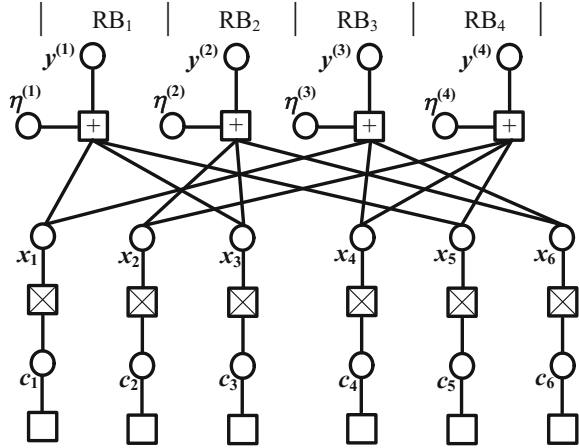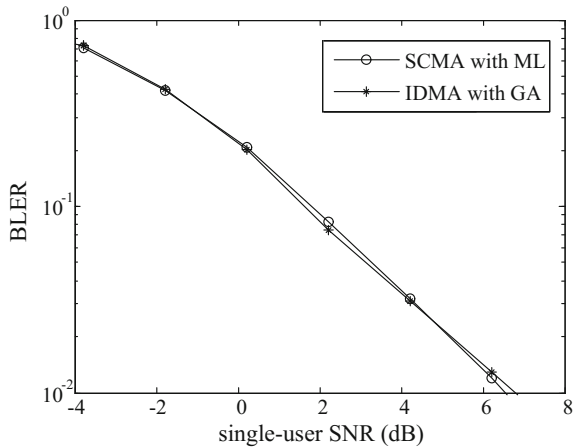representation of sparse
coding

Fig. 13.18 IDMA with GA
versus SCMA with ML. Ten
iterations for both schemes

Also, multi-user gain in MIMO is determined by the number of users concurrently
transmitting in each RB. Therefore, sparse coding may not be an efficient option in
MIMO (especially in massive MIMO).

Figure 13.18 compares IDMA and SCMA in quasi-static Rayleigh fading chan-
nels. The channels remain unchanged within each transmission. Both schemes are
with six users, two receiver antennas, sum-rate = 3 and equal transmitted power per
user. A rate-1/2 LTE turbo code is used followed the following transmitter structures:

- IDMA is with a length-2 spreading and QPSK modulation.
- SCMA is based on Fig. 13.17 with the 16-point modulation in [92].

We can see from Fig. 13.18 that the two schemes have similar performance. However,
SCMA in Fig. 13.18 is based on ML, while IDMA based on GA. The latter has much
lower complexity.

13.7   Summary

We have shown that the real attractiveness of NOMA is in systems without centralized control or without accurate CSI. It is difficult or too costly to establish orthogonality in such channels, so we have to resort to NOMA. Iterative processing holds the key; interference can be gradually resolved and CSI can be gradually refined through iterative processing. IDMA is a simple implementation technique for NOMA. The features of IMDA can be seen from its sparse graphic representation. The interleaved edge connections in IDMA facilitate iterative processing at the receiver. We have demonstrated that IDMA can offer significant performance gain in random access and MIMO systems. IDMA also offers lower detection complexity compare with other alternatives.

