8

Evolved Physical Layer

Chapter Editors: A. Burr, L. Clavier, G. Dimic, T. Javornik, W. Teich, M. Mostafa and J. Olmos

8.1 Introduction

In this chapter, we discuss progress on the physical layer (PHY) for next generation wireless systems, covering the range of "smart environments". However, we omit the PHY of cooperative systems for the majority of the chapter, since this has been covered in Chapter 7.

It is well-appreciated that interference is the major limiting factor in wireless systems of all kinds: the capacity of most networks is limited by interference, whether from outside the network or from within it. It is important to characterise this interference as well as to mitigate its effects, so in Section 8.1, we focus in particular on interference that does not necessarily follow the Gaussian distribution, but instead is impulsive in nature. We show that this may enable us to mitigate its effect on receivers. IA is another approach for mitigating interference within networks, also discussed in this section, especially in the context of realistic channel state information (CSI) at transmitters.

Also of increasing importance is the energy efficiency (EE) of wireless networks, and in keeping with the emphasis of the whole book, we consider this issue in Section 8.2. For completeness, the effect of relaying on EE is also included here, including modelling of the energy consumption of wireless terminals. The benefits of energy harvesting are also considered.

Multiple-input multiple-output by now has become a standard feature of wireless networks of all kinds, but weaknesses remain in multiple-input multiple-output (MIMO) techniques, which have motivated further research to improve performance in future network, as reported in Section 8.3. This includes pre-coding for adaptation on realistic channels, and also two novel approaches: beam-space MIMO and spatial modulation. The effect of certain channel imperfections, such as peak amplitude limitations, are also addressed.

Cooperative Radio Communications for Green Smart Environments, 305–340.

Significant effort has continued on the development of iterative methods in modulation and coding, given the power of the approach in wireless receivers, and this forms a major focus of the work reported in Section 8.4. Iterative interference cancellation (IIC) and iterative decoding and demodulation are considered, and non-coherent demodulation is also discussed.

In evaluating the performance of complete wireless networks, it is rarely feasible to fully simulate the operation of the PHY of all links involved. This leads to the requirement for link abstraction, that is, the provision of models of the PHY to incorporate into system-level simulations. Such models are discussed in the final section of the chapter.

8.2 Dealing with Interference

There are two ways to consider interference: first as a noise resulting from the activity of other devices in the networks or from other networks. It can be too complex for some simple devices to make advance signal processing to reduce the impact of this interference. It happens, however, that this interference in many situations does not exhibit a Gaussian behaviour but is rather impulsive. In this section, we first propose an overview of impulsive interference models as well as a discussion about the dependence structure. We also propose some receiver structures that outperform the classical linear receiver in such a context. The second approach is to minimise the impact of the noise. This can be done by properly shaping the transmitted signal in order that interference is in an orthogonal space to the useful signal at the receiver, which is called interference alignment (IA) and will be addressed in the second part of the section.

8.2.1 Impulsive Interference

Interference is, in many situations, not properly modelled with a Gaussian distribution. Pioneer works in communications can be found in Middleton [Mid77], who obtained general expressions based on series expansion. If is work is widely used, infinite sums are difficult to handle in practice and simplifications have to be made, for instance only considering the most significant terms [Vas84, GDK06, AB07, VTNH14]. More recently, many works have been done concerning time hopping ultra-wideband. Analytical solution have been proposed [FNKS02, SMM03] but to simplify the receiver design, many works have proposed empirical distributions to represent interference plus thermal noise (see Beaulieu and Young [BY09] for an overview). Even more recently, analysis of networks has attracted a lot of works relying on

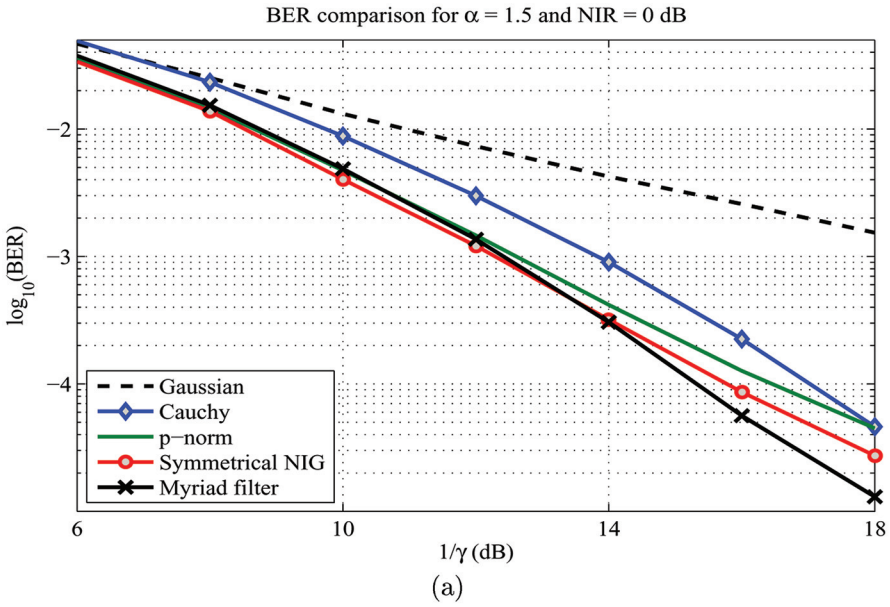
stochastic geometry. Although the first papers were published in the 90s [Sou92, TN95, IH98], an unbounded received power makes the interference fall in the attraction domain of a stable law. This can be seen as a consequence of the generalised central limit theorem [ST94, NS95]. The main advantage of the heavy-tailed stable distributions is their ability to represent rare events. In many communication situations, these events are in fact the events that will limit the system performance. The traditional Gaussian distribution ignores them leading to poor results. The proof of this result is generally done considering the log-characteristic function of the total interference (see for instance Sousa [Sou92], Win et al. [WPS09], and Ghannudi et al. [GCA+10]). This area of research is still active. Problems concerning the homogeneous position of users are studied, for instance based on cluster point process [GEAT10, GBA12] for general *ad hoc* networks or Poisson hole process for cognitive radio [LH12]. The dependence structure of interference is also attracting many works [GES12]: it is an important feature for the network analysis but difficult to handle.

8.2.2 Receiver Design

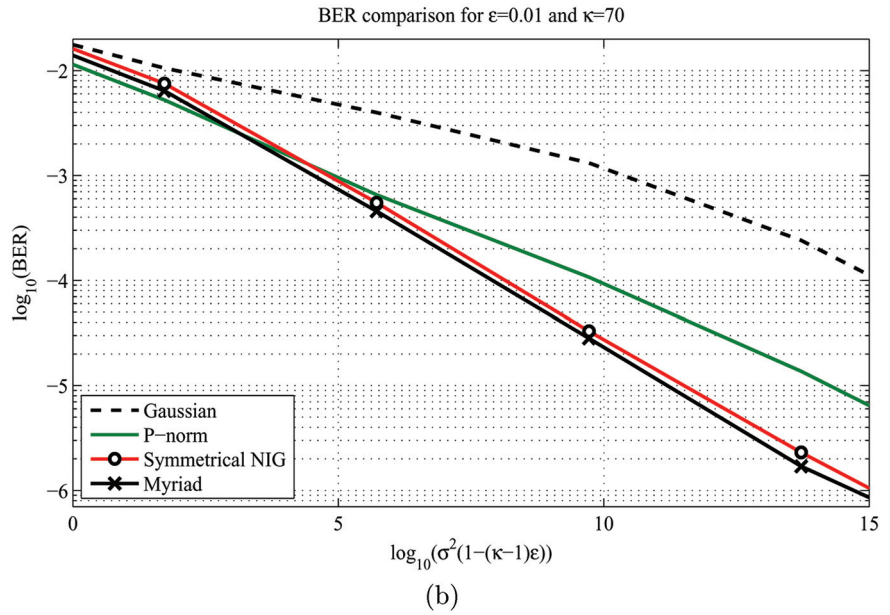
To illustrate the impact of impulsive interference, we consider a system where the symbol is repeated five times. We only represent here two significant examples: (i) a moderately impulsive case, represented by a mixture of α-stable ($\alpha = 1.5$) and Gaussian noises with, NIR = 0 dB (see Gu et al. [GPC+12] for more details); (ii) an ε-contaminated noise with $\varepsilon = 0.01$ and $k = 70$ reflecting rare but strong impulses (highly impulsive noise).

We show in Figure 8.1 the bit error rate (BER) performance of Gaussian, Cauchy, Myriad, symmetric normal inverse Gaussian (NIG), and p-norm receivers. When the noise involves an α stable impulsive interference, the BER is measured as a function of the inverse dispersion of the SαS distributions ($1/\gamma$) for network interference, since the increasing of inverse dispersion indicates the decreasing of the network interference strength, reflecting the conventional signal-to-noise-ratios (SNRs). For the ε-contaminated, the SNR at the receiver is used for the x-axis. The number of training samples for NIG, Myriad and p estimations is set to 1000 bits.

From the preceding figures and other approaches, we did not include here, we can make the following comments: linear receivers adapt poorly to impulsiveness even if optimal linear receivers [Joh96] are considered; a better approach is to find flexible families of distributions with parameters that can be easily and efficiently estimated. For instance, the Myriad approximation (based on Cauchy distributions) or the NIG distribution family [GPC+12] are



(a)



(b)

Figure 8.1 BER performance in (a) moderate impulsive environment and (b) α-contaminated noise.

good choices; log-likelihood ratio (LLR) approximation based approaches seem to have good potential. The intuitive approaches are the soft limiter and the hole puncher that limits the impact of the large values but remain less efficient than other non-linear solutions. On the other hand, the p-norm allows either a close to linear or linear behaviour when the thermal Gaussian noise is the main contribution to the noise and also approaches the *sharp* shape of the LLR when impulsiveness increases.

These works can be extended to more sophisticated receiver structures, for instance to channel coding [GC12b, MGCG13, DGCG14].

8.2.3 Dependence Structure

If many works consider independent and identically distributed (i.i.d.) interference random variables, space, time, or frequency diversity can result in vectors with dependent components. Mahmood et al. [MCA12] use the symmetric α-stable model, which, converted to its complex baseband form, is generally not isotropic, which means that the real and imaginary components are dependent. In Gulati et al. [GES12], the joint temporal statistics of interference in the network is derived along with an assumption of a bounded path-loss function. In Schilcher et al. [SBB12], closed-form expressions and calculation rules for the correlation coefficient of the overall interference are derived. Three sources of correlation are considered: node locations, channel, and traffic.

