# DECLARATION OF JUNE ANN MUNFORD

1. My name is June Ann Munford. I am over the age of 18, have personal knowledge of the facts set forth herein, and am competent to testify to the same.

2. I earned a Master of Library and Information Science (MLIS) from the University of Wisconsin-Milwaukee in 2009. I have over ten years of experience in the library/information science field. Beginning in 2004, I have served in various positions in the public library sector including Assistant Librarian, Youth Services Librarian and Library Director. I have attached my Curriculum Vitae as Appendix CV.

3. During my career in the library profession, I have been responsible for materials acquisition for multiple libraries. In that position, I have cataloged, purchased and processed incoming library works. That includes purchasing materials directly from vendors, recording publishing data from the material in question, creating detailed material records for library catalogs and physically preparing that material for circulation. In addition to my experience in acquisitions, I was also responsible for analyzing large collections of library materials, tailoring library records for optimal catalog

search performance and creating lending agreements between libraries during my time as a Library Director.

4. I am fully familiar with the catalog record creation process in the library sector. In preparing a material for public availability, a library catalog record describing that material would be created. These records are typically written in Machine Readable Catalog (herein referred to as "MARC") code and contain information such as a physical description of the material, metadata from the material's publisher, and date of library acquisition. In particular, the 008 field of the MARC record is reserved for denoting the date of creation of the library record itself. As this typically occurs during the process of preparing materials for public access, it is my experience that an item's MARC record indicates the date of an item's public availability.

5. Typically, in creating a MARC record, a librarian would gather various bits of metadata such as book title, publisher and subject headings among others and assign each value to a relevant numerical field. For example, a book's physical description is tracked in field 300 while title/attribution is tracked in field 245. The 008 field of the MARC record is reserved for denoting the creation of the library record itself. As this is the only date reflecting the inclusion of said materials within the library's collection, it is my experience

that an item's 008 field accurately indicates the date of an item's public availability.

6. I have reviewed Exhibit LGE1014, "Turbo-coded APSK Modulations Design for Satellite Broadband Communications" by R. De Gaudenzi, A. Guillen I Fabregas and A. Martinez as published in *International Journal of Satellite Communications and Networking*, July/August 2006.

7. Attached hereto as Appendix DEGAUDENZI01 is a true and correct copy of the MARC record for *International Journal of Satellite Communications and Networking* as held by the Penn State University library. I secured this record myself from the library's public catalog. The MARC record contained within Appendix DEGAUDENZI01 accurately describes the title, author, publisher, and ISSN number of *International Journal of Satellite Communications and Networking.* The 'Dates of Publication and/or Sequential Designation' field of this record visible on page 2 indicates this collection includes the July/August 2006 edition of *International Journal of Satellite Communications and Networking* containing "Turbo-coded APSK Modulations Design for Satellite Broadband Communications" by R. De Gaudenzi, A. Guillen I Fabregas and A. Martinez.

8. Attached hereto as Appendix DEGAUDENZI02 is a true and correct copy of "Turbo-coded APSK Modulations Design for Satellite Broadband Communications" as published in *International Journal of Satellite Communications and Networking*. I secured this PDF copy myself from Penn State University's digital collection. In comparing Exhibit LGE1014 to Appendix DEGAUDENZI02, it is my determination that Exhibit LGE1014 is a true and correct copy of "Turbo-coded APSK Modulations Design for Satellite Broadband Communications" by R. De Gaudenzi, A. Guillen I Fabregas and A. Martinez as published in *International Journal of Satellite Communications and Networking* July/August 2006.

9. The 008 field of the MARC record in Appendix DEGAUDENZI01 indicates the date of record creation. The 008 field of Appendix DEGAUDENZI01 indicates the Penn State University library first acquired this journal as of October 21, 2002 and holds the journal in perpetuity since that date. Considering this information, it is my determination that *International Journal of Satellite Communications and Networking* July/August 2006 and therefore "Turbo-coded APSK Modulations Design for Satellite Broadband Communications" was made available to the public shortly after its initial release in Summer 2006.

10. I have reviewed Exhibit LGE1018, "Bit-Interleaved Coded Modulation" by G. Caire, G. Taricco and E. Biglieri as published in *IEEE Transactions of Information Theory*, May 1998.