References

1. R. Ahlswede, Multi-way communication channels, in *Second International Symposium on Information Theory* (1971), pp. 103–135
2. H. Liao, A coding theorem for multiple access communications, in *Proceedings of the International Symposium on Information Theory*, Asilomar (1972)
3. D. Tse, P. Viswanath, *Fundamentals of Wireless Communication* (Cambridge University Press, 2005)
4. P. Viswanath, D.N.C. Tse, V. Anantharam, Asymptotically optimal water-filling in vector multiple-access channels. IEEE Trans. Inf. Theory **47**(1), 241–267 (2001)
5. W. Yu, W. Rhee, S. Boyd, J.M. Cioffi, Iterative water-filling for Gaussian vector multiple-access channels. IEEE Trans. Inf. Theory **50**(1), 145–152 (2004)
6. M. Kobayashi, G. Caire, An iterative water-filling algorithm for maximum weighted sum-rate of Gaussian MIMO-BC. IEEE J. Sel. Areas Commun. **24**(8), 1640–1646 (2006)
7. R.G. Gallager, An inequality on the capacity region of multiaccess multipath channels, in *Communications and Cryptography: Two Sides of One Tapestry* (Kluwer, Boston, 1994), pp. 129–139
8. R.R. Muller, A. Lampe, J.B. Huber, Gaussian multiple-access channels with weighted energy constraint, in *IEEE Information Theory Workshop*, June 1998, pp. 106–107
9. P. Wang, J. Xiao, L. Ping, Comparison of orthogonal and non-orthogonal approaches to future wireless cellular systems. IEEE Veh. Technol. Mag. **1**(3), 4–11 (2006)
10. P. Wang, L. Ping, On maximum eigenmode beamforming and multi-user gain. IEEE Trans. Inf. Theory **57**(7), 4170–4186 (2011)
11. N. Otao, Y. Kishiyama, K. Higuchi, Performance of non-orthogonal access with SIC in cellular downlink using proportional fair-based resource allocation, in *IEEE ISWCS 2012*, August 2012, pp. 476–480
12. Z. Ding, Y. Liu, J. Choi, Q. Sun, M. Elkashlan, I. Chih-Lin, H.V. Poor, Application of non-orthogonal multiple access in LTE and 5G networks. IEEE Commun. Mag. **55**(2), 185–191 (2017)
13. L. Dai, B. Wang, Y. Yuan, S. Han, I. Chih-Lin, Z. Wang, Non-orthogonal multiple access for 5G: solutions, challenges, opportunities, and future research trends. IEEE Commun. Mag. **53**(9), 74–81 (2015)
14. G. Wunder, P. Jung, M. Kasparick, T. Wild, F. Schaich, Y. Chen, L. Mendes, 5GNOW: non-orthogonal, asynchronous waveforms for future mobile applications. IEEE Commun. Mag. **52**(2), 97–105 (2014)

15. F. Schaich, B. Sayrac, S. Elayoubi, I.P. Belikaidis, M. Caretti, A. Georgakopoulos, B. Mouhouche, FANTASTIC5G: flexible air interface for scalable service delivery within wireless communication networks of the 5th generation. Trans. Emerg. Telecommun. Technol. **27**(9), 1216–1224 (2016)

16. W. Shin, M. Vaezi, B. Lee, D.J. Love, J. Lee, H.V. Poor, Non-orthogonal multiple access in multi-cell networks: theory, performance, and practical challenges. IEEE Commun. Mag. **55**(10), 176183 (2017)

17. L. Ping, L. Liu, K. Wu, W.K. Leung, Interleave division multiple-access. IEEE Trans. Wirel. Commun. **5**(4), 938–947 (2006)

18. P.A. Hoeher, H. Schoeneich, J.C. Fricke, Multi-layer interleave-division multiple access: theory and practice. Trans. Emerg. Telecommun. Technol. **19**(5), 523–536 (2008)

19. T. Yang, J. Yuan, Z. Shi, Rate optimization for IDMA systems with iterative joint multi-user decoding. IEEE Trans. Wirel. Commun. **8**(3), 1148–1153 (2009)

20. K. Kusume, G. Bauch, W. Utschick, IDMA vs. CDMA: analysis and comparison of two multiple access schemes. IEEE Trans. Wirel. Commun. **11**(1), 78–87 (2012)

21. K. Wu, K. Anwar, T. Matsumoto, BICM-ID-based IDMA: convergence and rate region analyses. IEICE Trans. Commun. **97**(7), 1483–1492 (2014)

22. G. Song, J. Cheng, Distance enumerator analysis for interleave-division multi-user codes. IEEE Trans. Inf. Theory **62**(7), 4039–4053 (2016)

23. C. Berrou, A. Glavieux, P. Thitimajshima, Near Shannon limit error-correcting coding and decoding: turbo-codes. 1, in *IEEE ICC'93*, vol. 2, May 1993, pp. 1064–1070

24. S. Benedetto, G. Montorsi, Unveiling turbo codes: some results on parallel concatenated coding schemes. IEEE Trans. Inf. Theory **42**(2), 409–428 (1996)

25. T. Richardson, R. Urbanke, *Modern Coding Theory* (Cambridge University Press, 2008)

26. C. Douillard, M. Jzquel, C. Berrou, D. Electronique, A. Picart, P. Didier, A. Glavieux, Iterative correction of intersymbol interference: turbo-equalization. Trans. Emerg. Telecommun. Technol. **6**(5), 507–511 (1995)

27. X. Wang, H.V. Poor, Iterative (turbo) soft interference cancellation and decoding for coded CDMA. IEEE Trans. Commun. **47**(7), 1046–1061 (1999)