Whilst it may be possible to work with multivariate impulsive interference models, and especially multivariate α-stable models, it is very challenging in practice due to the intractable nature of the multivariate distribution function and because it requires the modelling of a high-dimensional spectral measure. We take a different approach [YCP+15]: in order to maintain the suitable marginal behaviours offered by α-stable models and to consider a relevant dependence structure, we use a meta-distribution model (see discussion in Fang et al. [FFK02] and Abdous et al. AGR05]) and rely on a parametric dependence model based on a copula construction, which is justified by the Sklar's theorem:

Sklar's Theorem: Let X and Y be random variables with distribution functions F and G respectively and joint distribution function H. Then there exists a copula C such that for all (x,y) $\varepsilon\mathbb{R}\times\mathbb{R}$ one has

$$H(x,y) = C(F(x),G(y)). \tag{8.1}$$

If F and G are continuous, then C is unique; otherwise, C is uniquely determined on $Ran(F) \times Ran(G)$, the cartesian product of the ranges of the

marginal cumulative distributions. Conversely, if C is a copula and F and G are distribution functions, then the function H defined by Equation (8.1) is a joint distribution function with margins F and G.

8.2.4 Interference Alignment

Interference alignment was introduced in Cadambe and Jafar [CJ08] as a means to allow spatial or temporal coordination among a group of users. The degrees-of-freedom (DoF) of the channel are exploited to confine all interfering signals to a common subspace, leaving the remaining orthogonal subspace for interference-free communication. Early IA studies showed promise of greatly increasing the number of co-channel users that could be supported with interference-free signal reception, but were based on unrealistic channel state knowledge assumptions: perfect instantaneous global CSI was assumed known at all nodes throughout the network, which is impossible in practice.

8.2.4.1 Reducing feedback requirements

In order to assess the feasibility of applying IA techniques to future communications systems, researchers have recently been looking at the impact of more realistic assumptions on IA techniques. One of the main hurdles to overcome in the implementation of IA systems is that of reducing the amount of feedback required to realistic levels. Several approaches in reducing feedback depend on quantisation using constellations of unit vectors uniformly distributed across the Grassmann manifold. In each of the methods described here, rather than directly quantising a vector or vectorised matrix, a vector that is representative of the desired information is selected from the constellation, and its index within the constellation is instead fed back, offering significant efficiency by reducing system overhead.

In Colman and Willink [CW11], a cooperative iterative algorithm was presented to reduce the overall signal-to-interference plus noise ratio (SINR) in a MIMO interference limited application using constellation-based pre-coding. This algorithm was designed not to require global channel information; the algorithm proceeds with information that each mobile station can measure locally. In every iteration of the algorithm, each receiver unit assesses the SINR impact of the transmitters using pre-coding vectors that are in the vicinity of the current pre-coding vector for each transmitter. Candidate pre-coding matrices that would improve the SINR are given a positive designation while candidate pre-coding matrices that would degrade the SINR are given a negative designation. This impact information is fed back to the transmitters where

each transmitter selects the pre-coding vector resulting in the greatest positive system impact. This algorithm was shown in Colman and Willink [CW11] to provide improved SINRs when compared with an optimal IA algorithm that had its feedback quantised using the same number of feedback bits.

The same authors presented a second IA algorithm based on local information in Colman et al. [CMW14]. In this algorithm, the receivers use local CSI to determine desirable pre-coders for all interfering transmitters. These pre-coding matrices are then quantised, using a Grassmannian constellation of pre-coding matrices, and this information is fed back to the transmitters. The next pre-coding matrices used at the transmitters are then a linear combination of the quantised desirable pre-coding matrices. In studying these two algorithms, the authors demonstrated the significant degradation that quantisation and imperfect or delayed channel information make on the interference rejection capabilities of IA techniques. Furthermore, these degradation effects are exacerbated when the common assumption of equal power from all transmitters at all receivers is removed.

Rather than feeding back CSI information, it was proposed in El Ayach and Heath [EAH12] to return the differential in CSI to the transmitters, which requires less feedback in slowly varying channels. This work was extended in Xu and Zemen [XZ13] by tracking the basis expansion coefficients on the manifold instead of the channel impulse response. This difference allows the transmitters to predict future channel realisations, which can compensate for inherent feedback delay, improving performance. Furthermore, the differential information of these basis expansion coefficients was quantised using a Grassmannian constellation, showing that effective IA performance could be maintained even with as few as two bits per channel realisation.

Prediction techniques were also harnessed for improved robustness to time-varying channels in Xu and Zemen [XZ14]. In this work, effects on IA performance, due to mobility and feedback delays, were mitigated by enabling reduced-rank channel prediction. The time-varying channel was approximated in a reduced-rank representation of orthogonal bases in the delay domain. These delay domain coefficients were then quantised using a random vector code-book distributed throughout the Grassmann manifold. Simulation results showed that a significant improvement in sum-rate is realisable when compared with non-predictive feedback strategies.

By stripping away assumptions, a better idea can be made of the viability of emerging technologies. However, in order to obtain a realistic idea of the potential benefits of technologies such as IA, experimental implementations

are of critical importance. Such an implementation was described in Zetterberg and Moghadam [ZM12] using a six-node setup to test MIMO, IA, and coordinated multipoint (CoMP) technologies, where each node had two antenna elements. System feedback was accomplished via cables and the nodes were fixed throughout the measurements. Experimental and post-processing SINRs show demonstrated a significant degradation between the ideal and achievable SINRs, due to hardware impairments.

8.2.4.2 IA applications

As IA becomes a more mature topic in the research community, the concept is being adapted and broadened from its initial concepts and considered as a method to mitigate interference in specific applications. In Kafedziski and Javornik [KJ13] the concept of IA, which is usually implemented in the space and/or time domains, was extended to the frequency domain with frequency–space IA using orthogonal frequency division multiplexing (OFDM). By using multiple antennas and multiple frequencies, the dimensionality of the signal space was increased, allowing exploitation to increase the number of signal streams or the number of users. By using IA techniques, both inter-cell and intra-cell interference could be removed effectively, demonstrating that IA may be a good technique to enable a frequency reuse factor of one in future wireless networks.

In Chatzinotas et al. [CSO13], IA concepts were extended by examining the effect of frequency packing on IA techniques. In a frequency-packed system, the carrier frequencies of adjacent signals are brought closer together in the frequency domain. This requires advanced interference mitigation techniques such as IA to suppress the additional interference. The work in Chatzinotas et al. [CSO13] was applied to a dual satellite coexistence scenario of monobeam satellites as primary and multibeam satellites as secondary users. It was shown that as the number of carriers in increased within a specified band, the overall sum-rate could be improved through effective interference mitigation. It was further shown that, with a coordinated IA technique, the primary rate could be perfectly maintained.

8.3 Energy Efficiency

8.3.1 On the Need for Improved EE

The EE in communications is of interest because of the increase in the number of communication devices. By increasing the EE, it may be possible to slow growth of the total energy consumption related to communications when the

number of devices increases. The EE in wireless communications has been studied by a working group of the GreenTouch initiative, the EARTH-FP7 project, and other projects. The IC1004 Action has contributed to the EE analysis in cooperative radio communications. The results show that the EE can be increased by various advances at the PHY and medium access control (MAC) layer.

One studied method is the two-hop relay. The introduction of relay with no other improvements at the PHY and MAC layers comes at the price of an additional energy consumer—the relay transceiver, reduced spectral efficiency, and increased source-to-destination delay. However, several advances are proposed which improve the overall EE, in terms of [bit/J], while not compromising other performance metrics. A key is the improvement of the source-relay and relay–destination channels with respect to (w.r.t.) the source–destination channel (Section 8.3.2). To improve channel quality of the relay links, the focus was on appropriate relay placement [DBB14b], or the use of receive diversity [PDID13, IDPD13]. The better channels enable power gain/saving to improve the EE subject to performance constraints, or to trade-off the EE with performance.

Other methods, which do not assume relaying, propose the use of transmission policy optimisation when energy harvesting is used and partial information on energy availability is provided [ZBMP13], beamforming instead of sectoral antennas [GC12a], and a variation of automatic retransmission request (ARQ) [BDZ13, ZBD14].

8.3.2 EE Increase by Use of Two-Hop Relays

An EE comparison between the two-hop relay and direct link has been developed in Dimić et al. [ZDB14, DZB12a, DZB12b, DZB13, DBB14a, DBB14b]. The evolution of the model, dubbed Dimić-Zogović-Bajić, is described below.

The major power consumer in a wireless transmitter is the power amplifier. A wireless transceiver power consumption (WPTC) model for Class A, AB and B power amplifiers is proposed and justified [ZDB14]. It shows the total transmitter power consumption, p_{TX}, as a function of the transmit power, i.e., the power delivered to antenna, p_{t}: $p_{\mathrm{TX}} = f(p_{\mathrm{t}})$. It is a monotone, increasing function, which generalises the affine function between the minimum and maximum p_{t} points, $(p_{\mathrm{t,min}}, p_{\mathrm{t,max}})$. The receiver power consumption, P_{RX}, is modelled as a constant. The energy consumption is $E_{\mathrm{TX}} = P_{\mathrm{TX}}T$ and $E_{\mathrm{RX}} = P_{\mathrm{RX}}T$, where T is a unit of time.

The WTPC model is related to the aggregate channel loss along the link, *L*, by a simple link budget: $p_{\mathrm{r}} = l(p_{\mathrm{t}}, L)$, where p_{r} denotes the received signal power; and a simple, deterministic condition for successful reception: $p_{\mathrm{r}} \geq P_{\mathrm{th}}$, where P_{th} denotes a received power threshold [DZB12a, DZB12b]. The P_{th} encapsulates several receiver parameters into one scalar, in particular, interference power at the receiver and the SINR [DZB13]. A perfect adaptive transmission power control is introduced for link adaptation (LA), so that $p_{\mathrm{r}} = P_{\mathrm{th}}$, which yields $p_{\mathrm{t}} = g(L, P_{\mathrm{th}})$. Then, along the source-destination link

$$p_{\mathrm{TX}} = f(g(L, P_{\mathrm{th}})). \tag{8.2}$$

Equation (8.2) shows that the total power consumption has to adapt to a random variable *L*, where *L* is not under control of the link-or-network designer.

Equation (8.2) is applied to the direct link, and the first and second relay hops, denoted by subscript *i* = void, 1, 2, respectively. Then, P_{TX}, $P_{\mathrm{TX},1}$, $P_{\mathrm{TX},2}$ are evaluated in the space of channel losses (L, L_1, L_2). In (L, L_1, L_2), a feasible energy consumption ratio between the two-hop relay and the direct link is

$$\alpha = \frac{(E_{TX,1} + E_{RX,1}) + (E_{TX,2} + E_{RX,2})}{E_{TX} + E_{RX}}. \tag{8.3}$$

The ratio α is a scalar-valued function in (L, L_1, L_2), parameterised by the transceiver characteristics. This evaluation yields an equipotential plane in the space (L, L_1, L_2), which separates the regions of energy consumption corresponding to the given value of α. The DZB energy consumption model for non-identical, i.e., heterogeneous, transceivers at the source, relay, and destination is derived in [DZB13].

