Layered Design of Hierarchical Exclusive Codebook and Its Capacity Regions for HDF Strategy in Parametric Wireless 2-WRC

Jan Sykora, Alister Burr

***Abstract*—The paper addresses a Hierarchical Decode and Forward (HDF) strategy in the wireless 2-Way Relay Channel (2-WRC). This strategy uses a Hierarchical eXclusive Code (HXC) that allows full decoding of the hierarchical symbols at the relay. The HXC represents two data sources only through the exclusive law and requires side information on the complementary data at the destination (which naturally holds for the 2-WRC). The HDF strategy has the advantage over classical MAC stage relaying with joint decoding that its rate region extends beyond the classical MAC region. We present a layered design of the HXC codebook which uses an arbitrary outer state-of-the-art capacity approaching code (e.g. LDPC) and an inner layer with an exclusive symbol alphabet. We provide basic theorems showing that this scheme forms an HXC and we also evaluate its alphabet constrained rate regions. The rate regions depend on the relative channel phase parameters. Some channel rotation leads to catastrophic violation of the exclusive law. However these values appear only in a limited range of phases and have only mild impact on the mean capacity for some component symbol constellations.**

I. Introduction

A. Background And Motivation

Multi-node and multi-source wireless communication scenarios are currently under intensive investigation in the research community. Generally, these can be seen as similar to the Network Coding (NC) paradigm [1]. NC operates with a discrete (typical binary) alphabet over lossless discrete channels. It is in fact an operation on data rather than on the channel codewords. NC has great potential in substantially increasing the throughput of complicated communication networks. An extension of these principles into the wireless (signal space) domain is however non-trivial. Some attempts have been carried out using a simple concatenation of NC and a single-link physical layer modulation and coding technique. In [2], the authors construct various relaying schemes with complex field NC. The operation on the relay however requires individual decoding of incoming MAC phase symbols, i.e. the NC is built on top of the classical PHY layer technique. This has number of drawbacks and only limited optimality. An optimal solution is a direct signal space domain code synthesis.

B. Goals of this Paper and Contributions

We present a relaying strategy for the 2-Way Relay Channel (2-WRC) (also known as TWRC) based on Hierarchical eXclusive Code (HXC) relay processing. Hierarchical relay processing handles hierarchical data symbols which uniquely represent the individual data from source A and B only by providing side information on the complementary data at the final destination. We call this strategy a Hierarchical Decode and Forward (HDF) strategy. A reason for using the name *hierarchical* is that the relay decodes symbols hierarchically composed from the original two source symbols. The relay does not care about these individual symbols and treats them as one container. In a more complicated network topology, this encapsulation would occur in hierarchical levels.

An advantage of the hierarchical relay processing is its higher achievable rate region, extending outside the rate region of the classical MAC channel that would correspond to individual data stream decoding at the relay. A direct synthesis of the proper hierarchical exclusive code suitable for such processing appears to be too complex. Our paper offers a tractable low complexity solution.

The paper provides the following results and contributions.

1) We propose a *layered* hierarchical exclusive code design which combines a hierarchical inner symbol mapping layer and an outer ordinary single-user-link capacity approaching code (e.g. LDPC or turbo code).
2) We provide basic theorems showing that the layered scheme forms a *hierarchical exclusive code*.
3) We analyze the alphabet limited *achievable rate region* for the hierarchical MAC stage of the relaying and compare it to the alphabet limited and unconstrained cut-set bound capacity limits.
4) We analyze the impact of the *channel parameter values* on the above stated capacities. In this paper the parameters of interest are the channel fading coefficients.

IEEE-T-VT, F-D8 (14.6.2011)
Czech Technical University in Prague, jan.sykora@fel.cvut.cz
University of York, agb1@ohm.york.ac.uk
This work was supported by the European Science Foudation through COST Action 2100, FP7-ICT SAPHYRE project, the Grant Agency of the Czech Republic, grant 102/09/1624, and the Ministry of education, youth and sports of the Czech Republic, prog. MSM6840770014, and grant OC188.
Authors would like also to express an acknowledgement to Tad Matsumoto (Japan Advanced Institute of Science and Technology, Japan) for numerous discussions and collaboration on the subject.

C. Related Work

Limited code design and capacity region results are available for the simplest possible scenario of the 2-WRC. The authors of [3], [4], [5], [6] essentially follow the symbol-wise hierarchical exclusive constellation optimization. Another related area is lattice code based hierarchical code construction. It is followed by [7] and [8]. The authors of [7] and [8] provide lattice based code construction using the principles of [9] but do not investigate the impact of the channel parameters. The lattice based constructions have a drawback in overall complexity since they treat both exclusivity law and capacity approaching code performance in one complicated design. The lattice based approach also treats the problem in a rather conceptual and abstract manner in comparison with the practically implementable alphabet and code constructions treated in this paper. Nevertheless, the lattice based approach provides a direct connection to the pure information-theoretic investigations. A number of results related to the distributed coding and processing are available. Some authors ([10], [11]) approach a similar set of problems by a strategy called Compute and Forward which relies again on structured codes with a lattice based code construction ([9]). To the best of our knowledge, these are the only significant previous works on the 2-WRC respecting a true signal space nature of wireless communications for the hierarchical relay processing. There are also some other weakly related works using simply a traditional form of NC on top of the standard physical layer technique (like classical 2-source MAC), e.g. [12].

The authors of [13] address the problem of 2-WRC Physical-layer Network Coding design. They restrict themselves to a binary BPSK alphabet and most importantly by the non-parametric channel (no phase rotation). They encode the source nodes by the binary Repeat and Accumulate code and then they construct a Factor Graph SPA based decoder for virtual arithmetically superpositioned relay received symbols. The factor check node for the accumulator output is incorporated inside the corresponding mappings for the discrete Network Coded relay output symbol. This structure however strongly depends on the non-parametric channel assumption (it affects the virtual accumulator factor node update rules) and the usage of the repeater as the outer stage. The overall resulting scheme of [13] is limited by the alphabet (in this case BPSK) constrained hierarchical capacity, whereas our layered design provides a generic procedure for achieving the same target with an explicit rate region calculation.

The authors of [14] approach the problem of the code design for 2-WRC along the same lines as [7] and [8] using the lattice based design. It stands on the principles of [9] modified for the modulo sum of nested lattices. The design is not finite alphabet limited. On the other hand, the lattice based approach makes the inclusion of the joint multisource-relay channel parameters extremely difficult. The modulo-lattice operations would collapse under variable channel parameters. A positive feature of the lattice approach is that it provides, at least theoretically (but see also [15] for some recent developments), a tool to prove the achievability of the general Gaussian alphabet unconstrained channel rates. In [14], it is shown that the joint bi-directional achievable rate is $\frac{1}{2}\lg(\frac{1}{2}+\mathrm{SNR})$ for real valued lattice 2-WRC codes.

The authors of [16] address the scenario of Detect and Forward (symbol level detection at the relay, no codeword decoding) combined with discrete NC. They reveal that nonlinear NC maps can provide better throughput than the linear maps under the low signal to noise ratio regime. The nonlinear map helps to overcome unreliable symbol decisions at the relay.

Papers [6], [4] appear to be the closest related works. The papers treat the problem of designing the optimized constellation exclusive mapping regions for the relay. The authors take the pair-wise error probability (and hence free distance) as the optimization target. They reach the conclusion that, under a specific channel state, the minimal exclusive mapping fails to comply with the exclusive law (or gives poor performance in that region of the channel states). They suggest a solution based on adaptive change of the mapping regions and also allowing extended (higher cardinality) exclusive mapping. The work however leaves a number of unsolved or open problems and also adopts a number of ad-hoc assumptions. (1) The optimization target is the *per-symbol* free distance and the optimization does not respect multiplicity of the distances. The procedure is thus in fact strongly constrained to the uncoded case. (2) The problem of the cardinality of the exclusive mapping at the relay is treated in an isolated and ad-hoc manner. (3) The fact that some adaptive mapping modes use extended (non-minimal) mapping prevents the usage of the outer code layered design.

Our paper treats the above stated points in a different way. (1) We treat the problem in terms of the achievable rate region for coded communications. We develop a layered channel coding theorem allowing under precisely given conditions to apply standard channel coding. (2) We systematically define the relay exclusive mappings as minimal, extended and classical and we show how these affect individual parts of the system. (3) We evaluate true soft-output channel entropies. This leads to the achievable rate regions that can be closely approached with current state-of-the-art codes (turbo, LDPC). (4) We show that the layered code design is possible for the minimal mapping and allows the use of a common outer code with inner exclusive alphabet. The resulting scheme produces a correct exclusive end-to-end code.

II. System Model and Definitions

We consider a wireless 2-WRC system (Fig. 1) which has 3 physically separated nodes (source S, relay R, destination D) supporting two way communication through a common shared relay R. The source for data A is co-located with the destination for data B and vice-versa. The transmitted data (signal) of the source serves as Complementary Side-Information (C-SI) for the destination of the *reverse* link (hence the name Complementary). The system is wireless and all transmitted and received symbols are signal space symbols.