28. M.C. Reed, C.B. Schlegel, P.D. Alexander, J.A. Asenstorfer, Iterative multiuser detection for CDMA with FEC: near-single-user performance. IEEE Trans. Commun. **46**(12), 1693–1699 (1998)

29. F.N. Brannstrom, T.M. Aulin, L.K. Rasmussen, Iterative multi-user detection of trellis code multiple access using a posteriori probabilities, in *IEEE ICC 2001*, vol. 1, June 2001, pp. 11–15

30. M. Moher, An iterative multiuser decoder for near-capacity communications. IEEE Trans. Commun. **46**(7), 870–880 (1998)

31. L. Ping, L. Liu, W.K. Leung, A simple approach to near-optimal multiuser detection: interleave-division multiple-access, in *IEEE WCNC 2003*, vol. 1, March 2003, pp. 391–396

32. F.R. Kschischang, B.J. Frey, H.A. Loeliger, Factor graphs and the sum-product algorithm. IEEE Trans. Inf. Theory **47**(2), 498–519 (2001)

33. Y. Mao, A.H. Banihashemi, A heuristic search for good low-density parity-check codes at short block lengths, in *IEEE ICC 2001*, vol. 1, June 2001, pp. 41–44

34. R. Zhang, L. Xu, S. Chen, L. Hanzo, Repeat accumulate code division multiple access and its hybrid detection, in *IEEE ICC'08*, May 2008, pp. 4790–4794

35. Z. Chenghai, H. Jianhao, The shifting interleaver design based on PN sequence for IDMA systems, in *IEEE FGCN 2007*, vol. 2, December 2007, pp. 279–284

36. M. Noemm, T. Wo, P.A. Hoeher, Multilayer APP detection for IDM. Electron. Lett. **46**(1), 96–97 (2010)

37. D.L. Donoho, A. Maleki, A. Montanari, Message-passing algorithms for compressed sensing. Proc. Natl. Acad. Sci. **106**(45), 18914–18919 (2009)

38. J. Thorpe, Low-density parity-check (LDPC) codes constructed from protographs. IPN Prog. Rep. **42**(154), 42–154 (2003)

39. L. Ping, J. Tong, X. Yuan, Q. Guo, Superposition coded modulation and iterative linear MMSE detection. IEEE J. Sel. Areas Commun. **27**(6) (2009)