The operating region (OR) of the transceivers is the region in (L, L_1, L_2) where a triplet (P_{t}, $P_{\mathrm{t},1}$, $P_{\mathrm{t},2}$) is feasible. The OR is bounded by $p_{\mathrm{t,i,min}}$ and $p_{\mathrm{t,i,max}}$, and respective $P_{\mathrm{th,i}}$, with *i* = *void*, 1, 2 [DZB13]. The DZB model analysis is applied to a comparison of EE of two-hop transmission via Type la relay and the direct link in LTE-advanced (LTE-A), between the same Base station (BS) and user equipment (UE) [DBB14a, DBB14b]. It is shown how the transceivers parameters and (L, L_1, L_2) influence α and that the supportive relay may provide energy savings.

In the OR of the space (p_{t}, $p_{\mathrm{t},1}$, $p_{\mathrm{t},2}$), the required power levels to enable certain modulation and code scheme (MCS) (i.e., constellations and code rate pairs) are selected. These MCS levels define the data rate ratio between the two-hop relay data rate, r_{R}, and direct link data rate, *r*,

$$\beta = \frac{r_{\mathrm{R}}}{r} \tag{8.4}$$

Good channels L_1 and L_2 w.r.t. L facilitate increase in r_{R} w.r.t. r, and so β increases. Data rates of each relay hop must be 2 r_{R}

Overlapping the α-levels with the β-ratios quantifies the EE ratio

$$\eta = \frac{\beta}{\alpha}. \tag{8.5}$$

The increase in EE is achieved by maximising $p_{\mathrm{t},1}$ and $p_{\mathrm{t},2}$, and by selecting the MCS which maximises r_{R}, subject to the successful reception. In addition, a trade-off between increasing the energy and spectral efficiency versus reducing the generated interference when using a regenerative relay is proposed. Figure 8.2 shows α-levels in (L_1, L_2) space for a given L. Moving the operating point to the upper-right corner, by increasing $P_{\mathrm{t},1}$, $P_{\mathrm{t},2}$ improves EE.

In Tralli and Conti [TC11], it is shown that in the presence of other cells, the mutual interference can be a limiting factor and needs to be characterised. A framework for analysis and design of relay-assisted communications, which accounts for multichannel communications in each link, geometrical setting, interference, fading channel, and power allocation, while being general with

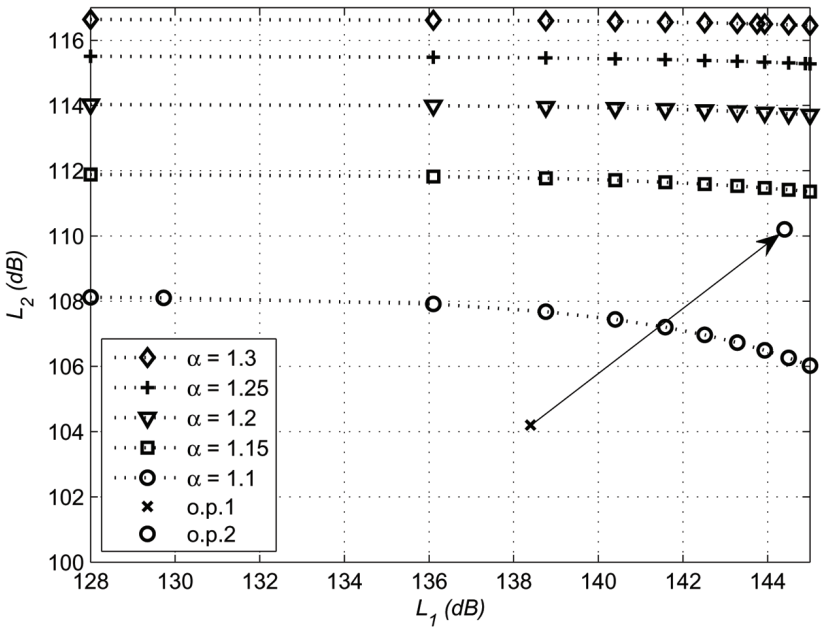


Figure 8.2 EE Example in (L_1, L_2): moving the operating point from o.p.l to o.p.2, by increasing $p_{\mathrm{t},1}$ and $p_{\mathrm{t},2}$ by 6 dB each, increases the constellation from 16QAM to 64QAM, which yields $\beta = 2$ and $\eta = 1.75$.

respect to MCS is proposed. The framework differs from the DZB model in the following:

1. There is a geometrical setting, with defined distances between the source, relay, and destination in the cell of interest and the interfering cell; DZB absorbs geometry in the channel losses (L, L_1, L_2);
2. The destination combines received signals from both the source and relay; DZB does not use receive combining;
3. Tralli and Conti [TC11] considers Rayleigh fading and outage probability (OP); DZB considers a simple condition for successful reception $p_{\mathrm{r}} \geq P_{\mathrm{th}}$;
4. Tralli and Conti [TC11] considers two power allocation techniques: (i) uniform (equal at the source and relay), and (ii) balanced (such that p_{r} is equal over source–destination and relay–destination links); DZB considers adaptive power control tightly linked to the total channel loss along a link;
5. Tralli and Conti [TC11] accounts for transmit power only, when evaluating EE; DZB considers total transceiver power consumption, p_{TX}.

The performance with and without relay-assistance is compared to quantify the benefits of relay and power allocation in terms of energy saving with respect to direct transmission [TC11]. The reference scenario assumes two source–relay–destination triplets. The first triplet is in the cell of interest, and the second is in an interfering cell. At each link, there are n_{T} transmitting antennas and 1 receive antenna.

The OP at the destination is evaluated for the source-destination data rate of R (*bit/s/Hz*). The OP optimal power allocation: in the absence of interference, with large enough link diversity, it tends to be balanced power allocation; in the presence of interference, it is not easy to find because of the mutual interference between the useful and interfering links. The average energy per bit required to support a given rate within the given coverage distance is evaluated. The use of relay with balanced power allocation provides power gains at medium and low rates for light frequency reuse, and only at low rates in heavy reuse. When power control is used, to guarantee the target OP, the average energy per bit is lower in both light and heavy reuse cases when relay is used Figure 8.3. The use of power control provides a 5-dB energy saving in all cases.

In Pejanović-Djurišić and Ilic-Delibašić [PDID13] and Ilić-Delibašić et al. [IDPD13], performance of two-hop relay system operating in Ricean-fading environment is studied. Similar to Tralli and Conti [TC11], both Pejanović-Djurišić and Ilic-Delibašić [PDID13] and Ilić-Delibašić et al. [IDPD13]

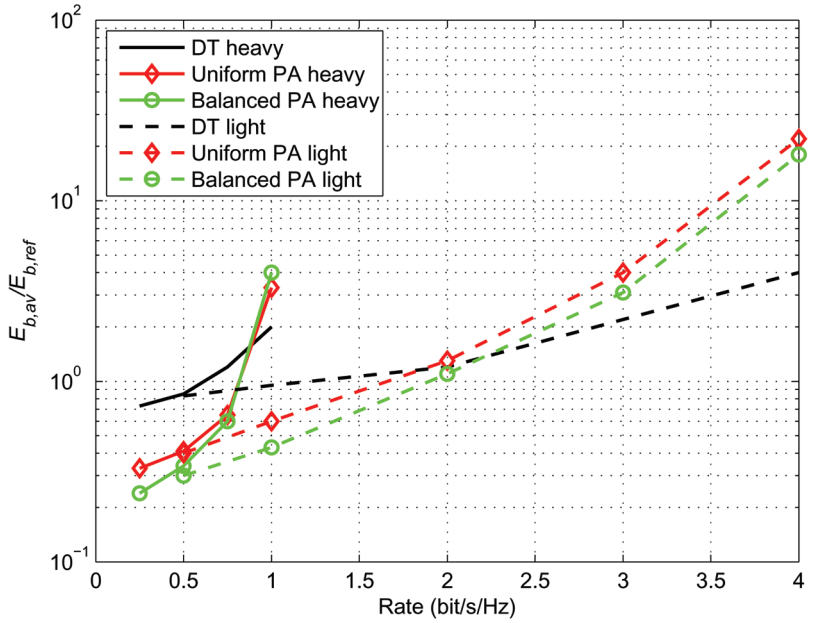


Figure 8.3 Normalised average bit energy required to guarantee the target rate at OP with power control.

account for transmit power only. Implementation of a dual-polarised antenna at the destination node is assumed.

The level of performance improvement for decode-and-forward (DF) relay systems is analysed, in terms of BER, accomplished by carefully planning the position of a relay node, to provide direct line-of-sight (LoS) signal component in both the source–relay link and in the relay–destination link. As the Ricean-fading K coeffcient increases, the SNR gain increases for the target BER. Therefore, it is desirable to provide LoS along both of the relay links. This result is aligned with the results of the DZB model.

If fading parameters along the source–relay and relay–destination links are different, the BER performance worsens. Typically, the relay–destination link has worse fading parameters, which motivates the implementation of polarisation diversity at the destination.

Two correlated and non-identical Ricean-fading channels are assumed for the relay-to-destination link, while maximal ratio combining (MRC) of the received signals is performed at the receiver. To determine effects of the polarisation diversity implementation, performance of the considered system

is compared with the standard DF relay system with single antenna terminals. Using polarisation diversity, the same BER values can be obtained with significantly lower SNR, despite a certain level of correlation and power imbalance between the vertical and horizontal diversity branches. Compared to the system with no diversity, the total needed transmit power for achieving the same level of the relay system performance is reduced, Figure 8.4. Thus, increased EE is enabled.