11. Attached hereto as Appendix CAIRE01 is a true and correct copy of the MARC record for *IEEE Transactions of Information Theory* as held by the Carnegie Library of Pittsburgh. I secured this record myself from the library's public catalog. The MARC record contained within Appendix CAIRE01 accurately describes the title, author, publisher, and ISSN number of *IEEE Transactions of Information Theory.* The 'Lib Has' field of this record visible on page 2 indicates this collection includes the May 1998 edition of *IEEE Transactions of Information Theory* containing "Bit-Interleaved Coded Modulation" by G. Caire, G. Taricco and E. Biglieri.

12. Attached hereto as Appendix CAIRE02 is a true and correct copy of "Bit-Interleaved Coded Modulation" as published in *IEEE Transactions of Information Theory*. I secured these scans myself from the Carnegie Library of Pittsburgh's collection. In comparing Exhibit LGE1018 to Appendix CAIRE02, it is my determination that Exhibit LGE1018 is a true and correct copy of "Bit-Interleaved Coded Modulation" by G. Caire, G. Taricco and E.

Biglieri as published in *IEEE Transactions of Information Theory,* May 1998.

13. The 008 field of the MARC record in Appendix CAIRE01 indicates the date of record creation. The 008 field of Appendix CAIRE01 indicates the Carnegie Library of Pittsburgh first acquired this journal as of November 1, 1975 and held the journal in perpetuity until 2008. Considering this information, it is my determination that *IEEE Transactions of Information Theory* May 1998 and therefore "Bit-Interleaved Coded Modulation" was made available to the public shortly after its initial release in Summer 1998.

14. I have reviewed Exhibit LGE1019, *DVB The Family of International Standards for Digital Video Broadcasting,* edited by Ulrich Reimers, 2nd Edition.

15. Attached hereto as Appendix REIMERS01 is a true and correct copy of the MARC record for *DVB - The Family of International Standards for Digital Video Broadcasting* as held by the University of Pittsburgh library. I secured this record myself from the library's public catalog. The MARC record contained within Appendix REIMERS01 accurately describes the title,

author, publisher, and ISBN number of *DVB - The Family of International Standards for Digital Video Broadcasting.*

16. Attached hereto as REIMERS02 are selections from *DVB - The Family of International Standards for Digital Video Broadcasting* as held by the University of Pittsburgh's library. I secured these scans myself from the University of Pittsburgh library. In comparing Exhibit 1008 to Appendix REIMERS02, it is my determination that Exhibit 1008 is a true and correct copy of *DVB - The Family of International Standards for Digital Video Broadcasting* 2nd Edition.

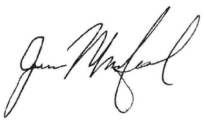
17. The 008 field of the MARC record in Appendix REIMERS01 indicates the date of record creation. The 008 field of Appendix REIMERS01 indicates the University of Pittsburgh library first acquired this book as of July 3, 2003. Considering this information, it is my determination that *DVB - The Family of International Standards for Digital Video Broadcasting* was made available to the public shortly after its initial acquisition in July 2003.

18. I have been retained on behalf of the Petitioner to provide assistance in the above-illustrated matter in establishing the authenticity and public

availability of the documents discussed in this declaration. I am being compensated for my services in this matter at the rate of $100.00 per hour plus reasonable expenses. My statements are objective, and my compensation does not depend on the outcome of this matter.

19. I declare under penalty of perjury that the foregoing is true and correct. I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code.

Dated: 11/28/2022



June Ann Munford

# APPENDIX CV

J. Munford
Curriculum Vitae

**Education**

University of Wisconsin-Milwaukee - MS, Library & Information Science, 2009
Milwaukee, WI

- Coursework included cataloging, metadata, data analysis, library systems, management strategies and collection development.
- Specialized in library advocacy, cataloging and public administration.

Grand Valley State University - BA, English Language & Literature, 2008
Allendale, MI

- Coursework included linguistics, documentation and literary analysis.
- Minor in political science with a focus in local-level economics and government.

**Professional Experience**

Researcher / Expert Witness, October 2017 – present
Freelance ● Pittsburgh, Pennsylvania & Grand Rapids, Michigan

- Material authentication and public accessibility determination. Declarations of authenticity and/or public accessibility provided upon research completion. Experienced with appeals and deposition process.

- Research provided on topics of public library operations, material publication history, digital database services and legacy web resources.

- Past clients include Alston & Bird, Arnold & Porter, Baker Botts, Fish & Richardson, Erise IP, Irell & Manella, O'Melveny & Myers, Perkins-Coie, Pillsbury Winthrop Shaw Pittman and Slayden Grubert Beard.