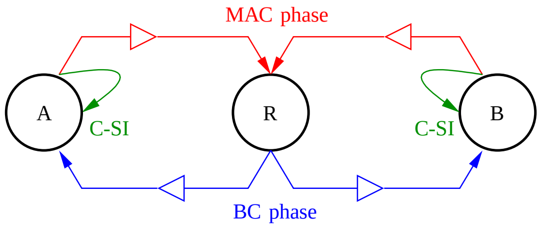

Figure 1. 2WRC with C-SI.

The channel model is linear frequency-flat with Additive White Gaussian Noise (AWGN). The signal transmission is half-duplex (the co-located transmitter (Tx) and receiver (Rx) are not allowed to operate simultaneously). The operation is split into a Multiple Access (MAC) and a Broadcast (BC) phase. The system is symmetric in its topology. This allows us to investigate it from the perspective of the data flow A with data B being only a nuisance parameter. All conclusions drawn for A should then be equally applicable to B. All receivers are assumed to have channel state information. As in some other papers ([6], [4]) the channels A-R and R-B are assumed to be symmetric. In the asymmetric case, the throughput would be given by the worse channel bottleneck. We do not however pursue this case in this paper.

A. MAC Phase

Now, we define all formal details. Subscripts A and B denotes the variables associated with node A and B respectively. Source data messages are $\mathbf{d}_A$, $\mathbf{d}_B$ and they are composed of data symbols $d_A, d_B \in \mathscr{A}_d = \{0, 1, \ldots, M_d - 1\}$, with alphabet cardinality $|\mathscr{A}_d| = M_d$. For notational simplicity, we omit the sequence number indices of individual symbols. Source node codewords are $\mathbf{c}_A, \mathbf{c}_B$ with code symbols $c_A, c_B \in \mathscr{A}_c$, $|\mathscr{A}_c| = M_c$. The encoding operation is performed by the encoders $\mathscr{C}_A, \mathscr{C}_B$ with codebooks $\mathbf{c}_A \in \mathscr{C}_A$ and $\mathbf{c}_B \in \mathscr{C}_B$. A signal space representation (with an *orthonormal* basis) of the transmitted channel symbols is $s_A = s(c_A)$, $s_B = s(c_B)$, $s_A, s_B \in \mathscr{A}_s \subset \mathbb{C}^N$. We assume a common channel symbol mapper $\mathscr{A}_s(\bullet)$. A signal space representation of the overall coded frame is $\mathbf{s}_A(\mathbf{c}_A)$ and $\mathbf{s}_B(\mathbf{c}_B)$. All vectors $\mathbf{d}_A, \mathbf{d}_B, \mathbf{c}_A, \mathbf{c}_B$ have size equal to the frame (codeword) length.

We use a slightly relaxed notation for the symbol and the codebook sets and the corresponding mapping and encoding operations. The notation uses the same letter symbol in two different contexts. This simplifies the reader's appreciation of the relatively large number of codebooks/alphabets involved. As an example, the codebook (the set of all codewords) is denoted by $\mathscr{C}$ and it is optionally supplemented by a subscript corresponding to the particular codeword. The notation $\mathscr{C}(\mathbf{d})$ denotes a codeword corresponding to the data input $\mathbf{d}$. A similar notation holds for symbol alphabets, denoted by $\mathscr{A}$, and corresponding mapping operations $\mathscr{A}(\bullet)$.

The useful part of the received signal at the relay is

$$u = s_A + h s_B. \tag{1}$$

It equivalently represents the parametric channel with both links parametrized according to flat fading assumed to be constant over the frame. It is obtained by a proper common rescaling of the true channel $x' = h_A s_A + h_B s_B + w'$ by $1/h_A$ and denoting $x = x'/h_A$, $h = h_B/h_A$, $w = w'/h_A$, $h_A, h_B, h \in \mathbb{C}^1$. The received signal of the equivalent channel at the relay is

$$x = u + w \tag{2}$$

where the circularly symmetric complex Gaussian noise has the variance σ_w^2 per complex dimension.

The Signal-to-Noise Ratio (SNR) is defined as the ratio of the real base-band symbol energy of one source (e.g. A, to have a fair comparison for reference cases) to the noise power spectrum density ratio $\gamma_x = \left(\bar{\mathscr{E}}_{s_A}/2\right)/N_{0w}$. Assuming orthonormal basis signal space complex envelope representation of the AWGN, we have $\sigma_w^2 = 2N_{0w}$ and thus $\gamma_x = \mathrm{E}\left[\|\mathbf{s}_A\|^2\right]/\sigma_w^2$.

B. BC Phase

The relay receives the signal x and processes it using a Hierarchical Decode and Forward (HDF) strategy. More details will be given in the next section. The output codeword and its code symbols are $\mathbf{c}_R$ and c_R. These are mapped into signal space channel symbols $v \in \mathscr{A}_R$ and signal space codewords $\mathbf{v}$ with the codebook $\mathbf{v} \in \mathscr{C}_R$ and broadcast to destinations A and B. At node B (the destination for data A), the received signal space symbols are

$$y_A = v + w_A \tag{3}$$

where the complex circularly symmetric AWGN w_A has variance σ_A^2 per complex dimension. We denote the signal space symbols at the node A (a destination for data B) similarly $y_B = v + w_B$. The SNR is defined as $\gamma_y = \mathrm{E}\left[\|\mathbf{v}\|^2\right]/\sigma_A^2$. Phase rotations of individual gains of R-DA and R-DB channels cannot affect the throughput regions. We assume a symmetric case for the R-DA and R-DB links.

III. Layered Hierarchical Exclusive Codebook Design

A. Network Coded Modulation

Network Coded Modulation (NCM) is the network structure aware modulation and coding. Each node, either source or relay, transmits the signal that can be processed in the remaining receiving nodes performing various joint relaying strategies (e.g. HDF) and utilizing all complementary side-information available at destinations. The NCM term itself (as introduced in our earlier works [17], [18], [19]) denotes a particular joint design of the multi-source codebook which (1) has signal-space codewords (hence the term “Modulation”), and (2) which respects the network structure and the position of the nodes involved in communication including other side-information the nodes might have available (hence the term “Network Coded”). The knowledge of the network structure is an essential part of NCM design (similarly as Trellis Coded Modulation uses the knowledge of the trellis structure).

B. Hierarchical Decode and Forward Strategy

1) Exclusive Law: The HDF strategy is based on relay processing which fully decodes the Hierarchical Data (HD) message $\mathbf{d}_{AB}(\mathbf{d}_A,\mathbf{d}_B)$ and sends out the corresponding codeword $\mathbf{v}=\mathbf{v}(\mathbf{d}_{AB})$ which represents the original data messages $\mathbf{d}_A$ and $\mathbf{d}_B$ only through the *exclusive law* [4]

$$\mathbf{v}(\mathbf{d}_{AB}(\mathbf{d}_A,\mathbf{d}_B)) \neq \mathbf{v}\left(\mathbf{d}_{AB}(\mathbf{d}'_A,\mathbf{d}_B)\right),\ \forall \mathbf{d}_A \neq \mathbf{d}'_A, \quad (4)$$

$$\mathbf{v}(\mathbf{d}_{AB}(\mathbf{d}_A,\mathbf{d}_B)) \neq \mathbf{v}\left(\mathbf{d}_{AB}(\mathbf{d}_A,\mathbf{d}'_B)\right),\ \forall \mathbf{d}_B \neq \mathbf{d}'_B. \quad (5)$$

The hierarchical data are a joint representation of the data from both sources such that it uniquely represents one data source given full knowledge of the other one. Assuming that the destination node B has perfect C-SI on the node's own data $\mathbf{d}_B$ it can then decode the message $\mathbf{d}_A$ (and similarly for node A). Data $\mathbf{d}_B$ will be called *complementary* data from the perspective of the data $\mathbf{d}_A$ operations. The exclusive law at the signal space codeword level implies the same for the HD messages

$$\mathbf{d}_{AB}(\mathbf{d}_A,\mathbf{d}_B) \neq \mathbf{d}_{AB}(\mathbf{d}'_A,\mathbf{d}_B),\ \forall \mathbf{d}_A \neq \mathbf{d}'_A, \quad (6)$$

$$\mathbf{d}_{AB}(\mathbf{d}_A,\mathbf{d}_B) \neq \mathbf{d}_{AB}(\mathbf{d}'_A,\mathbf{d}_B),\ \forall \mathbf{d}_B \neq \mathbf{d}'_B. \quad (7)$$

A code (codebook) satisfying the above criteria is called a Hierarchical eXclusive Code (HXC). Note that the mapping may in principle be linear or non-linear [16]. We will denote the mapping satisfying the exclusive law by the operator $\mathscr{X}(\centerdot,\centerdot)$.

The hierarchical data can be viewed as network coded data streams. But notice an important difference. When we simply concatenate independently the NC and then the channel coding, the data generated at the relay are obtained by jointly decoding both individual streams from A and B. The NC mainly serves for the BC stage transmission. The MAC phase is classical joint decoding. In our hierarchical approach this does not happen. The hierarchical data are directly obtained from the received signal *signal space* observations without the necessity of individual source data decoding. The MAC stage encoding must be such that the observation at the relay allows HXC mapping to the hierarchical data, which may then be encoded into $\mathbf{v}$. In other words, alongside the *complete signal path* (MAC and BC), the coding *must always be HXC* w.r.t. the *hierarchical data*.