8.3.3 EE Increase by Other Methods

When energy is harvested by wireless nodes from renewable sources, its availability becomes uncertain and its use for communications must be carefully designed. While the optimal power allocation has been derived in previous works when energy availability is fully known *a priori*, practical algorithms are needed when only causal and statistical information is available. In Zanella et al. [ZBMP13], the optimal transmission policy is studied

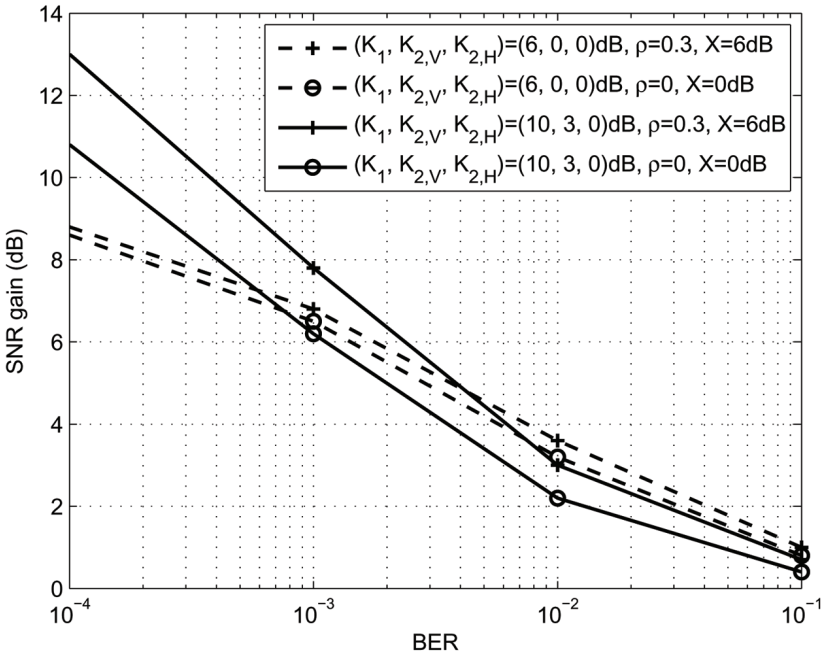


Figure 8.4 LoS along both relay hops and polarisation diversity at destination improve SNR gain and EE. K, Ricean coefficient; V, H, vertical, horizontal polarisation component; ρ, correlation coefficient between received signal envelopes; and X, cross-polar discrimination.

when only the statistical distribution of the energy arrival intervals is known and no information is available on the amount of energy that will be harnessed.

First, an exact solution for the case of a step-wise transmission power profile is obtained. This result is then extended to the time-continuous case. Within energy arrival intervals, the obtained power profile is shown to be non-increasing as a function of time, and non-decreasing as a function of the residual energy. For an exponentially distributed energy arrival process, the optimal policy can be expressed in a closed form. The numerical results for this case corroborate the main results. The optimal power allocation policy of Zanella et al. [ZBMP13] outperforms constant power allocation policy in every energy arrival interval, which increases the EE.

In Goncalves and Correia [GC12a], beamforming is studied to assess the efficiency in terms of radiated power that is possible to be saved when antenna arrays are placed instead of static sector antennas, for the universal mobile telecommunications system (UMTS) and long-term evolution (LTE) radio interface. Several multiple user scenarios were studied with different users' arrangements. Two simulators, one for UMTS and another for LTE, were developed to statistically evaluate the potential impact that adaptive antenna arrays have to reduce the radiated power, compared with actual BS static sector antennas. UMTS, besides signal improvement, has a lot of interference suppression potential due to its multiple access technique that separates users by codes in the same carrier frequency. LTE, due to the absence of co-channel inter-cell interference, is evaluated in terms of desired signal improvement. A model for UMTS is derived that describes the power improvement achieved as a function of the number of users and of radiator elements. For UMTS carriers near top capacity, a power reduction of the order of 90% is achievable. For LTE, significant power improvements are reached, especially for antenna arrays with eight elements, which can save near 65% of the radiated power.

The maximal number of transmission attempts in ARQ schemes is a positive integer, thus constraining the decision space [ZBD14]. A new retransmission scheme, r-ARQ, is proposed, where the expected value of the maximal number of allowed transmission attempts is a non-negative real number. Dependence of EE, throughput, delay, and jitter on the packet-loss and the expected value of maximal number of allowed transmission attempts have been derived. The packet-loss rate and the expected number of transmission attempts are compared for two examples of r-ARQ and traditional ARQ scheme. The examples show different behaviour, so that the r-ARQ should be designed specific to the optimisation problem.

8.4 MIMO Systems

This section addresses recent research in MIMO wireless communication systems. The research was focused on solving key weaknesses of MIMO systems, namely, (i) hardware and software complexity including needs for multiple RF chains and (ii) number of antennas and their spacing and size. Furthermore several documents looking at problems caused by distortions due non-ideal hardware and constraints due to limited feedback. This section is organised as follows. The first subsection looks at pre-coding techniques studying impact of realistic environment on its performance including limited feedback and realistic propagation environment. Two new approaches of applying pre-coding were also proposed: (i) considering CSI and data in pre-coding and (ii) usage of pre-coding for security in MIMO systems. The second and third subsection deals with spatial dimension of MIMO system namely spatial multiplexing and beamspace MIMO. Next section is devoted to the hardware impairments in MIMO system, while the topic of last section is adaptive MIMO systems.

8.4.1 Pre-Coding

The MIMO system consisting of very large number of antennas at the BS was evaluated in theoretical i.i.d Gaussian channel and measured residential-area channel in Gao et al. [GERT11] for three pre-coding scheme, namely dirty paper coding [APNG03], Zero-forcing pre-coding and minimum mean squared error (MMSE) pre-coding. It was found that the user channels, in the studied residential-area propagation environment, can be de-correlated by using reasonably large antenna arrays at the BS. With linear pre-coding, sum-rates as high as 98% of dirty paper coding capacity were achieved for two single-antenna users already at 20 BS antennas. This shows that even in realistic propagation environments and with a relatively limited number of antennas, we can see clear benefits with using an excessive number of BS antennas.

A new method was proposed to derive direction-of-arrival (DOA) information in a multiple-input single-output (MISO) or single-mode MIMO communications system that uses constellation-based limited feedback pre-coding in Colman [CWW12]. DOA estimation was accomplished by matching the distribution of selected codewords, $q(l)$, to a set of reference pre-coder index distributions generated using an idealised Ricean model under a range of K-factors. The DOA was estimated by minimising the Bhattacharyya distance between distributions. The method of estimating DOA was tested using time-varying channels generated with the WINNER2 rural LoS and non-line-of-sight (NLoS) models. The results presented in Figure 8.5 shows that good

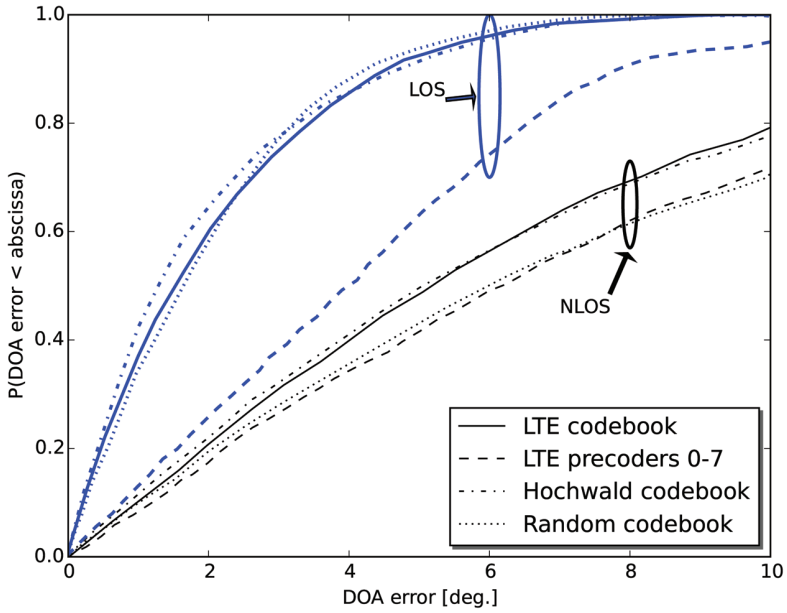


Figure 8.5 DOA error cdfs when generating $q(l)$ using 1000 samples over 4 m of mobile motion for the WINNER2 rural LoS, and NLoS scenarios.

direction estimates can be obtained with the algorithm presented, with median DOA estimation errors of less than 2°, regardless of the constellation used. Although the LTE codebook was designed with several directional pre-coders, the Hochwald constellation yields slightly better LoS error estimates.

The pre-coding approach was also applied to reduce the ability of a third-party eavesdropper to successfully receive pre-coded MIMO signals [Coll3]. Instead of transmitting the pre-coding matrix chosen by the intended receiver, a secondary constellation of pre-coding matrices clustered around this selected matrix is used in a pseudo-random sequence, effectively transmitting with a spatially hopped signature. Polar-cap constellations proposed in Choi et al. [CCLK12] were suggested as a good candidate for the secondary constellation, which would be shared only among a desired group of users. The results presented in the document show significant degradation of eavesdropper reception while maintaining of the performance of the desired user nearly unchanged. This effective method to control eavesdropping comes at little additional system cost; both transmission and reception require few additional resources other than the knowledge of the sequence and scale of the spatial hopping behaviour.

The focus of Farah et al. [NEGJ11] lays on various spreading and pre-coding techniques for the MIMO scenario in LTE wireless communication systems. Two spreading/pre-coding schemes were compared, namely large-delay cyclic delay diversity (CDD) and multi-carrier cyclic antenna frequency spreading (MC-CAFS). In CDD, cyclically delayed replicas of one symbol are transmitted from the different antennas at the same time. In an OFDM system the CDD approach increases the space diversity. While for MC-CAFS, one symbol is sent over all the available subcarriers and over all available antenna ports which results in diversity in space and frequency. The simulation results show that the MC-CAFS spreading scheme shows a better performance compared to the CDD scheme by capturing more frequency diversity.

The concept of jointly utilising the data information and channel state for designing symbol level pre-coders for MISO downlink channel was proposed in Alodeh et al. [ACO15]. The interference causing by simultaneous data streams is transformed to constructive interference, which improves the signal to interference noise ratio at the receiver. The document proposes a new algorithm, which is based on maximum ratio transmissions algorithm. The simulation results plotting EE show the proposed algorithm superior performance to constructive rotation zero forcing pre-coding algorithm [Masll] on the expense of higher complexity due to the need for solving a non-linear set of equations.

8.4.2 Spatial Modulations

The spatial modulation (SM), introduced by Mesleh et al. [MHS+08], was proposed to overcome inter-channel interference and to increase the spectral efficiency. This is attained through the adoption of a new pragmatic transmission approach which employs (i) the activation of a single antenna that transmits a given data symbol (constellation symbol) at each time instance and (ii) the exploitation of the spatial position (index) of the active antenna as an additional dimension for data transmission (spatial symbol). Both the constellation symbol and the spatial symbol depend on the incoming data bits. The spectral efficiency is increased by the logarithm to base 2 of the number of transmit antennas. The number of transmit antennas must be power of two. This approach resulting in the simplified hardware requirements of MIMO communication systems and avoids inter-channel interference, while still benefiting from increased spectral efficiency.