40. P.A. Hoeher, T. Wo, Superposition modulation: myths and facts. IEEE Commun. Mag. **49**(12) (2011)
41. R. Laroia, N. Farvardin, S.A. Tretter, On optimal shaping of multidimensional constellations. IEEE Trans. Inf. Theory **40**(4), 1044–1056 (1994)
42. X. Ma, L. Ping, Coded modulation using superimposed binary codes. IEEE Trans. Inf. Theory **50**(12), 3331–3343 (2004)
43. C.H.F. Fung, W. Yu, T.J. Lim, Precoding for the multiantenna downlink: multiuser SNR gap and optimal user ordering. IEEE Trans. Commun. **55**(1), 188–197 (2007)
44. L. Ping, L. Liu, Analysis and design of IDMA systems based on SNR evolution and power allocation, in *IEEE VTC2004-Fall*, vol. 2, September 2004, pp. 1068–1072
45. L. Liu, J. Tong, L. Ping, Analysis and optimization of CDMA systems with chip-level interleavers. IEEE J. Sel. Areas Commun. **24**(1), 141–150 (2006)
46. X. Yuan, L. Ping, C. Xu, A. Kavcic, Achievable rates of MIMO systems with linear precoding and iterative LMMSE detection. IEEE Trans. Inf. Theory **60**(11), 7073–7089 (2014)
47. Y. Zhang, K. Peng, J. Song, Enhanced IDMA with rate-compatible raptor-like quasi-cyclic LDPC code for 5G, in *IEEE GC Workshops 2017* (2017)
48. K. Takeuchi, T. Tanaka, T. Kawabata, Improvement of BP-based CDMA multiuser detection by spatial coupling, in *IEEE ISIT 2011*, July 2011, pp. 1489–1493
49. D. Truhachev, C. Schlegel, Spatially coupled streaming modulation, in *IEEE ICC 2013*, June 2013, pp. 3418–3422
50. S. Kudekar, T. Richardson, R.L. Urbanke, Spatially coupled ensembles universally achieve capacity under belief propagation. IEEE Trans. Inf. Theory **59**(12), 7761–7813 (2013)
51. C. Schlegel, D. Truhachev, Multiple access demodulation in the lifted signal graph with spatial coupling. IEEE Trans. Inf. Theory **59**(4), 2459–2470 (2013)
52. A. Yedla, Y.Y. Jian, P.S. Nguyen, H.D. Pfister, A simple proof of Maxwell saturation for coupled scalar recursions. IEEE Trans. Inf. Theory **60**(11), 6943–6965 (2014)
53. S. Kumar, A.J. Young, N. Macris, H.D. Pfister, Threshold saturation for spatially coupled LDPC and LDGM codes on BMS channels. IEEE Trans. Inf. Theory **60**(12), 7389–7415 (2014)
54. D.J. Costello, L. Dolecek, T. Fuja, J. Kliewer, D. Mitchell, R. Smarandache, Spatially coupled sparse codes on graphs: theory and practice. IEEE Commun. Mag. **52**(7), 168–176 (2014)
55. D.G. Mitchell, M. Lentmaier, D.J. Costello, Spatially coupled LDPC codes constructed from protographs. IEEE Trans. Inf. Theory **61**(9), 4866–4889 (2015)
56. C. Liang, J. Ma, L. Ping, Towards Gaussian capacity, universality and short block length, in *IEEE ISTC 2016*, September 2016, pp. 412–416
57. C. Liang, J. Ma, L. Ping, Compressed FEC codes with spatial-coupling. IEEE Commun. Lett. **21**(5), 987–990 (2017)
58. H.H. Chung, Y.C. Tsai, M.C. Lin, IDMA using non-Gray labelled modulation. IEEE Trans. Commun. **59**(9), 2492–2501 (2011)
59. D.P. Bertsekas, R.G. Gallager, P. Humblet, *Data Networks*, vol. 2 (Prentice-Hall, Englewood Cliffs, NJ, 1987)
60. S. Ghez, S. Verdu, S.C. Schwartz, Stability properties of slotted Aloha with multipacket reception capability. IEEE Trans. Autom. Control **33**(7), 640–649 (1988)
61. M.H. Ngo, V. Krishnamurthy, L. Tong, Optimal channel-aware ALOHA protocol for random access in WLANs with multipacket reception and decentralized channel state information. IEEE Trans. Signal Process. **56**(6), 2575–2588 (2008)
62. C. Xu, P. Wang, S. Chan, L. Ping, Decentralized power control for random access with iterative multi-user detection, in *IEEE ISTC 2012*, August 2012, pp. 11–15
63. C. Xu, L. Ping, P. Wang, S. Chan, X. Lin, Decentralized power control for random access with successive interference cancellation. IEEE J. Sel. Areas Commun. **31**(11), 2387–2396 (2013)
64. H. Lin, K. Ishibashi, W.Y. Shin, T. Fujii, A simple random access scheme with multilevel power allocation. IEEE Commun. Lett. **19**(12), 2118–2121 (2015)
65. M. Zou, S. Chan, H.L. Vu, L. Ping, Throughput improvement of 802.11 networks via randomization of transmission power levels. IEEE Trans. Veh. Technol. **65**(4), 2703–2714 (2016)