2) Relay Hierarchical Codebook Cardinality: In order for the exclusive law to hold, the relay hierarchical codebook cardinality must satisfy $|\mathscr{C}_R| \geq \max(|\mathscr{C}_A|,|\mathscr{C}_B|)$. This guarantees the invertibility of the mapping from the hierarchical symbol and the C-SI. On the other hand, the cardinality is upper-bounded by all possible $\mathbf{d}_A$ and $\mathbf{d}_B$ message combinations $|\mathscr{C}_R| \leq |\mathscr{C}_A||\mathscr{C}_B|$. The lower bound case requires perfect C-SI on the complementary data at the destination. The upper bound is in fact a classical MAC channel with joint decoding of both data messages at the relay and requires *no* C-SI. Any situation in between those two extreme cases requires partial C-SI at the destination.

The HXC having the minimal cardinality $|\mathscr{C}_R| = \max(|\mathscr{C}_A|,|\mathscr{C}_B|)$ will be called Minimal-HXC (M-HXC). This case puts minimum *throughput* requirements on the BC channel at the price of requiring perfect C-SI at the destination. All codebooks with a higher cardinality will be called Extended-HXC (E-HXC) or Non-Minimal-HXC.

C. Layered HXC Design for Perfect C-SI

1) Error Correcting and Hierarchical Exclusive Mapper Layers: It is evident that a direct design of the HXC codebook $\mathscr{C}_R$ providing the mapping $\mathbf{v}(\mathbf{d}_{AB}(\mathbf{d}_A,\mathbf{d}_B))$ is highly complex. An alternative approach is a layered design where the inner layer (closer to the channel symbols) provides the *exclusivity property* whereas the outer layer provides the error correcting (classical channel coding) functionality. The outer layer code can be an arbitrary state-of-the-art capacity approaching code, e.g. turbo code or LDPC. The dividing line between the layers can take various forms. Both layers can have various levels of internal structure (memory) and various levels of error correcting capability.

The simplest case is that in which the inner layer is simply a signal space channel symbol memoryless mapper. An alphabet memoryless mapper

$$c_{AB} = \mathscr{X}_c(c_A,c_B) \quad (8)$$

fulfilling the exclusive law will be called Hierarchical eXclusive Alphabet (HXA). The outer layer then carries all the capacity achieving (error correcting) responsibility. The layered system model is shown in Fig. 2.

2) Symbol-wise Relay HDF Processing: There is also an additional advantage of the layered design. The exclusive property at the symbol level allows a simple determination of the soft *per-symbol* measure decoding metric for hierarchical symbols at the relay. It can either be directly used by the full relay decoder or properly source encoded and sent on a per-symbol basis without decoding. The latter has the major advantage of lowering the latency of the relay processing. The topic of per-symbol *only* relay processing is out of this paper's scope.

3) Layered Design: The layered design relies on the following two lemmas. Source nodes have a common codebook $\mathscr{C}_A = \mathscr{C}_B = \mathscr{C}$.

Lemma 1 (Coding distributes over the exclusive law): Assume arbitrary *linear* one-to-one code mappings with a common codebook

$$\mathbf{c}_A = \mathscr{C}(\mathbf{d}_A),\ \mathbf{c}_B = \mathscr{C}(\mathbf{d}_B),\ \mathbf{c}_{AB} = \mathscr{C}(\mathbf{d}_{AB}) \quad (9)$$

where $\mathbf{d}_A,\mathbf{d}_B,\mathbf{d}_{AB} \in \mathrm{GF}(M^n)$ and $\mathbf{c}_A,\mathbf{c}_B,\mathbf{c}_{AB} \in \mathrm{GF}(M^{\tilde{n}})$, $\tilde{n} > n$. Then, there exist two *minimal* exclusive mappings (for data and codewords)

$$\mathbf{d}_{AB} = \mathscr{X}_{\mathbf{d}}(\mathbf{d}_A,\mathbf{d}_B),\ \mathbf{c}_{AB} = \mathscr{X}_{\mathbf{c}}(\mathbf{c}_A,\mathbf{c}_B) \quad (10)$$

such that the following holds

$$\mathscr{C}\left(\mathscr{X}_{\mathbf{d}}(\mathbf{d}_A,\mathbf{d}_B)\right) = \mathscr{X}_{\mathbf{c}}\left(\mathscr{C}(\mathbf{d}_A),\mathscr{C}(\mathbf{d}_B)\right). \quad (11)$$

This means that encoding hierarchical data is the same as individually encoding the original A and B data and then constructing a hierarchical codeword from the result.

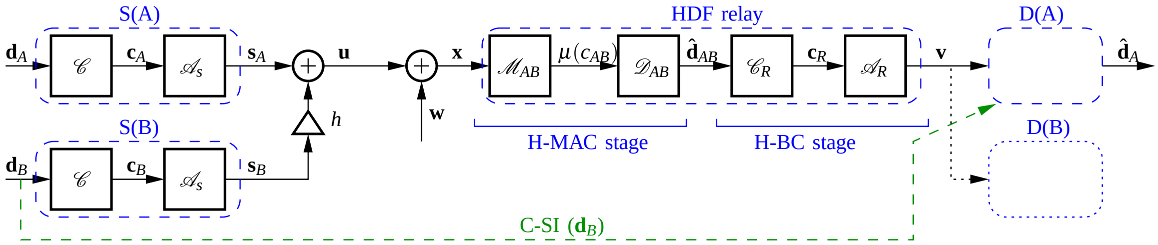

Figure 2. The system model for layered HXC design.

Proof: The minimal hierarchical exclusive mapping for the two data or codewords on GF(M^n) and GF($M^{\tilde{n}}$) respectively with a common alphabet size M can always be done by addition on (with a side effect of being a symbol-by-symbol operation)

$$\mathscr{X}_{\mathbf{d}}(\mathbf{d}_A, \mathbf{d}_B) = \mathbf{d}_A \oplus \mathbf{d}_B, \ \ \mathscr{X}_{\mathbf{c}}(\mathbf{c}_A, \mathbf{c}_B) = \mathbf{c}_A \oplus \mathbf{c}_B. \quad (12)$$

The $\oplus$ operation is symbol-wise over data or code symbol alphabet $GF(M)$. Then, we use the linearity of the code

$$\begin{aligned} \mathscr{C}\left(\mathscr{X}_{\mathbf{d}}(\mathbf{d}_A, \mathbf{d}_B)\right) &= \mathscr{C}\left(\mathbf{d}_A \oplus \mathbf{d}_B\right) \\ &= \mathscr{C}(\mathbf{d}_A) \oplus \mathscr{C}(\mathbf{d}_B) \\ &= \mathscr{X}_{\mathbf{c}}\left(\mathscr{C}(\mathbf{d}_A), \mathscr{C}(\mathbf{d}_B)\right). \end{aligned} \quad (13)$$

■

Corollary 1: A direct consequence of the code distributive law is the ability to decode hierarchical data from the hierarchical codeword. This is exactly what the HDF relay does.

Lemma 2 (Exclusive law decomposition over symbols): Assume that the symbol mapping obeys the exclusive law for each individual symbol. Then the exclusive law holds also for the complete codeword

$$c_{AB} = \mathscr{X}_c(c_A, c_B) \implies \mathbf{c}_{AB} = \mathscr{X}_{\mathbf{c}}(\mathbf{c}_A, \mathbf{c}_B). \quad (14)$$

Proof: Codewords differ $\mathbf{c}_A \neq \mathbf{c}'_A$ if they differ in at least one symbol $c_A \neq c'_A$. Then the exclusive symbols c_{AB} must differ $c_{AB} \neq c'_{AB}$ and thus also $\mathbf{c}_{AB}$ differs $\mathbf{c}_{AB} \neq \mathbf{c}'_{AB}$. Hence the GF addition construction of the exclusive codeword may be done on a symbol-by-symbol basis. ■

We design the layered HXC based on these two lemmas. The MAC and BC phases of the HDF strategy will be denoted Hierarchical-MAC (H-MAC) and Hierarchical-BC (H-BC). They operate, unlike classical MAC/BC, with *hierarchical* symbols/codewords at the relay. The goal is to ensure that the relay output obeys the exclusive law w.r.t. data $\mathbf{v} = \mathscr{X}_{\mathbf{v}}(\mathbf{d}_A, \mathbf{d}_B)$. The relay symbol space output $\mathbf{v}$ is uniquely given by $\mathbf{c}_R$, on a symbol-by-symbol basis $v = \mathscr{A}_R(c_R)$.