The simple receiver hardware requirements of the spatial modulation is well suited to mobile terminals and initiated a research of its application

as a fourth generation (4G) wireless communication techniques [WMJ$^+$]. A sensitivity of spatial modulation to variations in sub-channel properties using both statistical channel models and measured urban channel data was investigated in Thompson et al. [WAM$^+$]. The uncoded system performance was studied when transmitting over multi-antenna channels with variable correlations between spatial sub-channels and sub-channel power balance. The analysis aimed at finding out whether the variations in system performance over the measured channels can be recreated with the simple channel model used to create the synthesised channel model. The technique applied to create the synthesised channel is based on Kroneker model applying independent and identically distributed random variable, with complex normal distribution to create un-correlated MIMO channel and correlation matrix to introduce correlation between sub-channels, while the measured channel was obtained by extensive multi-antenna channel measurement in Bristol at 2 GHz with 20 MHz bandwidth using channel sounders configured as a 4×4 MIMO system. The simulation results when varying sub-channel correlation showed significantly more variation in BER results for the synthesised channels compared to the measured ones. The results for the spatial modulated system BER sensitivity to sub-channel branch power variations for the synthesised channels showed a strong correlation between the chosen measure of branch power imbalance and the SNR required for a specific BER performance. This strong relationship was not found for the measured channels (Figure 8.6).

8.4.3 BeamSpace MIMO

Beam-space MIMO system was proposed in Kalis et al. [KKP08] as a means to address the two key weaknesses of conventional MIMO systems namely the antenna size and the need for multiple RF chains. The research effort on Beam-space MIMO focuses on the development of functional MIMO transmission schemes with the use of a single RF chain while maintaining small antenna size using on electronically steerable passive array radiators (ESPARs). Beam-space MIMO applies a set of orthonormal basis antenna patterns to provide efficient multiplexing and beamforming, further using small-sized antenna arrays. Beam-space MIMO systems outperform conventional MIMO systems in terms of system capacity Kalis et al. [KKP08]. Research presented in Maliatsos et al. [MVK13a] and [MVK13b] makes first steps toward practical system design and focuses on basic channel estimation and a design and implementation of V-BLAST receiver for Beam-space MIMO. Exploiting the fact that there is equivalent algebraic representation of BS and conventional

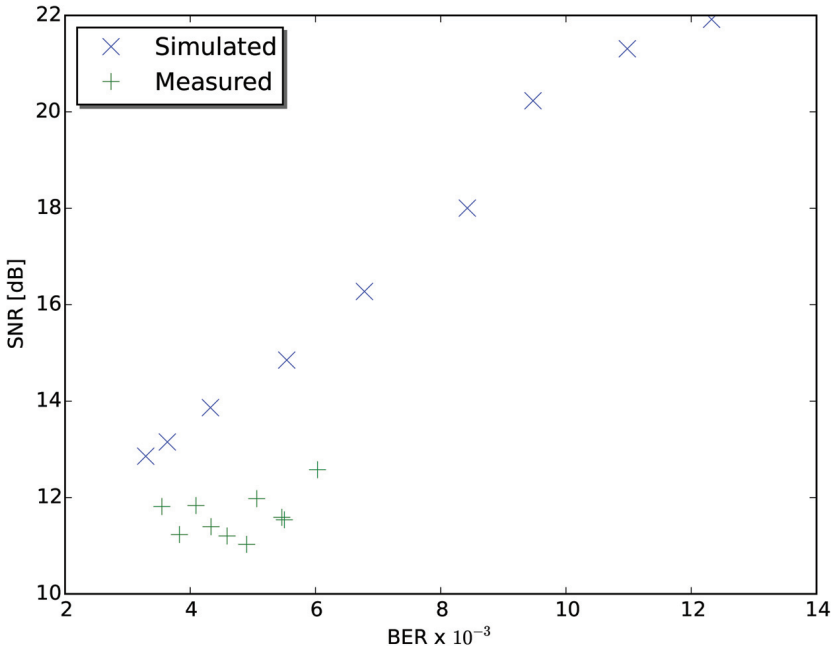


Figure 8.6 Scatter plot showing the required SNR for a BER of 10^{-3} for the measured and synthesisd channels.

MIMO systems, the algorithms and the training sequences applied in conventional MIMO systems are transferred in beamspace domain. Least Square and MMS estimators were applied to find channel estimates, which were used for basis antenna pattern reconfiguration. Beam-space MIMO system with channel estimation was evaluated applying system level simulations using rich scattering wireless world initiative new radio (WINNER) $\hat{a}\check{A}\acute{Z}$B2$\hat{a}\check{A}\acute{Z}$ channels. The performance comparison between three-element Beam-space MIMO and conversional MIMO system is shown in Figure 8.7. It is clear that Beam-space MIMO systems outperform conventional MIMO for small inter-element distances despite the fact that noise power is increased due to oversampling. However, it became clear [MVK13a] that this weakness limits Beam-space MIMO performance in low and medium SNRs.

In real-world systems the transmitter is not able to have full access to CSI. This fact initiated studies in a limited feedback technique in Beam-space MIMO based on conventional MIMO algorithm with limited feedback [MVK13b]. The codebooks from the beamspace channel partitioning study

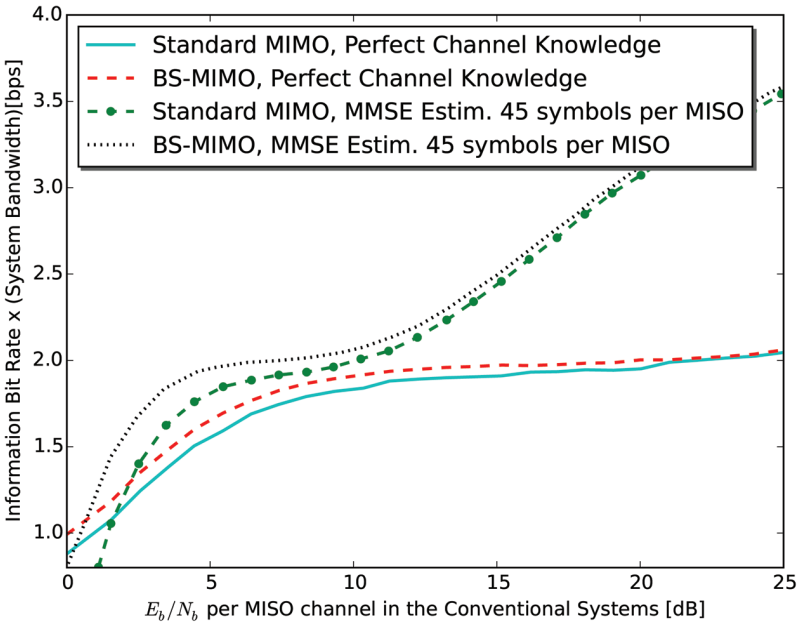


Figure 8.7 Performance comparison between three element Beam-space MIMO and conventional MIMO systems in terms of information bit rate in the PHY.

were analysed to map the computed patterns in reactance values providing a guide for the accuracy requirements in the antenna circuit design. Since CSI in the Transmitter (CSIT) is limited, the well-known V-BLAST reception (with MMSE detection) is adapted in the beamspace and adaptive pattern reception schemes were presented and analysed. The simulation results show a gap between system with complete CSIT (CSI at the transmitter) and the V-BLAST receiver without CIST, while the systems with limited feedback (4 and 5 bits) outperform conventional V-BLAST algorithm [MVK13b].

8.4.4 Hardware Realisation and Interference Impact on the Performance of MIMO Systems

Hardware implementation of MIMO system may significantly degrade its performance. Phase noise, non-linearities in transmission chain and I&Q imbalance have huge impact. The results in Zetterberg [Zetll] and Bjornson et al. [BZBO13] reveal that the MIMO systems distorted by hardware impairments have a finite capacity limit in the high-SNR regime—this is

fundamentally different from the unbounded asymptotic capacity for ideal transceivers. Furthermore, the bounds in Equation (8.6)

$$M \log_2 \left(1 + \frac{1}{\kappa^2}\right) \leq C_{N_\mathrm{t} N_\mathrm{r}}(\infty) \leq M \log_2 \left(1 + \frac{N_\mathrm{t}}{M\kappa^2}\right) \tag{8.6}$$

where N_r number of receiver antennas, N_t number of transmitter antennas, $M = \min(N_\mathrm{t}, N_\mathrm{r})$, and κ is level of impairments. The bounds hold for any channel distribution and are only characterised by the number of antennas and the level of impairments.

The BER performance of MIMO-OFDM/TDM based on MMSE-FDE in a peak-limited and frequency selective fading channel was analysed in Ligata et al. [LGJ14]. The analysis was based on the Gaussian approximation of the residual intersymbol interference (ISI) after MMSE–FDE and the noise due to peak limitation. The theoretical MMSE equalisation weights for MIMO-OFDM/TDM in a peak-limited channel were derived to capture the negative effect of HPA saturation. The performance was evaluated by both numerical Monte-Carlo method and theoretical analysis. The theoretical results confirmed that the BER performance of MIMO-OFDM/TDM using MMSE-FDE in a peak-limited and frequency-selective channel is a function of the OFDM/TDM system design, i.e. *K* on number of block in OFDM/TDM system. This is also illustrated in Figure 8.8, where *L* denotes number of reflected ray and ϕ high-power saturation level.

An evaluation of the information theoretic capacity of MIMO communication system with the presence of interference is given in Webb et al. [WBN04]. Three different assumptions of CSI of both desire signal and interference signal was considered: namely (i) full information of channel and interference information are known at the transmitter, (ii) neither channel nor interference information are known at the transmitter, and (iii) only channel information without interference is known at the transmitter. Outage capacity with varies numbers of interferers, array size and SINR power was compared. Alignment of interferer subspace gives an upper and lower bound for the MIMO capacity with interference.