Library Director, February 2013 - March 2015
Dowagiac District Library ● Dowagiac, Michigan

- Executive administrator of the Dowagiac District Library. Located in

Southwest Michigan, this library has a service area of 13,000, an annual operating budget of over $400,000 and total assets of approximately $1,300,000.

● Developed careful budgeting guidelines to produce a 15% surplus during the 2013-2014 & 2014-2015 fiscal years while being audited.

● Using this budget surplus, oversaw significant library investments including the purchase of property for a future building site, demolition of existing buildings and building renovation projects on the current facility.

● Led the organization and digitization of the library's archival records.

● Served as the public representative for the library, developing business relationships with local school, museum and tribal government entities.

● Developed an objective-based analysis system for measuring library services - including a full collection analysis of the library's 50,000+ circulating items and their records.

November 2010 - January 2013
Librarian & Branch Manager, Anchorage Public Library ● Anchorage, Alaska

● Headed the 2013 Anchorage Reads community reading campaign including event planning, staging public performances and creating marketing materials for mass distribution.

● Co-led the social media department of the library's marketing team, drafting social media guidelines, creating original content and instituting long-term planning via content calendars.

● Developed business relationships with The Boys & Girls Club, Anchorage School District and the US Army to establish summer reading programs for children.

June 2004 - September 2005, September 2006 - October 2013
Library Assistant, Hart Area Public Library
Hart, MI

● Responsible for verifying imported MARC records and original MARC

cataloging for the local-level collection as well as the Michigan Electronic Library.

● Handled OCLC Worldcat interlibrary loan requests & fulfillment via ongoing communication with lending libraries.

**Professional Involvement**

Alaska Library Association - Anchorage Chapter
- Treasurer, 2012

Library Of Michigan
- Level VII Certification, 2008
- Level II Certification, 2013

Michigan Library Association Annual Conference 2014
- New Directors Conference Panel Member

Southwest Michigan Library Cooperative
- Represented the Dowagiac District Library, 2013-2015

**Professional Development**

Library Of Michigan Beginning Workshop, May 2008
Petoskey, MI
- Received training in cataloging, local history, collection management, children's literacy and reference service.

Public Library Association Intensive Library Management Training, October 2011
Nashville, TN
- Attended a five-day workshop focused on strategic planning, staff management, statistical analysis, collections and cataloging theory.

Alaska Library Association Annual Conference 2012 - Fairbanks, February 2012
Fairbanks, AK
- Attended seminars on EBSCO advanced search methods, budgeting, cataloging, database usage and marketing.

**Depositions**

2019 ● Fish & Richardson
IPR Petitions of 865 Patent, Apple v. Qualcomm (IPR2018-001281 / 39521-00421IP & IPR2018-01282 / 39521-00421IP2)

2019 ● Erise IP
Implicit, LLC v. Netscout Systems, Inc (Civil Action No. 2:18-cv-53-JRG)

2019 ● Perkins-Coie
Adobe Inc. v. RAH Color Technologies LLC (Cases IPR2019-00627, IPR2019-00628, IPR2019-00629 and IPR2019-00646)

2020 ● O'Melveny & Myers
Maxell, Ltd. v. Apple Inc. (Case 5:19-cv-00036-RWS)

2021 ● Pillsbury Winthrop Shaw Pittman LLP
Intel v. SRC (Case IPR2020-1449)

**Limited Case History & Potential Conflicts**

Alston & Bird
- Nokia (v. Neptune Subsea, Xtera)

Arnold & Porter
- Ivantis (v. Glaukos)

Erise I.P.
- Apple
  - v. Future Link Systems (IPRs 6317804, 6622108, 6807505, and 7917680)
  - v. INVT
  - v. Navblazer LLC (Case No. IPR2020-01253)

v. Qualcomm (IPR2018-001281, 39521-00421IP, IPR2018-01282, 39521-00421IP2)

v. Quest Nettech Corp, Wynn Technologies (Case No. IPR2019-00XXX, RE. Patent Re38137)

- Fanduel (v CGT)

- Garmin (v. Phillips North America LLC, Case No. 2:19-cv-6301-AB-KS Central District of California)

- Netscout

v. Longhorn HD LLC)

v. Implicit, LLC (Civil Action No. 2:18-cv-53-JRG)

- Sony Interactive Entertainment LLC

v. Bot M8 LLC

v. Infernal Technology LLC

- Unified Patents (v GE Video Compression, Civil Action No. 2:19-cv-248)