Theorem 1 (Layered HXC Design): Assume the following:

1) There is a common outer layer codebook $\mathbf{c}_A = \mathscr{C}(\mathbf{d}_A), \mathbf{c}_B = \mathscr{C}(\mathbf{d}_B)$ for source nodes A and B, and the data forms a *minimal*[1] hierarchical exclusive mapping $\mathbf{d}_{AB} = \mathscr{X}_{\mathbf{d}}(\mathbf{d}_A, \mathbf{d}_B)$.
2) The symbol-by-symbol inner layer mapper has an exclusive mapping w.r.t. the useful signal at the relay

$$c_{AB} = \mathscr{X}_c(c_A, c_B) \quad (15)$$

where $u(c_A, c_B) = \mathscr{A}_u(c_{AB})$. The useful signal can have multiple instances corresponding to one hierarchical symbol. The $\mathscr{A}_u(c_{AB})$ is generally a *set* or a multi-symbol *class*.
3) The information outer code rate obeys

$$R_A = R_B = R_{AB} \leq I(c_{AB}; x). \quad (16)$$

Then the relay H-MAC stage decoder $\mathscr{D}_{AB}$ output is $\hat{\mathbf{d}}_{AB} = \mathbf{d}_{AB}$ and the H-BC stage output fulfills the exclusive law

$$\mathbf{v} = \mathscr{X}_{\mathbf{v}}(\mathbf{d}_A, \mathbf{d}_B). \quad (17)$$

Proof: Lemma 2 implies $\mathbf{c}_{AB} = \mathscr{X}_{\mathbf{c}}(\mathbf{c}_A, \mathbf{c}_B)$. This together with Lemma 1 gives $\mathbf{c}_{AB} = \mathscr{C}(\mathbf{d}_{AB})$. The hierarchical code symbols are observed at the relay by $x = \mathscr{A}_u(c_{AB}) + w$. Providing that the code information rate on the symbols c_{AB} does not exceed $I(c_{AB}; x)$ the H-MAC stage relay decoder $\mathscr{D}_{AB}$ can decode the symbols d_{AB} reliably and thus $\hat{d}_{AB} = d_{AB}$. The data $\mathbf{d}_{AB}$ represents the minimal exclusive mapping and thus fully represents both data streams $\mathbf{d}_A, \mathbf{d}_B$ providing the C-SI is available. This gives a rectangular rate region (16). The rate is achievable since it is known from the literature that there exist *linear* capacity approaching single-source single-destination codes. Under the assumption that the individual source code symbols c_A, c_B are uniformly distributed and the mapping is minimal, the distribution of c_{AB} is also uniform. The H-BC relay encoder encodes this information in a one-to-one manner into $\mathbf{v} = \mathbf{v}(\mathbf{d}_{AB})$ which gives (17). ■

The theorem states that we can achieve capacity by using ordinary capacity approaching codes on top of HXA symbols and the only bottleneck is the finite cardinality of the H-MAC channel hierarchical symbols. A graphical demonstration of the theorem is in Fig. 3.

It is useful to realize that the exclusivity requirements are important only on data $\mathbf{d}_A, \mathbf{d}_B$ and codeword level $\mathbf{c}_A, \mathbf{c}_B$. The symbol space mapping itself does *not* necessarily need to produce disjoint sets $\mathscr{A}_u(c_{AB}) \cap \mathscr{A}_u(c'_{AB}) = \emptyset$, for $c_{AB} \neq c'_{AB}$. A non-zero intersection would of course decrease the value $I(c_{AB}; x)$ but (16) should still hold. The zero set/class overlap is in fact just an exclusivity mapping extension for sets (symbol class output).

[1] Note that the requirement for a minimal mapping prevents this layered approach from being applied to the mapping of [6], since for some channel states this requires a non-minimal mapping (5-ary for QPSK sources).

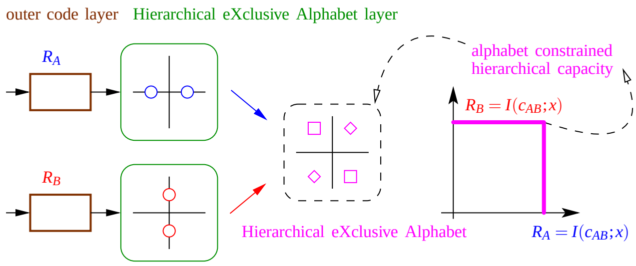

Figure 3. Layered HXC Design theorem graphical demonstration.

4) Equal Information Rates and Uniform-Input Capacity: For the present, we assume equal data information rates $R_A = R_B$ with common shared codebook and mappers. A solution of the asymmetric case is not dealt with here.

The mutual information $I(c_{AB};x)$ is not a capacity in a strict sense since we do not maximize over the input distribution. We will however assume that this distribution is given and it is a uniform distribution of a discrete channel input. This will be called a *uniform-input (alphabet constrained) capacity*. We denote that by letter C but the uniform-input assumption must be kept in mind.

D. Relay Hierarchical Decoding with Soft HXA Metric and Common Codebook

The soft-output H-MAC relay demodulator produces the symbol-wise metric $\mu(c_{AB})$ on the hierarchical data c_{AB} which is fed into the hierarchical H-MAC stage decoder $\mathscr{D}_{AB}$. The metric must properly reflect the fact that the useful signal $\mathscr{A}_u(c_{AB})$ is a set (class) with the cardinality generally higher than one and also depending on the channel parameters.

The symbol-wise metric is the likelihood of the given symbol $\mu(c_{AB}) = p(x|c_{AB})$. The condition in the $p(x|c_{AB})$ PDF is in fact a union of individual pairs $\{c_A, c_B\}$ conforming with the hierarchical symbol $c_{AB} = \mathscr{X}_c(c_A, c_B)$. Then

$$p(x|c_{AB}) = p\left(x| \bigcup_{c_A,c_B:\mathscr{X}_c(c_A,c_B)=c_{AB}} \{c_A,c_B\}\right) = \frac{p\left(x\cap\left(\bigcup_{c_A,c_B:\mathscr{X}_c(c_A,c_B)=c_{AB}}\{c_A,c_B\}\right)\right)}{p\left(\bigcup_{c_A,c_B:\mathscr{X}_c(c_A,c_B)=c_{AB}}\{c_A,c_B\}\right)}. \tag{18}$$

Pairs $\{c_A, c_B\}$ form a partition (disjoint subsets). Then

$$p(x|c_{AB}) = \frac{\sum_{c_A,c_B:\mathscr{X}_c(c_A,c_B)=c_{AB}} p(x|c_A,c_B)p(c_A,c_B)}{\sum_{c_A,c_B:\mathscr{X}_c(c_A,c_B)=c_{AB}} p(c_A,c_B)}. \tag{19}$$

An alternative notation form of the result can be obtained using the indicator (Kronecker delta) function $\mathscr{I}_c(c_A,c_B,c_{AB}) = \delta[c_{AB} - \mathscr{X}_c(c_A,c_B)]$

$$p(x|c_{AB}) = \frac{\sum_{c_A,c_B} p(x|c_A,c_B)\mathscr{I}_c(c_A,c_B,c_{AB})}{\sum_{c_A,c_B} \mathscr{I}_c(c_A,c_B,c_{AB})} \tag{20}$$

where we also utilized the assumption of uniformly distributed component code symbols $p(c_A,c_B) = \text{const}$.

In a *special case* of a *minimal* exclusive code (M-HXC) and uniformly distributed c_A, c_B, the hierarchical symbols have $\Pr\{c_{AB}\} = 1/\sum_{c_A,c_B} \mathscr{I}_c(c_A,c_B,c_{AB}) = 1/M_c$ and the conditional PDF is simply

$$p(x|c_{AB}) = \frac{1}{M_c} \sum_{c_A,c_B:\mathscr{X}_c(c_A,c_B)=c_{AB}} p_w\left(x - u(c_A,c_B)\right). \tag{21}$$

We sum only over such pairs c_A, c_B that are compliant with the given hierarchical symbol c_{AB}. There are M_c such uniformly probable cases for a *minimal* exclusive code and equal cardinality symbols. An alternative notation of the result with the indicator function is

$$p(x|c_{AB}) = \frac{1}{M_c} \sum_{c_A,c_B} p_w\left(x - u(c_A,c_B)\right) \mathscr{I}_c(c_A,c_B,c_{AB}). \tag{22}$$

It is important to realize that the decoding metric is *not* simply a distance, even in the simplest case of the Gaussian channel and minimal HXC. The signal space 2-source codebook $u(c_A,c_B)$ thus cannot be generally optimized with the distance criterion (compare this to [6]). In the case of extended mapping, the expression (20) is even more complicated due to the necessity of respecting particular c_A, c_B mapping on c_{AB}.

IV. Throughput Rate Region

A. Hierarchical MAC Rate Region

1) Hierarchical Mutual Information: It follows from the Layered HXC Design theorem that the throughput rate region is rectangular and given by the hierarchical mutual information $I(c_{AB};x)$. It is evaluated for given chosen symbol alphabets $\mathscr{A}_s$ with given channel parameters

$$x = u\left(s(c_A) + hs(c_B)\right) + w. \tag{23}$$

The hierarchical code symbols are mapped on the useful signal by $u(c_A,c_B) = \mathscr{A}_u(c_{AB})$. The transmitted signals, the useful H-MAC signal and the received signal are signal space representations $s(c_A), s(c_B), u, x \in \mathbb{C}^N$. The mutual information evaluation must respect the fact that the $\mathscr{A}_u(c_{AB})$ is generally a set (class) of multiple possible symbols for each particular c_{AB}.