8.4.5 Adaptive MIMO Systems

Adaptive modulation is shown to be an effective approach to increase the effective data rate in multi-carrier communication systems [CCB95]. However; the transmitter and the receiver should have a perfect knowledge about the bit allocation per subcarriers. Transferring bit allocation per subcarriers from

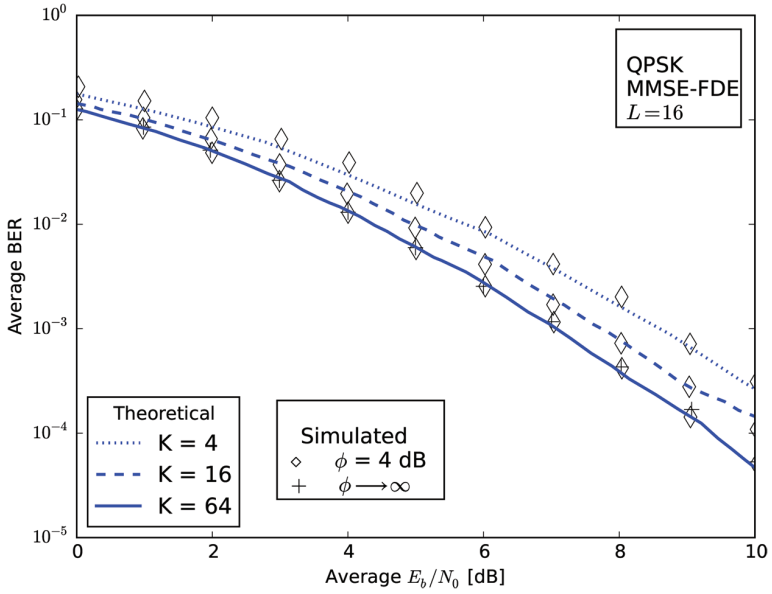


Figure 8.8 BER versus $E_\mathrm{b} = N_0$ with K as a parameter.

transmitter to the receiver may significantly reduce the effective link throughput. A signal-assisted maximum *a posterior* (MAP) modulation classification algorithm was proposed in Haring and Kisters [HK14] in time-division-duplex wireless communications. The method utilises the received signal form, truncated signalling information and channel reciprocity in wireless TDD systems. Simulations show a significant system performance improvement of the proposed schemes over conventional schemes applied in typical indoor environment (Figure 8.9).

8.5 Modulation and Coding

This chapter summarises the results of the COST IC1004 project obtained in the fields of modulation and coding. An important aspect has been the application of iterative methods. In Section 8.5.1 iterative methods are utilised for equalisation, leading to a low complexity solution and a dramatic improvement of the power-to-speed ratio. Section 8.5.2 deals with iterative (turbo) detection for concatenated systems. In Section 8.5.3 advanced coding schemes and iterative decoding methods are analysed. Section 8.5.4 finally considers the non-coherent transmission methods differential phase shift keying (DPSK) and OFDM-MFSK.

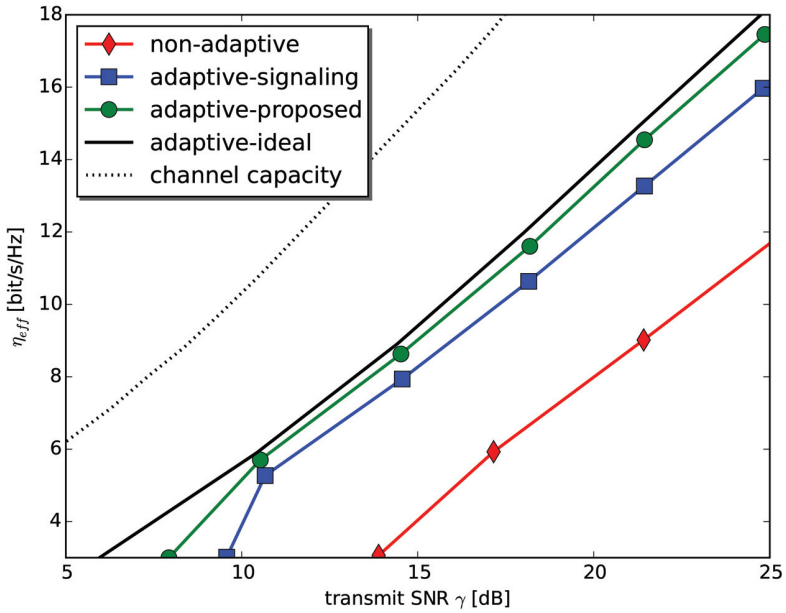


Figure 8.9 Effective bandwidth efficiency η_{eff} at target frame error rate (FER) $FER_{\text{target}} = 10^{(1)}$ versus transmit SNR for 4×4 SVD transmission.

8.5.1 Iterative Interference Cancellation

Multipath propagation on the physical channel between transmitter and receiver leads to ISI. In case of multiuser, multi-subchannel, multi-antenna transmission systems, and combinations thereof, additional interference emerge [Lin99]. To cope with this interference, equalisation has to be applied at the receive side. Because of the high computational complexity of the optimum equaliser, suboptimum schemes are usually utilised, mainly iterative ones because of their good complexity–performance trade-off. We introduce here recurrent neural network (RNN) as a suboptimum IIC technique.

Discrete-time RNN can be adapted to perform IIC without the need for a training phase. Stability and convergence proofs in this case are available. Properties without which an iterative scheme often is suspect. The same applies for continuous-time RNNs, which serve as promising computational models for an analogue hardware implementation. The later one combines reduced energy consumption with the fast signal processing required by the IIC for high-data rate transmission. An improvement of the power-to-speed ratio of up to four orders of magnitude compared to a standard digital solution has been obtained [OSM+14].

8.5.1.1 Transmission model

The uncoded discrete-time (on symbol basis) vector-valued block transmission model for linear modulation schemes is shown in Figure 8.10.

In details:

- SRC (SNK) represents the vector-valued digital source (sink).
- $\mathbf{q}(\hat{\mathbf{q}}) \in \{0,1\}^{N_\mathrm{q}}$ is the vector of source (hard estimated) bits of length N_q.
- $\hat{\mathbf{s}} \in \psi^{N_\mathrm{t}}$ is the (hard estimated) transmit vector of length N_t. $\psi = \{\psi_1, \psi_2, ..., \psi_{2^m}\}$ represents the symbol alphabet where $m \in \mathbb{N}/\{0\}$. $M = 2^m$ represents the cardinality of the symbol alphabet. In this case, there exist M^{N_t} possible transmit vectors. The mapping from $\mathbf{q}$ to $\mathbf{s}$ is bijective and performed by $\boldsymbol{\mathcal{M}}$. We notice that the number of possible transmit vectors increases exponentially with the block size N_t.
- $\mathbf{r}$ is the receive vector of length N_t.
- $\mathbf{R}_\mathrm{s}$ is the (discrete-time) channel matrix for the block transmission of size $N_t \times N_\mathrm{t}$. It includes the transmit and receive filter as well as the impulse response of the channel. Using a channel matched filter as receive filter, $\mathbf{R}_s$ is hermitian and positive semi-definite.
- $\mathbf{n}$ is a sample function of an additive Gaussian noise vector process of length N_t with zero mean and covariance matrix $\boldsymbol{\Phi} = \frac{\mathrm{N}_0}{2} \cdot \mathbf{R}_\mathrm{s}$. $\frac{\mathrm{N}_0}{2}$ is the double-sided noise power spectral density.
- IC: Interference cancellation (a suboptimum equalization technique).

The model in Figure 8.10 is a general model and fits to different transmission schemes like OFDM, code division multiple access (CDMA), Multicarrier code division multiple access (MC-CDMA), MIMO and MIMO-OFDM. The relation with the original physical continuous-time model can be found in Lindner [Lin99]. Mathematically it is described as follows

$$\begin{aligned} \mathbf{r} &= \mathbf{R}_\mathrm{s} \cdot \mathbf{s} + \mathbf{n} \\ \mathbf{r} &= \underbrace{\mathbf{D} \cdot \mathbf{s}}_{\textbf{Signal}} + \underbrace{\mathbf{C} \cdot \mathbf{s}}_{\textbf{Interference}} + \underbrace{\mathbf{n}}_{\textbf{Additive noise}} \end{aligned} \tag{8.7}$$

$\mathbf{D}$ contains only the diagonal elements of $\mathbf{R}_\mathrm{s}$. Off-diagonal elements of $\mathbf{D}$ are zeros. $\mathbf{C}$ contains only the off-diagonal elements of $\mathbf{R}_\mathrm{s}$. Diagonal elements of

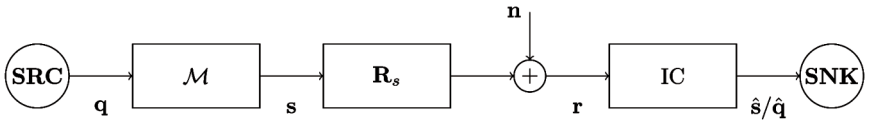


Figure 8.10 Uncoded discrete-time (on symbol basis) vector-valued block transmission model for linear modulation schemes.

$\mathbf{C}$ are zeros. We notice that the non-diagonal elements of $\mathbf{R}_{\mathrm{s}}$, namely $\mathbf{C}$ lead to interference between the elements of the transmit vector at the receive side (cf. Equation (8.7)). For interference-free transmission $\mathbf{C} = 0$.

The interference cancellation (suboptimum equalization) in Figure 8.10 has to deliver a vector $\hat{\mathbf{s}}$ which must be as similar as possible to the transmit vector $\mathbf{s}$ given that ψ and $\mathbf{R}_{\mathrm{s}}$, are known at the receive side. Mapping $\hat{\mathbf{s}}$ to $\hat{\mathbf{q}}$ for an uncoded transmission is straightforward. Equalisation can be seen as a classification process and represents a non-linear discrete optimisation problem.

8.5.1.2 IIC: A generic structure

The optimum equalizer, the maximum likelihood (ML) one, calculates for every receive vector $\mathbf{r}$ a distance (more precisely the Mahalanobis distance because of the coloured noise $\mathbf{n}$) to all possible transmit vectors $\mathbf{s}$ of cardinality $M^{N_{\mathrm{t}}}$ and decides in favour of that transmit vector $\hat{\mathbf{s}}$ with the minimum distance to the receive vector $\mathbf{r}$. Thus, the computational complexity of the optimum equalizer is too high for a realistic block length N_{t}. The computational complexity in general increases exponentially with the block length N_{t}. Therefore suboptimum equalizer schemes, mainly iterative ones, are applied. Figure 8.11 shows a generic structure of IIC.

The basic idea is to obtain an estimate (soft or hard) of the transmit vector. Based on this estimate, the interference is recreated (linear function of the estimate). The rebuild interference is then subtracted from the receive vector $\mathbf{r}$. This process is repeated on symbol or vector basis (serial or parallel update). Depending on the estimation function we distinguish:

- *Linear* IIC: The estimation function is linear, as an example the zero-forcing block linear equaliser (ZF-BLE) [HYKY03].
- *Non-linear* IIC: The estimation function is non-linear, as an example the multistage detector (MSD) [VA90].

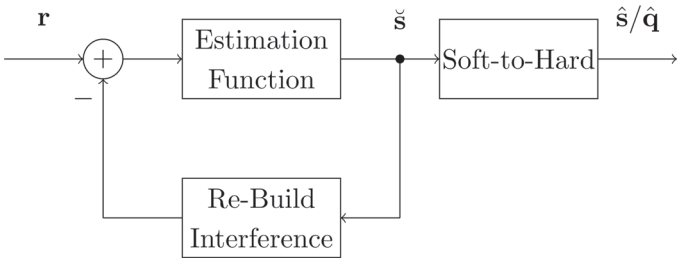


Figure 8.11 A generic structure of IIC, $\mathbf{r}$ is the receive vector, $\breve{\mathbf{s}}$ is the "soft" estimated transmit vector, $\hat{\mathbf{s}}$ is the "hard" estimated transmit vector.