Fish & Richardson

- Apple

v. LBS Innovations

v. Masimo (IPR 50095-0012IP1, 50095-0012IP2, 50095-0013IP1, 50095-0013IP2, 50095-0006IP1)

v. Neonode

v. Qualcomm (IPR2018-001281, 39521-00421IP, IPR2018-01282, 39521-00421IP2)

- Dish Network

v. Realtime Adaptive Streaming, Case No 1:17-CV-02097-RBJ)

v. TQ Delta LLC

- Huawei (IPR 76933211)
- Kianxis
- LG Electronics (v. Bell Northern Research LLC, Case No. 3:18-cv-2864-CAB-BLM)
- Metaswitch
- MLC Intellectual Property (v. MicronTech, Case No. 3:14-cv-03657-SI)
- Realtek Semiconductor
- Quectel
- Samsung (v. Bell Northern Research, Civil Action No. 2:19-cv-00286-JRG)
- Texas Instruments

Irell & Manella

- Curium

O'Melveny & Myers

- Apple (v. Maxell, Case 5:19-cv-00036-RWS)

Perkins-Coie

- TCL Industries (v. Koninklijke Philips NV, PTAB Case Nos. IPR2021-00495, IPR2021-00496, and IPR2021-00497)

Pillsbury Winthrop Shaw Pittman

- Intel (v. FG SRC LLC, Case No. 6:20-cv-00315 W.D. Tex)

# APPENDIX CAIRE01

*screenshot-iiisy1.einetwork.net-2022.07.27-15_51_26*
*https://iiisy1.einetwork.net/search~S1?/o1752552/o1752552/1%2C1%2C1%2CE/marc&FF=o1752552&1%2C1%2C*
*27.07.2022*



```
LEADER 00000cas  2200000 a 4500
001    1752552
003    OCoLC
005    20150414134057.0
008    751101c19639999nyubr p       0   a0eng
010    58035217
016 7  SR0052230|2DNLM
022    0018-9448
030    IETTAW
032    256880|bUSPS
035    600008|a00008
040    DLC|cMUL|dNSD|dYUS|dMUL|dIUL|dOCL|dIXA|dNSD|dAIP|dDLC|dAIP
       |dDLC|dNST|dOCL|dNST|dNSD|dWAU|dNST|dWaOLN|dUtOrBLW
042    nsdp|alc
050 00 Q350|b.I2
069 1  SR0052230
082 00 001.53/9|219
210 0  IEEE trans. inf. theory
222  0 IEEE transactions on information theory
245 00 IEEE transactions on information theory /|cProfessional
       Technical Group on Information Theory.
246 10 Transactions on information theory
246 10 Information theory
246 23 Institute of Electrical and Electronics Engineers
       transactions on information theory
260    [New York] :|b[Institute of Electrical and Electronics
       Engineers],|c[01963]-
300    volumes ;|billustrations, portraits ;|c28-32 cm
310    Bimonthly,|bMar. 1968-
321    Quarterly,|bJan. 1963-Jan. 1968
336    text|btxt|2rdacontent
337    unmediated|bn|2rdamedia
338    volume|bnc|2rdacarrier
362 0  Vol. IT-9, no. 1 (Jan. 1963)-
500    Title from cover.
510 1  Applied science & technology index|x0003-6986|b1991-
510 1  Index to IEEE publications|x0099-1368
510 2  Chemical abstracts|x0009-2258
510 2  Computer & control abstracts|x0036-8113|b1968-
510 2  Electrical & electronics abstracts|x0036-8105|b1968-
510 2  Electronics and communications abstracts journal
       (Riverdale)|x0361-3313
510 2  Engineering index annual (1968)|x0360-8557
510 2  Engineering index bioengineering abstracts|x0736-6213
510 2  Engineering index energy abstracts|x0093-8408
510 2  Engineering index monthly (1984)|x0742-1974
510 2  Excerpta medica
510 2  International aerospace abstracts|x0020-5842
510 2  ISMEC bulletin|x0306-0039
510 2  Mathematical reviews|x0025-5629
510 2  Physics abstracts|x0036-8091|b1968-
510 2  Ship abstracts|x0346-1025
510 0  Engineering index monthly|x0013-7960
510 0  Nuclear science abstracts|x0029-5612
550    Vols. for 1963-Apr. 1964 issued by the Professional
       Technical Group on Information Theory of the IEEE; July
       1964-   by the Information Theory Group of the IEEE; <Jan.
       1968>-May 1989 by the IEEE Information Theory Group; July
       1989-   by the IEEE Information Theory Society.
650  0 Information theory|vPeriodicals.
650  0 Telecommunication|vPeriodicals.
710 2  Institute of Electrical and Electronics Engineers.
       |bProfessional Technical Group on Information Theory.
710 2  Institute of Electrical and Electronics Engineers.
       |bInformation Theory Group.
710 2  IEEE Information Theory Group.
710 2  IEEE Information Theory Society.
780 00 |tIRE transactions on information theory|x0096-1000
       |w(DLC)sf 84001167|w(OCoLC)2012604
850    CoU|aDLC|aFU|aNcD|aNcRS|aNjR|aPU|aTxU|aViW|aWaU
935    ADM-7810
```