The hierarchical mutual information (uniform-input capacity) is

$$C_{AB} = I(c_{AB};x) = \mathscr{H}[x] - \mathscr{H}[x|c_{AB}] \tag{24}$$

$$= \mathscr{H}[x] - \mathscr{H}[x|\mathscr{A}_u(c_{AB})]. \tag{25}$$

The channel parameter h and symbol space alphabets $s(\bullet)$ are *implicitly* hidden in the hierarchical class $\mathscr{A}_u(c_{AB})$ definition and we do not use an explicit notation for that.

2) Received Signal Entropy: From the entropy definition, we get (all integrals are over $\mathbb{C}^N$ support)

$$\mathscr{H}[x] = -\int p(x) \lg p(x)\, dx \tag{26}$$

where the PDF is

$$p(x) = \frac{1}{M_c^2} \sum_{c_A,c_B} p_w(x - u(c_A,c_B)) \tag{27}$$

and $p_w(w)$ is the PDF of complex signal space AWGN representation

$$p_w(w) = \frac{1}{\left(\pi\sigma_W^2\right)^N} \exp\left(-\frac{\|w\|^2}{\sigma_w^2}\right). \tag{28}$$

We assumed uniformly distributed code symbols $c_{A,}c_B$.

3) Conditional Entropy: An evaluation of the conditional entropy $\mathscr{H}[x|c_{AB}] = \mathscr{H}[x|\mathscr{A}_u(c_{AB})]$ is slightly more complicated compared to the classical point-to-point single user channel. Generally, it is *not* the entropy of the AWGN $\mathscr{H}[x|c_{AB}] \neq \mathscr{H}[w]$ as it would have been in the single user channel. This is because, given c_{AB}, the actual useful signal is a multi-symbol class $\mathscr{A}_u(c_{AB})$.

From the definition, we get

$$\mathscr{H}[x|c_{AB}] = -\sum_{c_{AB}} \Pr\{c_{AB}\} \int p(x|c_{AB}) \lg p(x|c_{AB})\, \mathrm{d}x. \tag{29}$$

Then we apply the results of the $p(x|c_{AB})$ evaluation from Section III-D.

B. MAC and Cut-Set Bound Reference

As a reference case, we consider the classical MAC rate region first and second order limits

$$I(c_A; x|c_B),\ I(c_A, c_B; x). \tag{30}$$

The nth order limit in the cut-set bound is understood as the limit for the sum-rate containing n individual rates $\sum_{i\in\mathscr{R}} R_i$, where $|\mathscr{R}| = n$. Notice that the first order rate limit $I(c_A; x|c_B)$ is in fact a cut-set bound [20] with finite alphabet limitation. The second order limit $I(c_A, c_B; x)$ corresponds to the joint decode and forward strategy where individual data streams from A and B are separately decoded at the relay.

For comparison, we evaluate $I(c_A; x|c_B)$ as a limiting performance criterion. The value R_0 is in fact the mutual information of the single user A channel as if there is no user B at all. Notice that this can be quite easily modeled by setting the channel B transfer $h = 0$. The uniform-input alphabet constrained capacity is found as $C_0 = I(c_A; x|c_B)$.

The second order cut-set bound limit is in fact a sum-rate. We evaluate the sum-rate rescaled to one user (assuming symmetric rates from nodes A and B)

$$C_s = \frac{1}{2} I(c_A, c_B; x). \tag{31}$$

All capacities and regions are depicted in Fig. 4.

C. Unconstrained Capacity

We also include the true alphabet unconstrained cut-set bound capacity for given input variance which is

$$C_u = \lg\left(1 + \frac{\mathrm{E}\left[\|\mathbf{s}_A\|^2\right]}{\sigma_w^2}\right). \tag{32}$$

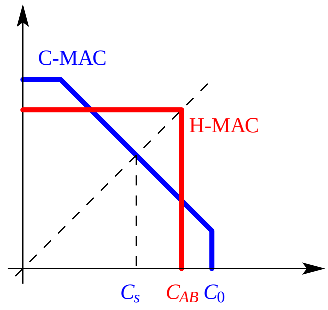

Figure 4. Capacity regions for H-MAC.

D. Hierarchical BC

In order to evaluate end-to-end system performance, we also assess the BC phase. We make the realistic assumption that the channel is reciprocal (as in [6]). Such a channel is assumed to have the same signal space alphabet, the same signal-to-noise ratio and the same Gaussian noise variance. The usage of the same alphabet implies that the *minimal* HXC mapping is used. Under the above-stated assumptions and under the assumption of perfect C-SI at the destination, the capacity of the H-BC link is equal to the first order cut-set bound of the H-MAC link

$$C_{\mathrm{HBC}} = C_0. \tag{33}$$

As a direct consequence, the H-MAC must be always the performance bottleneck for the *reciprocal* channel case. An analysis of the impact of the imperfect (partial) C-SI is treated in [18].

V. Numerical Results and Discussion

The hierarchical rate and single-user (alphabet limited cut-set bounds) rates were evaluated for a number of signal alphabets $\mathscr{A}_s$ and channel parameters. The alphabet $\mathscr{A}_s$ is indexed by symbols $c_A, c_B \in \{0, \ldots, M_c - 1\}$. We used a natural constellation index mapping for BPSK, QPSK and 8PSK constellations $\{\exp(\mathrm{j}\, i 2\pi/M_c\}_{i=0}^{M_c-1}$. The constellation denoted by QPSK-cross has indexing defined by $\{1, -1, \mathrm{j}, -\mathrm{j}\}$. 16QAM constellation uses a natural row-wise index mapping starting from the south-west corner. The exclusive *minimal* hierarchical mapping is $c_{AB} = c_A \oplus c_B$. The amplitude of $h = |h| \exp(\mathrm{j}\,\psi)$ is kept constant ($|h| = 1$) in our setup to respect the symmetry of the rates from A and B.

Figs. 6, 7, 8, 9, 10 show comparisons of the hierarchical symmetric capacity with various channel parameters with the alphabet constrained cut-set bound (both first and second order bound) and the unconstrained AWGN cut-set bound capacity. There are two main observations. First, we see that HXA outperforms the classical C-MAC capacity second order cut-set bound (sum-rate per user) and closely approaches the first order cut-set bound for medium to high signal to noise ratios. This means that the 2-WRC behaves at the MAC stage as if there was just one user alone. In the region of very low signal to noise ratio, the C-MAC is marginally better. However this is strongly affected by the choice of constellations, which is ad-hoc and not optimized. An optimization of the component constellation could improve this. The second observation shows a relatively small dispersion of the capacity over all possible channel relative phase parameters for some component symbol constellations. This is strongly dependent on the constellation. For example, the QPSK-cross constellation is particularly vulnerable to this phenomenon. Again, proper selection and optional optimization of the alphabet components is an important aspect.

Encouraged by the previous observation, we also plot a relative H-MAC capacity degradation due to the phase channel relative rotation ψ (Figs. 11, 12, 13). Various channel phases cause a capacity degradation due to the movement of the composite hierarchical points u. Their mutual position influences the metric (20), the decision regions, and thus also the capacity. However we can observe two types of behavior. In the first case, the capacity is just a little lowered when the points occupy less favorable mutual positions. But for some constellations and a given indexing catastrophic failure of the exclusivity law may occur, where several points belonging to different hierarchical indices c_{AB} reach same position. Fortunately these phases are relatively isolated. This observation encourages an evaluation of the mean capacities, which indicates that they have only a limited impact on the mean capacity over all phase rotations. The mean capacity is quite close to the maximum and, most importantly, it also exceeds the C-MAC capacity.

The results for the mean value over all channel phases have a direct practical implication. They suggest the use of *phase scrambling* schemes. The phase scrambling pattern would be known to the relay, which could easily unscramble it. This operation would effectively erase the influence of the particular critical phase channel rotation at the price of slightly lower throughput.

There are two possible phase scrambling approaches. The first (inspired by traditional information theory, but less practical) uses a block-wise phase scrambling. Each block is assumed to be long enough (at least approximately) to support the capacity achieving code. The transmitters are assumed to have channel state information and accordingly they switch codebooks with rate corresponding to the actual channel state. The mean achievable rate is then $\mathrm{E}_{\psi}[C_{AB}(\psi)]$.