- *Hard* IIC: The elements of $\check{\mathbf{s}}$ belong to ψ. In this case, there is no need for the soft-to-hard block in Figure 8.11.
- *Soft* IIC: The elements of $\check{\mathbf{s}}$ do not belong necessarily to ψ.

The superiority of non-linear soft IIC is widely accepted.

8.5.1.3 Recurrent neural networks

This class of neural networks has been attracting a lot of interest because of their widespread applications. They can be either trained to approximate a MIMO system ([GM04]; system identification), or they can be considered as dynamical systems. In the later case, one of the most important properties of these networks is their ability to solve optimisation problems without the need for a training phase, which is desirable in many engineering fields like signal processing, communications, automatic control, etc. In this case, the training process, always associated with computational complexity, time and free parameter optimisation can be avoided. This relies on the ability of these networks (under specific conditions) to be Lyapunov-stable. These conditions have been shown to be fulfilled for iterative interference cancellation-based recurrent neural network (IIC-RNN).

8.5.1.4 Discrete-time RNNs for IIC

The discrete-time IIC-RNN is well known in the literature [ETL+02]. The iterative process of the discrete-time IIC-RNN is described as

$$\begin{aligned}\mathbf{u}[\kappa] &= [-\mathbf{D}^{-1}\cdot\mathbf{C}]\cdot\check{\mathbf{s}}[\kappa-1]+\mathbf{D}^{-1}\cdot\mathbf{r},\\ \check{\mathbf{s}}[\kappa] &= \boldsymbol{\theta}(\mathbf{u}[\kappa]).\end{aligned} \tag{8.8}$$

$\boldsymbol{\theta}(\cdot)$ is the non-linear optimum estimation function given for each component as [Eng03]:

$$\theta(u)=\frac{\sum_{j=1}^{M}\psi_j\cdot\exp\left\{-\frac{1}{2}\cdot(\beta_r\cdot\psi_{j,r}^2+\beta_i\cdot\psi_{j,i}^2)\right.}{\sum_{j=1}^{M}\exp\left\{-\frac{1}{2}\cdot(\beta_r\cdot\psi_{j,r}^2+\beta_r\cdot\psi_{j,r})\right.}\frac{\left.+\beta_r\cdot\psi_{j,r}\cdot u_r+\beta_i\cdot\psi_{j,i}\cdot u_i\right\}}{\left.+\beta_r\cdot\psi_{j,r}\cdot u_r+\beta_i\cdot\psi_{j,i}\cdot u_i\right\}} \tag{8.9}$$

The subscripts *r*, *i* indicate the real and imaginary part of a complex-valued variable, respectively. β_r, β_i are positive free parameters, usually inversely proportional to the variance of the noise.

Successive over relaxation (SOR) is a well-known method to improve the convergence of linear iteration processes. SOR has been applied for the IIC-RNN Equation (8.8) in Mostafa et al. [MTL10, TW12]. Particularly in Teich and Wallner [TW12] SOR for discrete-time IIC-RNN has been applied for digital transmission schemes based on multiple sets of orthogonal spreading codes. It has been shown that the bandwidth efficiency can be increased by up to 50% without loss in power efficiency. Resulting interference between different non-orthogonal codes can be combated by IIC-RNN. By applying SOR to IIC-RNN the bandwidth efficiency can be increased even a bit further (without loss of power efficiency). Even more, the complexity of IIC-RNN (number of iterations) could be reduced substantially by the SOR.

The local asymptotical stability of the discrete-time IIC-RNN with/without SOR has been proven in the sense of Lyapunov in Mostafa et al. [MTL14c].

8.5.1.5 Continuous-time RNNs for IIC

For high data rates, the power consumption of iterative detection algorithms (equalization and channel decoding) is expected to become a limiting factor. The growing demand for jointly high data rate transmission and power efficient detection revives the analogue implementation option. Continuous-time RNNs serve as promising computational models for analogue hardware implementation [CU93]. The continuous-time IIC-RNN is described by the following differential equation

$$\begin{aligned} \frac{d\mathbf{u}(t)}{d_t} &= -\mathbf{u}(t) + [-\mathbf{D}^{-1} \cdot \mathbf{C}] \cdot \check{\mathbf{s}}(t) + \mathrm{D}^{-1} \cdot \mathbf{r}, \\ \check{\mathbf{s}}(t) &= \boldsymbol{\theta}(\mathbf{u}(t)). \end{aligned} \tag{8.10}$$

First work toward continuous-time IIC-RNN was limited to binary phase shift keying (BPSK) modulation [KM96]. This has been extended to square quadrature amplitude modulation (QAM) in Mostafa et al. [MTL12, Mosl4]. The local asymptotical stability proof in the sense of Lyapunov is given in Kuroe et al. [KHM02] and has been extended in Mostafa [Mosl4].

An initial analogue implementation of the continuous-time IIC-RNN for a BPSK symbol alphabet [OSM+14] showed an improvement of the power-to-speed ratio of up to four orders of magnitude, compared to a standard digital solution, e.g., a low-power Intel XScale central processing unit (CPU): clock frequency 600 MHz, power consumption 0.5 W.

The analogue implementation of the continuous-time IIC-RNN requires an analogue implementation of the optimum estimation function Equation (8.9). It has been proven in Mostafa et al. [MTL14a] that the optimum estimation function for square QAM symbol alphabets can be approximated by a sum of a

limited number of properly shifted and weighted hyperbolic tangent functions, which can be realised in analogue by differential amplifiers. A comparison between the dynamical behaviour of the discrete-time and continuous-time IIC-RNN can be found in Mostafa et al. [MTLllb, MTL12].

8.5.2 Concatenated Systems, Bit-Interleaved Coded Modulation (BICM) and Iterative Detection

The mobile evolution of the second generation digital terrestrial television broadcasting is the scenario considered in Vargas et al. [VWMG12]. digital video broadcasting-next generation handheld (DVB-NGF) is a MIMO system with BICM. Specifically an iterative receiver concept is employed, where the MIMO demapper and the channel decoder iteratively exchange extrinsic information (turbo detection). The focus is on implementation aspects such as suboptimal demodulators (soft MMSE demodulation with *a priori* information from the decoder) and quantisation effects (optimum LLR quantisation). FER simulations were performed for an outdoor mobile scenario with a user velocity of 60 km/h. For a low-code rate of 1/3, the soft MMSE demodulator shows no performance degradation compared to a much more complex max-log reference receiver. For a high-code rate of 11/15, the soft MMSE demodulator suffers from a performance loss of up to 2 dB compared to the max-log receiver. With turbo detection SNR gains between 1 and 2 dB can be achieved.

The system considered in Wu et al. [WAM14] is an interleave division multiple access (IDMA) uplink over frequency selective fading channels in the low-power regime. Performance close to capacity is achieved by using optimised block-interleaved coded modulation-iterative detection (BICM-ID) with a very low-rate single parity check code concatenated with an irregular repetition code together with a doped accumulator and extended mapping. The parameters of the system (code parameters, doping ratio, and modulation mixing ratio) have been optimised for an additive white Gaussian noise (AWGN) channel with an extrinsic information transfer chart (EXIT) constrained binary switching algorithm. As detector, a joint frequency domain turbo equalisation and IDMA signal detection algorithm is proposed. With this optimised system and for frequency selective fading channels, a performance close to the OP can be achieved. For a multiuser scenario, the detector complexity can be reduced substantially by implementing detection ordering according to the received power.

He et al. [HZAM13] considered a parallel wireless sensor network (WSN). A non-negative constrained iterative algorithm is presented to estimate the

observation error probabilities for WSN with an arbitrary number of sensor nodes. The idea is to exploit the correlation among the observed data from different sensor nodes, representing the correlation of the links between the sensing object and the sensor nodes. To utilise these correlations, a multi-dimensional iterative receiver is proposed. The estimated observation error probabilities expressed as LLR are exchanged between the decoders (global iteration). As usual, each decoder performs an extrinsic LLR exchange (local iteration). For an AWGN channel Monte Carlo simulations show, that for large sensor networks the required SNR to achieve a given error rate can be improved by up to 9 dB by using a sufficient number of global iterations. For a Rayleigh-fading channel a SNR improvement of about 4 dB is obtained at a BER of 10^{-3}.

Az et al. [AAM13] also considered WSN. A detection technique based on factor graphs is proposed to estimate the position of illegal radio in a wireless system. To improve the geo-location estimation, received signal strength (RSS)- and DOA-based factor graphs are combined in a single factor graph with empirically found weighting factors. The performance of DOA is improved by introducing a modified variance approximation of the target location based on a Taylor series expansion of the tangent functions.

8.5.3 Advanced Coding Schemes and Iterative Decoding

An analogue realisation of iterative decoding algorithm based on high-order RRNs is the work presented in Mostafa et al. [MTLlla] and [MTL14b]. Besides a large decoding speed, the main advantages of an analogue very large-scale integration (VLSI) implementation are the improved EE (higher power-to-speed ratio) and, going hand in hand with that, the reduced area consumption compared to a digital VLSI implementation. This is similar as in the case of an analogue realisation of IIC (see Section 8.5.1). Starting from the iterative threshold decoding algorithm, Mostafa et al. [MTLlla] derives the specific structure and parameters so that a high-order RNN can serve as an analogue decoder:

- each code symbol is represented by one neuron;
- the *L*-value of each code symbol is represented by the inner state of the neuron;
- each neuron has an external input given by the intrinsic *L*-value of the corresponding code symbol;
- the non-linearity of the neuron is a tangent hyperbolic function with slope equal to 0.5;
- the code structure is represented in the non-linear feedback matrix and the non-linear feedback vector function of the network

The equation of motion of the resulting continuous-time dynamical system is given by a non-linear system of first-order differential equations. For a specific example (tail-biting convolutional code with memory one), simulation results show, that the BER performance of the analogue decoder (continuous-time dynamical system) is equivalent to the one of the digital counterpart (discrete-time dynamical system) and even matches ML performance.

In Mostafa et al. [MTL14b], the iterative decoding algorithm *belief propagation* and *iterative threshold decoding* are described as dynamical systems, where both, discrete-time and continuous-time systems are considered. Specifications for the stability of fix points or equilibrium points are given for the linear as well as the non-linear case. Furthermore, conditions are derived, so that the continuous-time dynamical system shares the same (stable) equilibrium points as the discrete-time dynamical system (twin dynamical system). This eventually allows to represent an iterative decoding algorithm (discrete-time dynamical system) by an analogue VLSI realisation. For repetition codes the dynamics of belief propagation and iterative threshold decoding is linear and a closed form solution for the stability of the equilibrium/fixed points is given.