| | |
|---|---|
| *Location* | Carnegie Library of Pittsburgh - Off Site Storage (Please ask for assistance) |
| *Lib Has* | Volume 9 (1963) to Volume 15 (1969). |
| *Location* | CLP Main (Oakland) Second Floor - Magazines |
| *Lib Has* | Volume 16 (1970) to Volume 54 (2008). |

*screenshot-iiisy1.einetwork.net-2022.07.27-15_51_10*
*https://iiisy1.einetwork.net/search/o1752552*
*27.07.2022*

A Quick Survey Catalog Home My Account Login Help

The CATALOG
Connecting Public Libraries in Allegheny County, PA

Your Library | Tips & Resources | Booklists | Find Articles | Interlibrary Loan | Get a Library Card

Start Over | Request | Add to Wish List | Add to Book Cart | MARC Display | More Information | Another Search

CONTROL NO 1752552 All Libraries System Sorted Search

Show available items only

(Search History) Record: Previous Next

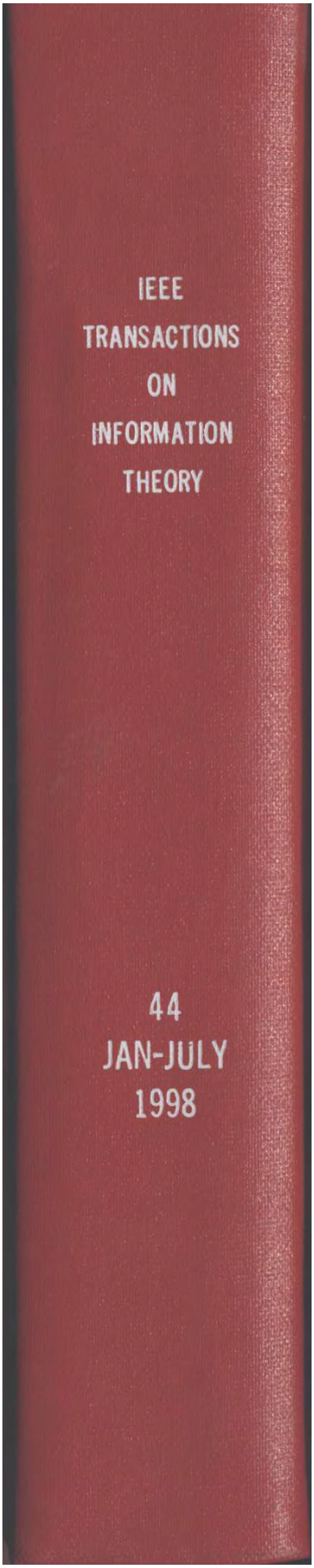
| | |
|---|---|
| **Title** | **IEEE transactions on information theory / Professional Technical Group on Information Theory.** |
| **Publisher** | [New York] : [Institute of Electrical and Electronics Engineers], [©1963]- |
| **Description** | volumes : illustrations, portraits ; 28-32 cm |
| **Frequency** | Bimonthly, Mar. 1968-<br>Quarterly, Jan. 1963-Jan. 1968 |
| **Pub Date** | Vol. IT-9, no. 1 (Jan. 1963)- |

| | |
|---|---|
| *Location* | Carnegie Library of Pittsburgh - Off Site Storage (Please ask for assistance) |
| *Lib Has* | Volume 9 (1963) to Volume 15 (1969). |
| *Location* | CLP Main (Oakland) Second Floor - Magazines |
| *Lib Has* | Volume 16 (1970) to Volume 54 (2008). |