Much more practical is the second option which requires *no* channel state information at the transmitters and performs the scrambling symbol-wise. It exploits the idea from [21] (Sec. 3.2): the capacity of the receiver-only channel side information channel with i.i.d. channel (per-symbol) channel states and the transmitted signal with channel state independent distribution is again given by $\mathrm{E}_{\psi}[C_{AB}(\psi)]$. The codebook for a such system is constructed for the mean rate $\mathrm{E}_{\psi}[C_{AB}(\psi)]$. Both the real channel phase and the scrambling phase are known at the receiver. Assuming that the scrambling is uniform and i.i.d. per symbol and the real phase shift is block-constant the resulting channel state is also uniform and i.i.d. The system

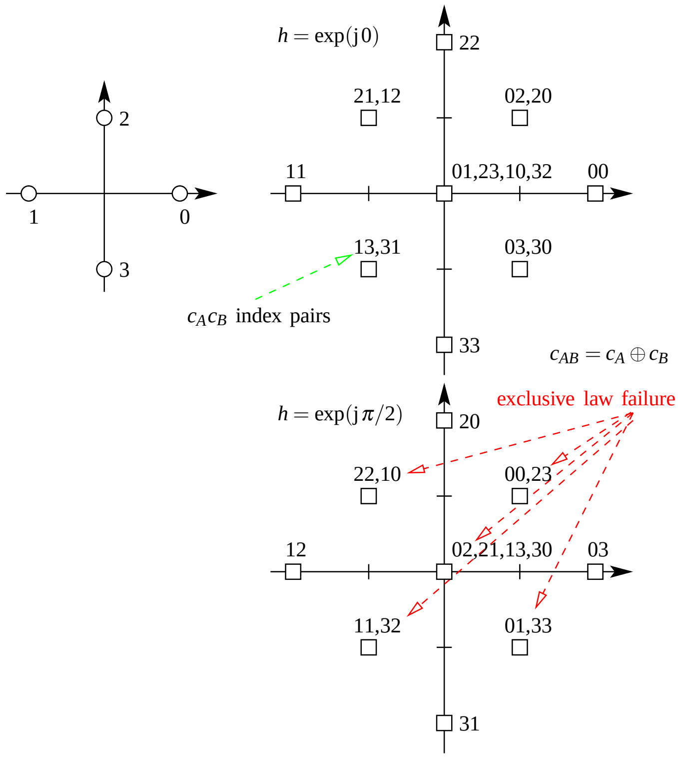

Figure 5. An example of the catastrophic exclusive law failure for QPSK-cross (cross indexing) and specific phase rotation.

then can use a single codebook with rate $\mathrm{E}_{\psi}[C_{AB}(\psi)]$ and this would be achievable.

The average rates $\mathrm{E}_{\psi}[C_{AB}(\psi)]$ are lower than the rates $\max_{\psi}[C_{AB}(\psi)]$ (see $\mathrm{E}_{\psi}[C_{AB}(\psi)]$ curves in Figs. 6, 7, 8, 9, 10) but provide a simple way to overcome the problem with a critical phase shift which can substantially degrade the throughput. The mean rate is then a sensible performance measure. The performance gain can be either measured by the outage probability (see [19]) or by $\mathrm{E}_{\psi}[C_{AB}(\psi)]/\min_{\psi}[C_{AB}(\psi)]$ ratio.

The fixed inner mapper can be used under the channel fading in the same way as any other symbol mapper in traditional *non-adaptive* point-to-point links. An adaptive (feedback) code/mapper design is outside the focus of the present paper. The performance of the nonadaptive system can be evaluated under two conditions. For non-ergodic channel behaviour during the codeword frame, the outage capacity is the proper evaluator, as addressed in [19]. In this paper, we have focused rather on a performance indicator suitable for the ergodic channel behaviour during the codeword frame, that is, the mean capacity. Our goal is simply to capture the impact of the phase variations *only*, since phase shift can directly cause exclusive law failure even at high SNR and with no amplitude fading. This means that it is more critical than the amplitude variations which cause hierarchical relay constellation scaling (with obvious impacts on achievable rate) but no unexpected exclusive law failures.

In order to assess the impact of the phase and amplitude

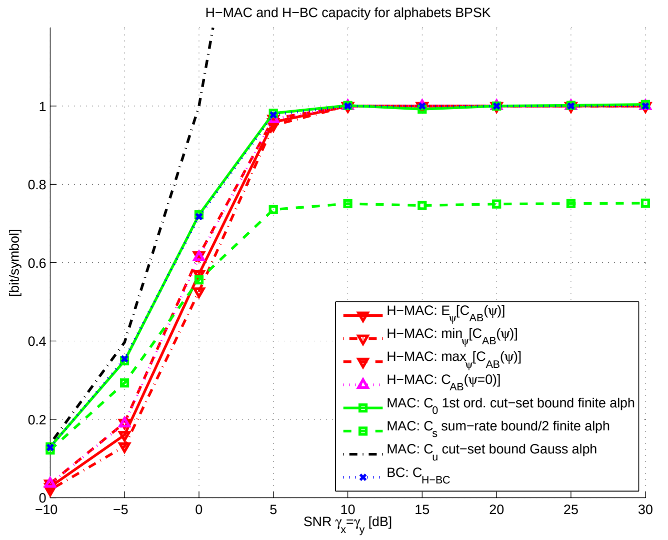

Figure 6. A comparison of capacities $C_{AB}, C_0, C_s, C_{\mathrm{HBC}}$ for BPSK alphabet and various channel phase parameters.

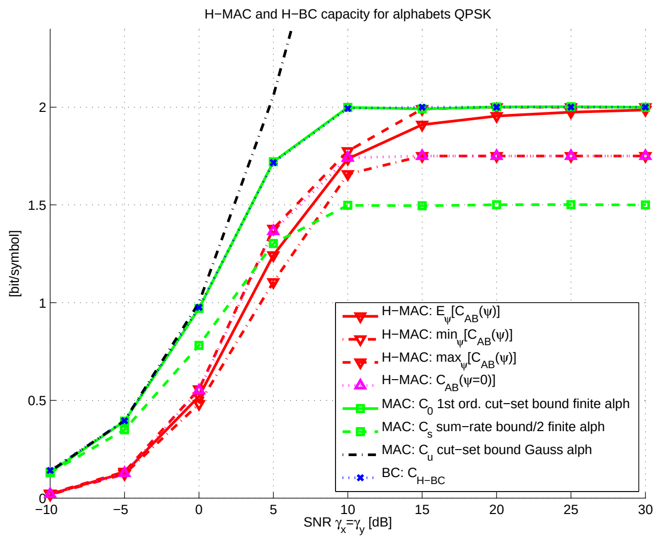

Figure 7. A comparison of capacities $C_{AB}, C_0, C_s, C_{\mathrm{HBC}}$ for QPSK alphabet and various channel phase parameters.

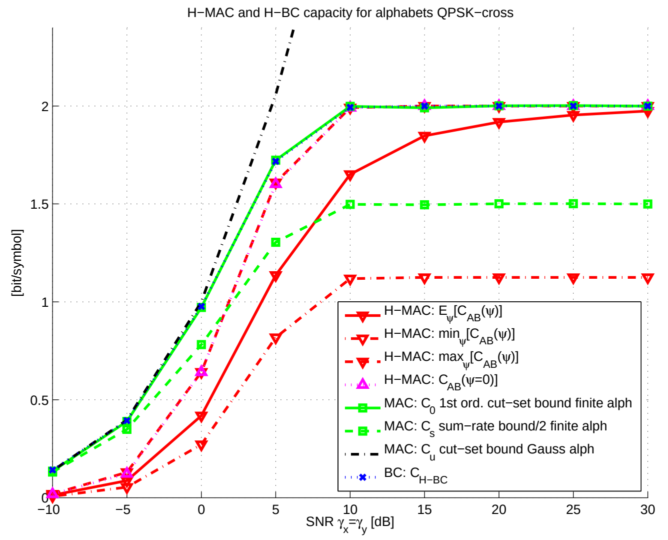

Figure 8. A comparison of capacities $C_{AB}, C_0, C_s, C_{\mathrm{HBC}}$ for QPSK-cross alphabet and various channel phase parameters.

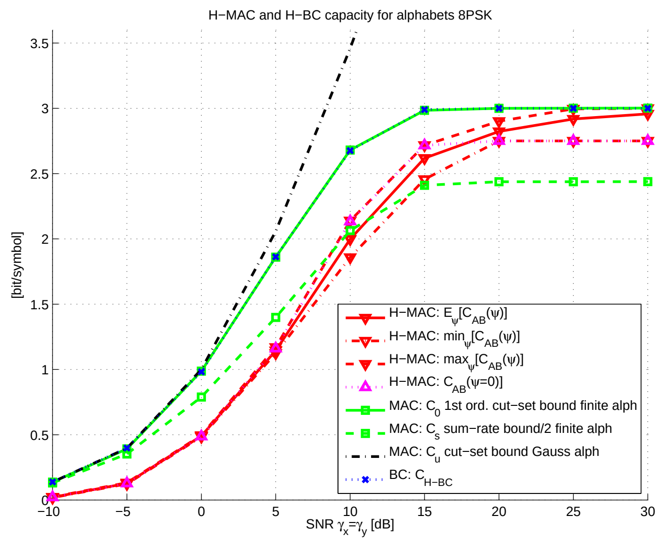

Figure 9. A comparison of capacities $C_{AB}, C_0, C_s, C_{\mathrm{HBC}}$ for 8PSK alphabet and various channel phase parameters.

variations which occur in practice on the performance of the scheme, we have also evaluated the hierarchical capacity in a Rayleigh fading environment. We assume that both links SA-R (h_A) and SB-R (h_B) are independent Rayleigh block constant fading channels with unity variance of the channel transfer. We evaluate the rates $C_{AB} = I(c_{AB}; x)$. These rates depend on the channel state. We evaluate their mean values over all channel states $\mathrm{E}[C_{ABi}(h_A, h_B)]$. In order to asses a stochastic description of the capacity fluctuations, we also evaluate the outage capacity $P_{\mathrm{out}} = \Pr\{C_{AB}(h_A, h_B) < R_{AB}\}$. It is evaluated by plotting the maximum achievable rate R_{AB} at the given probability of the outage P_{out}.