66. C. Xu, X. Wang, L. Ping, Random access with massive-antenna arrays, in *IEEE VTC2016-Spring*, May 2016, pp. 1–5
67. Y. Hu, C. Xu, L. Ping, NOMA and IDMA in random access, invited paper, in *IEEE VTC2018-Spring*, June 2018, pp. 1–5
68. J. Mietzner, R. Schober, L. Lampe, W.H. Gerstacker, P.A. Hoeher, Multiple-antenna techniques for wireless communications-a comprehensive literature survey. IEEE Commun. Surv. Tutor. **11**(2) (2009)
69. G. Caire, N. Jindal, M. Kobayashi, N. Ravindran, Multiuser MIMO achievable rates with downlink training and channel state feedback. IEEE Trans. Inf. Theory **56**(6), 2845–2866 (2010)
70. D. Gesbert, S. Hanly, H. Huang, S.S. Shitz, O. Simeone, W. Yu, Multi-cell MIMO cooperative networks: a new look at interference. IEEE J. Sel. Areas Commun. **28**(9), 1380–1408 (2010)
71. F. Rusek, D. Persson, B.K. Lau, E.G. Larsson, T.L. Marzetta, O. Edfors, F. Tufvesson, Scaling up MIMO: opportunities and challenges with very large arrays. IEEE Signal Process. Mag. **30**(1), 40–60 (2013)
72. J. Hoydis, S. Ten Brink, M. Debbah, Massive MIMO in the UL/DL of cellular networks: how many antennas do we need? IEEE J. Sel. Areas Commun. **31**(2), 160–171 (2013)
73. L. Lu, G.Y. Li, A.L. Swindlehurst, A. Ashikhmin, R. Zhang, An overview of massive MIMO: benefits and challenges. IEEE J. Sel. Top. Signal Process. **8**(5), 742–758 (2014)
74. P.A. Hoeher, N. Doose, A massive MIMO terminal concept based on small-size multi-mode antennas. Trans. Emerg. Telecommun. Technol. **28**(2) (2017)
75. C. Xu, Y. Hu, C. Liang, J. Ma, L. Ping, Massive MIMO, non-orthogonal multiple access and interleave division multiple access. IEEE Access **5**, 14728–14748 (2017)
76. S. Vishwanath, N. Jindal, A. Goldsmith, Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels. IEEE Trans. Inf. Theory **49**(10), 2658–2668 (2003)
77. W. Yu, Uplink-downlink duality via minimax duality. IEEE Trans. Inf. Theory **52**(2), 361–374 (2006)
78. E.G. Larsson, O. Edfors, F. Tufvesson, T.L. Marzetta, Massive MIMO for next generation wireless systems. IEEE Commun. Mag. **52**(2), 186–195 (2014)
79. H. Schoeneich, P.A. Hoeher, Iterative pilot-layer aided channel estimation with emphasis on interleave-division multiple access systems. EURASIP J. Appl. Signal Process. **2006**, 250–250 (2006)
80. M. Zhao, Z. Shi, M.C. Reed, Iterative turbo channel estimation for OFDM system over rapid dispersive fading channel. IEEE Trans. Wirel. Commun. **7**(8) (2008)
81. C. Novak, G. Matz, F. Hlawatsch, IDMA for the multiuser MIMO-OFDM uplink: a factor graph framework for joint data detection and channel estimation. IEEE Trans. Signal Process. **61**(16), 4051–4066 (2013)
82. J. Ma, L. Ping, Data-aided channel estimation in large antenna systems. IEEE Trans. Signal Process. **62**(12), 3111–3124 (2014)
83. P. Hoeher, F. Tufvesson, Channel estimation with superimposed pilot sequence, in *IEEE GLOBECOM'99*, vol. 4 (1999), pp. 2162–2166
84. A. Ashikhmin, T. Marzetta, Pilot contamination precoding in multi-cell large scale antenna systems, in *IEEE ISIT 2012*, July 2012, pp. 1137–1141
85. J. Ma, C. Liang, C. Xu, L. Ping, On orthogonal and superimposed pilot schemes in massive MIMO NOMA systems. IEEE J. Sel. Areas Commun. **35**(12), 2696–2707 (2017)
86. Y. Chen, Low-cost superimposed pilots based receiver for massive MIMO in multicarrier system, in *IEEE VTC2017-Spring*, June 2017
87. Nokia, Alcatel-Lucent Shanghai Bell, Performance of interleave division multiple access (IDMA) in combination with OFDM family waveforms, document R1-165021, in *3GPP TSG-RAN WG1 #85* (2016)
88. Qualcomm Incorporated, RSMA, document R1-164688, in *3GPP TSG-RAN WG1 #85* (2016)
89. Samsung, Link level performance evaluation for IGMA, document R1-166750, in *3GPP TSG RAN WG1 Meeting #86* (2016)

90. S. Chen, B. Ren, Q. Gao, S. Kang, S. Sun, K. Niu, Pattern division multiple access—a novel nonorthogonal multiple access for fifth-generation radio networks. IEEE Trans. Veh. Technol. **66**(4), 3185–3196 (2017)
91. H. Nikopour, H. Baligh, Sparse code multiple access, in *IEEE PIMRC'13*, September 2013, pp. 332–336
92. HiSilicon, Huawei, LLS results for uplink multiple access document R1-164037, *3GPP TSG RAN WG1 Meeting #85*, May 2016
93. L. Ping, L. Liu, K.Y. Wu, W.K. Leung, Approaching the capacity of multiple access channels using interleaved low-rate codes. IEEE Commun. Lett. **8**(1), 4–6 (2004)