Hu et al. [HKD11] considered the IEEE 802.11a WiFi system. Based on a modification of the BCJR algorithm the authors propose to improve the forward error correction (FEC) decoding in the PHY by exploiting known or predictable symbols from higher open system interconnection (OSI) layers. For the MAC header error rate an improved power efficiency of 1.5 dB was obtained. However, for the overall MAC FER the gain reduced to 0.35 dB. Introducing an additional interleaving to better distribute the known symbols, the gain could be increased to about 1 dB.

A theoretical performance analysis of a code based on the concatenation of a low-density parity check (LPDC) code and a rateless code is given in Botos and Bota [BB12]. The approximate analysis of various parameters is performed for a transmission over a Rayleigh-fading channel. The performance metrics considered are the probability of message non-recovery at a given SNR or the minimum SNR required to ensure an imposed probability of message non-recovery, the spectral efficiencies, and delay (time required to successfully transmit a message). Parameter of the evaluation are the message length and overhead and depth of the rateless code.

8.5.4 Non-Coherent Detection

In Zhu and Burr [ZB12], the authors consider a serially concatenated system composed of an (outer) recursive systematic convolutional (RSC) code concatenated with DPSK modulation. To compensate for the loss due to

non-coherent detection, Zhu and Burr [ZB12] propose to employ a turbo receiver. To do so they view the differential encoder of the DPSK modulator as a rate one recursive non-systematic convolutional encoder (accumulator), which can be represented by a two-state trellis. In the iterative receiver, soft information is exchanged between the phase shift keying (PSK) soft demapper, the differential decoder of the inner code and the RSC decoder of the outer code. Simulation results are given for a correlated Rayleigh-fading channel. Depending on the number of iterations the power efficiency could be improved by up to 5 dB, compared to a coded PSK scheme with non-iterative detection. Compared to the case of having perfect CSI at the receiver, introducing a simple *a posteriori* probability (APP) channel estimator leads to a performance degradation of 0.5–1 dB.

Peiker-Feil et al. [PFWTL12] and [PFTL13] also consider a non-coherent transmission scheme. OFDM-MFSK has been proposed as a robust transmission scheme suitable for (fast) time-varying and frequency selective channel [Wet 11]. OFDM is concerned with the frequency-selective behaviour of the channel, whereas M-ary frequency shift keying (MFSK) allows an incoherent detection and is thus able to cope with a (fast) time-variance of the channel. Peiker et al. [PFWTL12] show that OFDM-MFSK can be described as a non-coherent transmission based on vector sub-spaces. This gives new insight in the design and analysis of advanced non-coherent transmission schemes based on OFDM-MFSK [PF14].

One of the major disadvantages of OFDM-MFSK is its poor bandwidth-efficiency. This problem is addressed in Peiker-Feil et al. [PFTL13]. Combining MFSK with multi-tone frequency shift keying (FSK) the bandwidth efficiency of OFDM-MFSK could be increased by 50%, compared to classical OFDM-MFSK. At the same time, the loss in power efficiency in an AWGN channel was limited to about 1.5 dB.

8.6 Link Abstraction

8.6.1 On the Need for Link Abstraction

System-level simulators of mobile communication systems often rely on simplified look-up-tables (LUTs) generated offline by a link-level simulator. Link abstraction techniques aim at obtaining LUTs to predict the block error rate (BLER) for multistate channels. A multistate channel arises when the received LLRs, within a given codeword, show very different reliabilities due to one or several reasons such as: (i) the frequency selective fading in OFDM, (ii) the LLR combination prior to hybrid ARQ (H-ARQ) decoding,

and (iii) the unequal error protection with high order modulations in BICM. BLER prediction is needed at the UE and at the BS to decide about the most suitable modulation and code rate to apply for next transmission time interval (TTI) and also for large scale system evaluations.

8.6.2 Truncated Shannon Bound

The Shannon bound gives an upper limit on the throughput of a communication link. The state of the art in modulation and coding schemes is a few dB from this bound, since the Shannon bound can be reached only in the limit of infinite block length and decoding complexity. The throughput of practical modulation and coding schemes used in wireless standards, like LTE/LTE-A and worldwide interoperability for microwave access (WiMAX), can be approximated by a function similar to the Shannon bound, but shifted to the right and/or compressed on the vertical scale by a scaling factor less than 1.

Taking into account that in a real system the capacity is zero below a minimum SNR threshold (γ_0) and is upper bounded by the capacity of the highest MCS ($C_{\max}$) for SNR higher than $\gamma_{\max,}$ the expression for the truncated Shannon bound (TSB) throughput is [Bur11]:

$$C_{\mathrm{TSB}}(\gamma) = \begin{cases} 0 & \gamma < \gamma_0 \\ W\alpha \log_2(1+\gamma/\gamma_{\mathrm{sh}}) & \gamma_0 \leq \gamma < \gamma_{\max} \\ C_{\max} & \gamma_{\max} < \gamma \end{cases} \tag{8.11}$$

where α and γ_{sh} must be adjusted to optimally approximate the actual throughput function.

Given that the SNR is random, the condition to apply to find α and γ_{sh} is that the average throughput given by the TSB should be as close as possible to the average throughput given by the true throughput function ($C_{\mathrm{thr}}(\gamma)$), which depends on the specific SNR thresholds of each MCS and its corresponding throughputs. The average throughputs are computed by averaging $C_{\mathrm{TSB}}(\gamma)$ (resp. $C_{\mathrm{thr}}(\gamma)$) over all possible values of the SNR.

In a MIMO link, we must also take into account the random nature of the channel matrix and the fact that the direction of the interference is relevant [Bur03]. To obtain the average throughput of a MIMO link the TSB can be applied on a stream basis (given the cumulative distribution function (CDF) of the SNR experienced on each stream). In general, the CDF of the stream SNR has to be estimated using link-level Monte Carlo simulation.

8.6.2.1 Link abstraction for multi-user MIMO (MU-MIMO)

Assuming an OFDM system with N subcarriers and a SINR vector $\{\gamma_1, \gamma_2, ..., \gamma_N\}$, the received bit information rate (RBIR) is defined as the average information capacity of a coded bit in the received codeword, i.e.:

$$\mathrm{RBIR} = \frac{1}{\beta} \frac{\sum_{\kappa=1}^{N} SI(\gamma_\kappa)}{\sum_{\kappa=1}^{N} M_\kappa} \ (\mathrm{bits/coded\ bit}), \tag{8.12}$$

where $\mathrm{SI}(\gamma_\kappa)$ and M_κ are, respectively, the mutual information (MI) at modulation symbol level and the modulation order of subcarrier κ, and β is a parameter that accounts for the capacity loss of a practical code.

In a MU-MIMO link, the function $\mathrm{SI}(\cdot)$ in Equation (8.12) can include the known interference statistics in the computation of the MI at modulation symbol level, thus modelling the performance of an "interference aware" receiver [LKK11].

From the RBIR the effective signal-to-noise ratio (ESNR), defined as the SNR that would give the same RBIR under AWGN channel [BAS+05] can be obtained. The ESNR allows BLER prediction using AWGN BLER LUTs.

8.6.2.2 Link abstraction for H-ARQ

In LTE, incremental redundancy (IR) H-ARQ each newly received redundancy version (RV) is soft-combined with the previous ones to increase the decoding probability. The key issue is to find an approximated method to compute the RBIR at the turbo decoder taking into account that any coded bit may have been received more than once and that in each transmission it may have experienced a different quality, due to the unequal error protection of 16 and 64 QAM and to the changing SNR conditions of the different subcarriers and H-ARQ RVs.

Let's assume that a sample coded bit is transmitted exactly twice using 16QAM in two different subcarriers with $\mathrm{SNR}_1 = \xi_1$ and $\mathrm{SNR}_2 = \xi_2$. In 16QAM the bit to symbol mapping creates two different channel qualities: A and B. Assuming, for example, that the first transmission uses bit channel A and the second transmission uses bit channel B, a possible procedure to obtain the RBIR contribution of the sample coded bit is [OSRL12]:

$$\mathrm{RBIR}(\xi_1, \xi_2) = I[I^{-1}(\mathrm{RBIR_A}(\xi_1)) + I^{-1}(\mathrm{RBIR_B}(\xi_2))], \tag{8.13}$$

where $I(\gamma) = 1 - exp(-\gamma/10^{-0.12})$ approximates the AWGN RBIR for BPSK and $\mathrm{RBIR_A}$ (ξ_1) and $\mathrm{RBIR_B}$ (ξ_2) are the AWGN RBIR of the 16QAM bit channels. Since both the subcarrier SNRs and the 16QAM channel quality are random, only the average value of Equation (8.13) can be computed.

The proposed complete methodology is to classify all the coded bits into disjoint subsets based on the number of times that the bit has been received inside each of the RVs. For each of these subsets we should know the set $\{\xi\}$ of SNRs and the cardinal of the subset relative to the codeword size (called "repetition factor"). The repetition factors may be computed off-line (only once) for each MCS and stored in a LUT. Finally, the explained methodology for the case of two transmissions is extrapolated and applied to compute the contribution of each subset to the global RBIR and the contributions are added together after being weighted by the repetition factors.

In order to simplify the computations for H-ARQ link abstraction it is also possible to compute a global ESNR for the whole set of received RVs [LKKO12], using:

$$\gamma_{\text{eff}} = \frac{r_m}{r_{\text{eff}}} \cdot \text{SI}^{-1} \left[\frac{1}{J \cdot M} \sum_{j=1}^{J} \sum_{m=1}^{M} \text{SI}(\gamma_{m,j}) \right] \tag{8.14}$$

where *J* is the number of subcarriers, *M* is the number of received RVs, $\gamma_{m,j}$ - is the SNR of subcarrier *j* at RV m, r_m is the mother code rate (r_m = 1/3 for LTE) and r_{eff} is the effective code rate defined as the number of information bits divided by the total number of received coded bits (including repetitions) per codeword. The predicted BLER is then given by the reference AWGN BLER curve for the modulation in use and code rate r_m.

8.6.2.3 A system level simulator applying link abstraction

In Litjens et al. [LZN+13] the system level performances of non-orthogonal spectrum sharing (NOSS), based on transmit beamforming, are compared to those of orthogonal spectrum sharing (OSS) and fixed spectrum assignment (FSA). The ESNR concept is used to model the beamforming SINR performance and to aggregate a set of physical resource block (PRB)-specific SINRs to a single effective SINR which is then used to predict the user throughput. When transmit beamforming is in use two UEs are simultaneously served from two cooperating co-located BSs at the same PRB. In this case the link abstraction LUT provides an ESNR for each UE which depends on the beamforming technique being applied (SR, nash bargaining (NB), or zero forcing (ZF)).