Results for the full Rayleigh fading case are in Fig. 14. Fig. 15 shows a comparison of these results with two additional cases. The first is the case where we keep *only the phase fluctuations* by using channel gains $h_A/|h_A|$ and $h_B/|h_B|$ instead h_A and h_B. The second is on the other hand the situation where we keep *only the magnitude fluctuations* by using channel gains $|h_A|$ and $|h_B|$ instead h_A and h_B. We see that full fading degrades the ergodic capacity performance by around 5 dB compared to phase only fluctuation. We observe also that the phase fluctuations have a significant effect in the medium to high SNR regime and cause substantial spreading of the capacity distribution, while in the low SNR regime this effect is relatively small. The amplitude fluctuations cause a similar spreading in the low and medium SNR range. This suggests that the additional performance loss in the full fading case may be attributed simply to the expected capacity loss due to fading,

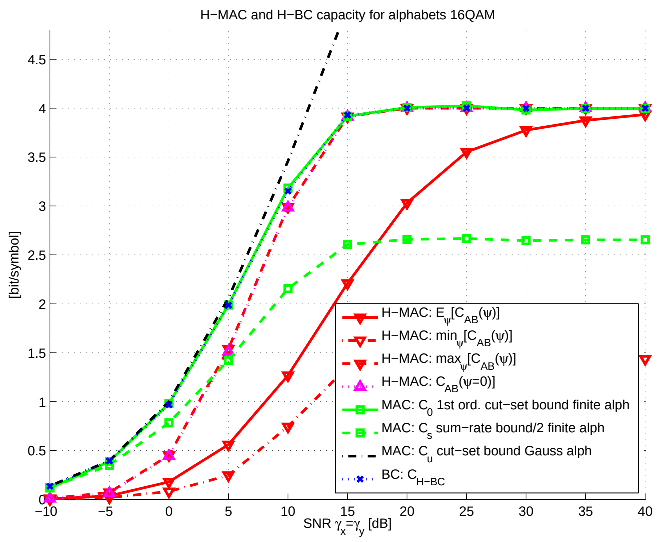

Figure 10. A comparison of capacities $C_{AB}, C_0, C_s, C_{\mathrm{HBC}}$ for 16QAM alphabet and various channel phase parameters.

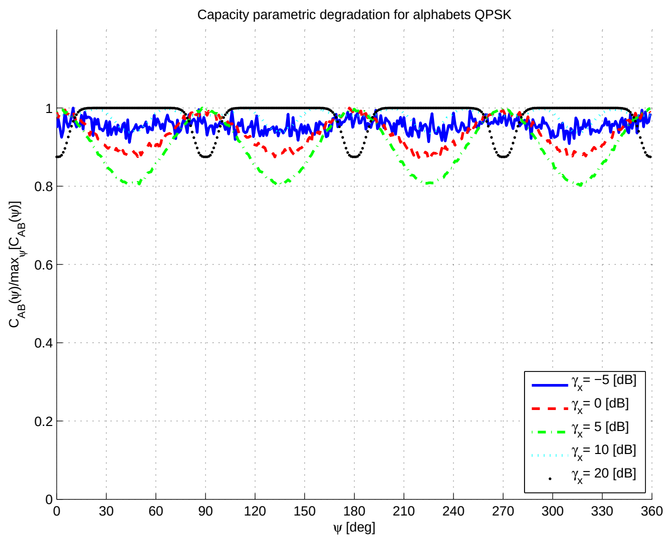

Figure 12. A degradation of the capacity $C_{AB}(\psi)/\max_{\psi'}[C_{AB}(\psi')]$ as a function or channel relative phase rotation for QPSK alphabet.

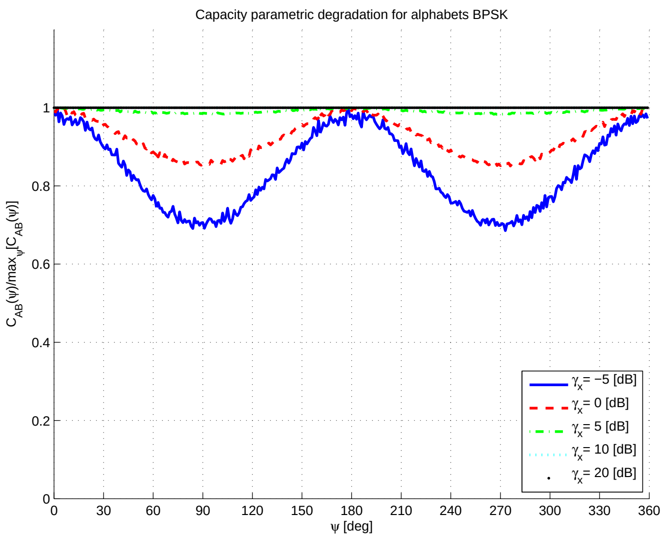

Figure 11. A degradation of the capacity $C_{AB}(\psi)/\max_{\psi'}[C_{AB}(\psi')]$ as a function or channel relative phase rotation for BPSK alphabet.

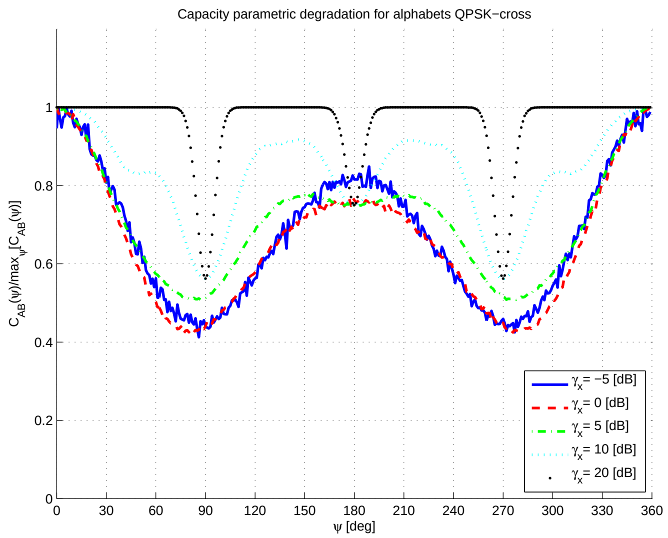

Figure 13. A degradation of the capacity $C_{AB}(\psi)/\max_{\psi'}[C_{AB}(\psi')]$ as a function or channel relative phase rotation for QPSK-cross alphabet.

rather than to any exclusive law failures. It is worth noting that the addition of amplitude fading actually increases the 10% outage capacity at high SNR, because it makes exclusive law failures less likely to occur.

VI. Conclusions

The capacity results presented show that layered design of the HXC can be a viable solution. It can utilize an enormous range of currently available state-of-the-art outer capacity approaching codes. The relay H-MAC decoder can utilize a soft decoding metric due to the symbol-by-symbol based hierarchical processing at the relay H-MAC stage. The penalty paid for the layered design (w.r.t. the cut-set bound) is negligible in the moderate to high SNR regions. The penalty is higher for very small SNR where it is essentially determined by the component alphabets, which here have been chosen on an ad-hoc basis.

The capacity results also show that the channel parameters have a significant impact for particular channel symbol indexing patterns and the corresponding hierarchical mappings. This is clearly visible on the graphs showing the capacity degradation as a function of the phase. Some channel rotations have a catastrophic effect on the capacity for a given exclusive mapping. Some channel rotations cause the points corresponding to different hierarchical codewords to fall into the same useful signal u and in fact, the useful signal is *no longer exclusive*. This highlights the importance of code design that properly takes the channel parameters into an account,

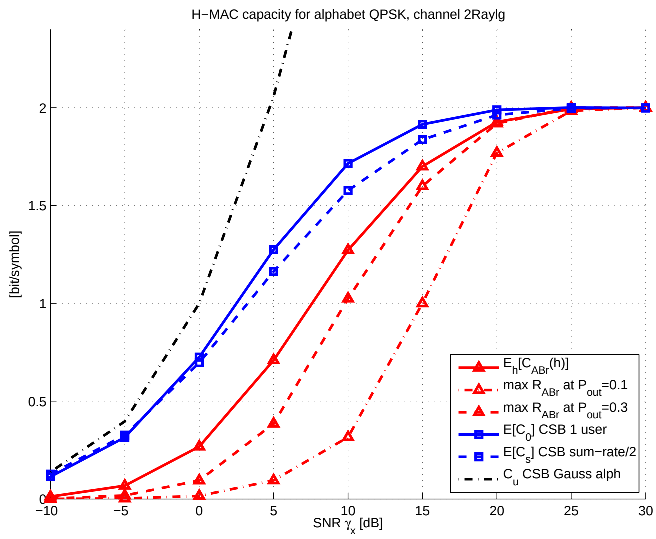

Figure 14. Mean capacity $\mathrm{E}[C_{AB}(h_A,h_B)]$ and outage capacity $P_{\text{out}} = \Pr\{C_{AB}(h_A,h_B) < R_{AB}\}$ for QPSK alphabet in Rayleigh fading channel.

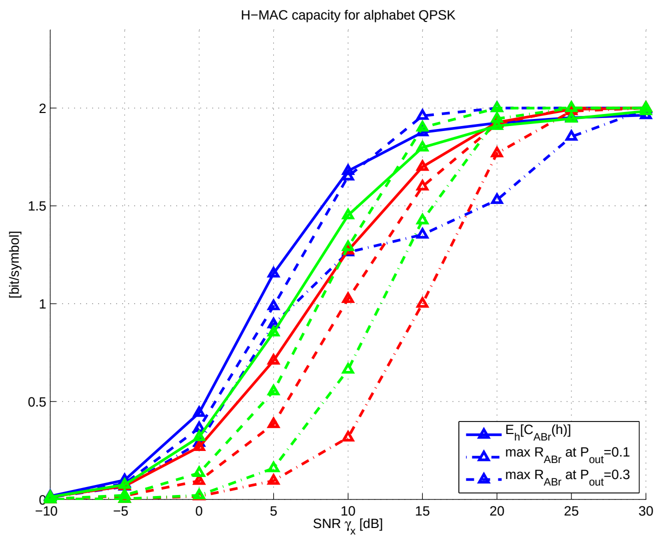

Figure 15. Mean capacity $\mathrm{E}[C_{AB}(h_A,h_B)]$ and outage capacity $P_{\text{out}} = \Pr\{C_{AB}(h_A,h_B) < R_{AB}\}$ for QPSK alphabet in Rayleigh fading channel with full fading (red), phase fluctuation only (blue), and amplitude fluctuation only (green).

e.g. by being an invariant to this. Signal phase predistortion to avoid such situations is feasible only in very specific situations, e.g. for channels which are invariant for several frames. It is not generally suitable for short packet oriented transmissions. However, on the other side, we observe that the catastrophic effects are relatively limited to a narrow range of phases and have only mild impact on the mean capacity for some constellations. This observation may lead to number of phase scrambling practical solutions.

VII. Appendix

A. Monte-Carlo Evaluation of the Integrals

All integrals involved in the entropy evaluations are complicated multi-fold integrals over the N-dimensional complex plane. Their direct evaluation by traditional numerical integration procedures is known to be complex and very slowly converging. Better computational efficiency is achieved by Monte-Carlo numerical evaluation of the integrals. This is not to be confused with Monte-Carlo simulation. We use it for the numerical evaluation of the integrals, not for running the system with some particular coding and signal processing algorithms. The method is usable for evaluation of any summation or integration of the form

$$I = \int f(x) g(x)\, \mathrm{d}x. \tag{34}$$

We consider $f(x)$ as a PDF, and hence the integral can be evaluated as

$$I = \mathrm{E}_{f(x)}[g(x)]. \tag{35}$$

Providing that we are able to generate random values with the distribution $f(x)$: $x \sim f(x)$ then the integral is approximated by evaluating the empirical mean $\hat{\mathrm{E}}[\bullet]$ of $g(x)$. Generation of the random variable with given a density is a relatively easy task particularly for linear Gaussian mixture densities. We can also always resort to building a full system model producing the desired random variable and feeding that system with its all random excitations — typically data and noise sources.

A numerical evaluation of $\mathscr{H}[x]$ gives

$$\mathscr{H}[x] \approx -\hat{\mathrm{E}}_{x\sim p(x)}\left[\lg\left(\frac{1}{M_c^2}\sum_{c_A,c_B} p_w(x-u(c_A,c_B))\right)\right] \tag{36}$$

where x is generated with the distribution (27). The conditional entropy approximation is obtained as

$$\mathscr{H}[x|c_{AB}] = -\hat{\mathrm{E}}_{x,c_{AB}\sim p(x,c_{AB})}\left[\lg p(x|c_{AB})\right]. \tag{37}$$

References

[1] R. W. Yeung, S.-Y. R. Li, N. Cai, and Z. Zhang, *Network Coding Theory*. now Publishers, 2006.

[2] T. Wang and G. B. Giannakis, “Complex field network coding for multiuser cooperative communications,” *IEEE J. Sel. Areas Commun.*, vol. 26, pp. 561–571, Apr. 2008.

[3] P. Popovski and T. Koike-Akino, “Coded bidirectional relaying in wireless networks,” in *Advances in Wireless Communications* (V. Tarokh, ed.), Springer, 2009.

[4] T. Koike-Akino, P. Popovski, and V. Tarokh, “Denoising maps and constellations for wireless network coding in two-way relaying systems,” in *Proc. IEEE Global Telecommun. Conf. (GlobeCom)*, 2008.

[5] T. Koike-Akino, P. Popovski, and V. Tarokh, “Denoising strategy for convolutionally-coded bidirectional relaying,” in *Proc. IEEE Internat. Conf. on Commun. (ICC)*, 2009.

[6] T. Koike-Akino, P. Popovski, and V. Tarokh, “Optimized constellations for two-way wireless relaying with physical network coding,” *IEEE J. Sel. Areas Commun.*, vol. 27, pp. 773–787, June 2009.

[7] W. Nam, S.-Y. Chung, and Y. H. Lee, “Capacity bounds for two-way relay channels,” in *Proc. Int. Zurich Seminar on Communications*, 2008.

[8] I.-J. Baik and S.-Y. Chung, “Network coding for two-way relay channels using lattices,” in *Proc. IEEE Internat. Conf. on Commun. (ICC)*, 2008.

[9] U. Erez and R. Zamir, “Achieving $1/2\log(1+\mathrm{SNR})$ on the AWGN channel with lattice encoding and decoding,” *IEEE Trans. Inf. Theory*, vol. 50, pp. 2293–2314, Oct. 2004.

[10] B. Nazer and M. Gastpar, "Computation over multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 53, pp. 3498–3516, Oct. 2007.

[11] B. Nazer and M. Gastpar, "Compute-and-forward: Harnessing interference through structured codes," *IEEE Trans. Inf. Theory*, 2010. Submitted for publication.

[12] A. Y.-C. Peng, S. Yousefi, and I.-M. Kim, "On error analysis and distributed phase steering for wireless network coding over fading channels," *IEEE Trans. Wireless Commun.*, vol. 8, pp. 5639–5649, Nov. 2009.

[13] S. Zhang and S.-C. Liew, "Channel coding and decoding in a relay system operated with physical-layer network coding," *IEEE J. Sel. Areas Commun.*, vol. 27, pp. 788–796, June 2009.

[14] M. P. Wilson, K. Narayanan, H. D. Pfister, and A. Sprintson, "Join physical layer coding and network coding for bidirectional relaying," *IEEE Trans. Inf. Theory*, vol. 56, pp. 5641–5654, Nov. 2010.

[15] N. Sommer, M. Feder, and O. Shalvi, "Low-density lattice codes," *IEEE Trans. Inf. Theory*, vol. 54, pp. 1561–1585, Apr. 2008.

[16] T. Koike-Akino, P. Larsson, P. Popovski, and V. Tarokh, "Non-linear network coding in two-way relaying discrete channels," in *Int. Conf. on Wireless Com. and Sig. Proc. (WCSP)*, 2009.

[17] J. Sykora and A. Burr, "Wireless network coding: The network aware PHY layer," in *Proc. Int. Symp. of Wireless Communication Systems (ISWCS)*, (York, United Kingdom), Sept. 2010. tutorial.

[18] J. Sykora and A. Burr, "Network coded modulation with partial side-information and hierarchical decode and forward relay sharing in multi-source wireless network," in *Proc. European Wireless Conf. (EW)*, (Lucca, Italy), pp. 1–7, Apr. 2010.

[19] J. Sykora and A. Burr, "Hierarchical alphabet and parametric channel constrained capacity regions for HDF strategy in parametric wireless 2-WRC," in *Proc. IEEE Wireless Commun. Network. Conf. (WCNC)*, (Sydney, Australia), pp. 1–6, Apr. 2010.

[20] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. John Wiley & Sons, 1991.

[21] A. J. Goldsmith and P. P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Trans. Inf. Theory*, vol. 43, pp. 1986–1992, Nov. 1997.

PLACE PHOTO HERE

Jan Sykora (M'03, Jan.Sykora@fel.cvut.cz, http://radio.feld.cvut.cz/~sykora, MIEEE) received the M.Sc. and Ph.D. degrees in Electrical Engineering from Czech Technical University in Prague, Czech Republic, in 1987 and 1993, respectively. Since 1991, he has been with the Faculty of Electrical Engineering, Czech Technical University in Prague, where he is now a Professor of Radio Engineering. His research includes work on wireless communication and information theory, cooperative and distributed modulation, wireless network coding and distributed signal processing, MIMO systems, nonlinear space–time modulation and coding, and iterative processing. He has served on various IEEE conferences as a Technical Program and Organizing Committee member and chair. He has led a number of industrial and research projects financed by EU and national agencies.

PLACE PHOTO HERE

Alister Burr (M'92, agb1@ohm.york.ac.uk, http://www.elec.york.ac.uk/staff/agb1.html, BSc (Soton), PhD (Bristol), MIET, MIEEE, CEng) is Professor of Communications at the University of York, U.K. He has many years research experience in wireless networks, especially adaptive modulation and coding techniques, and also in evaluating overall network capacity, including realistic propagation modelling. He is currently Chair WG1, on Radio Systems, of COST Action 2100 "Pervasive Mobile and Ambient Wireless Networks", and Associate Editor of IEEE Communications Letters.