**Details**

| | |
|---|---|
| **Note** | Title from cover. |
| **Indexed In:** | Applied science & technology index 0003-6986 1991-<br>Index to IEEE publications 0099-1368<br>Chemical abstracts 0009-2258<br>Computer & control abstracts 0036-8113 1968-<br>Electrical & electronics abstracts 0036-8105 1968-<br>Electronics and communications abstracts journal (Riverdale) 0361-3313<br>Engineering index annual (1968) 0360-8557<br>Engineering index bioengineering abstracts 0736-6213<br>Engineering index energy abstracts 0093-8408<br>Engineering index monthly (1984) 0742-1974<br>Excerpta medica<br>International aerospace abstracts 0020-5842<br>ISMEC bulletin 0306-0039<br>Mathematical reviews 0025-5629<br>Physics abstracts 0036-8091 1968-<br>Ship abstracts 0346-1025<br>Engineering index monthly 0013-7960<br>Nuclear science abstracts 0029-5612 |
| **Issued By** | Vols. for 1963-Apr. 1964 issued by the Professional Technical Group on Information Theory of the IEEE; July 1964- by the Information Theory Group of the IEEE; <Jan. 1968>-May 1989 by the IEEE Information Theory Group; July 1989- by the IEEE Information Theory Society. |
| **Subject** | Information theory -- Periodicals.<br>Telecommunication -- Periodicals. |
| **Addt'l Author** | Institute of Electrical and Electronics Engineers. Professional Technical Group on Information Theory.<br>Institute of Electrical and Electronics Engineers. Information Theory Group.<br>IEEE Information Theory Group.<br>IEEE Information Theory Society. |
| **Key Title** | IEEE trans. inf. theory<br>IEEE transactions on information theory |
| **Addt'l Title** | Institute of Electrical and Electronics Engineers transactions on information theory |
| **Continues** | IRE transactions on information theory |
| **ISSN** | 0018-9448 |

# APPENDIX CAIRE02


IEEE
TRANSACTIONS
ON
INFORMATION
THEORY
44
JAN-JULY
1998

CARNEGIE
LIBRARY OF
PITTSBURGH

Science and Technology Department

# Bit-Interleaved Coded Modulation

Giuseppe Caire, *Member, IEEE*, Giorgio Taricco, *Member, IEEE*, and Ezio Biglieri, *Fellow, IEEE*

*Abstract*—**It has been recently recognized by Zehavi that the performance of coded modulation over a Rayleigh fading channel can be improved by bit-wise interleaving at the encoder output, and by using an appropriate soft-decision metric as an input to a Viterbi decoder. The goal of this paper is to present in a comprehensive fashion the theory underlying bit-interleaved coded modulation, to provide tools for evaluating its performance, and to give guidelines for its design.**

*Index Terms*—**Bit-interleaving, channel capacity, coded modulation, cutoff rate, fading channel.**

## I. Introduction and Motivations

EVER since 1982, when Ungerboeck published his landmark paper on trellis-coded modulation [19], it has been generally accepted that modulation and coding should be combined in a single entity for improved performance. Of late, the increasing interest for mobile-radio channels has led to the consideration of coded modulation for fading channels. Thus at first blush it seemed quite natural to apply the same "Ungerboeck's paradigm" of keeping coding combined with modulation even in a situation (the Rayleigh fading channel) where the code performance depends strongly, rather than on the minimum Euclidean distance of the code, on its minimum Hamming distance (the "code diversity"). Several results followed this line of thought, as documented by a considerable body of work aptly summarized and referenced in [14] (see also [5, Ch. 10]). Under the assumption that the symbols were interleaved with a depth exceeding the coherence time of the fading process, new codes were designed for the fading channel so as to maximize their diversity. This implied in particular that parallel transitions should be avoided in the code, and that any increase in diversity would be obtained by increasing the constraint length of the code.

A notable departure from Ungerboeck's paradigm was the core of [24]. Schemes were designed aimed at keeping as their basic engine an off-the-shelf Viterbi decoder for the *de facto* standard, 64-state rate-$1/2$ convolutional code. This implied giving up the joint decoder/demodulator in favor of two separate entities.

Based on the latter concept, Zehavi [26] recognized that the code diversity, and hence the reliability of coded modulation over a Rayleigh fading channel, could be further improved. Zehavi's idea was to make the code diversity equal to the smallest number of distinct *bits* (rather than *channel symbols*) along any error event. This is achieved by bit-wise interleaving at the encoder output, and by using an appropriate soft-decision bit metric as an input to the Viterbi decoder. Further results along this line were recently reported in [2], [13], and [1] (for different approaches to the problem of designing coded-modulation schemes for the fading channels see [20] and [6]).

This paper is based on Zehavi's findings, and in particular on his result, rather surprising *a priori*, that on some channels there is a downside to combining demodulation and decoding. Our goal is to present in a comprehensive fashion the theory underlying bit-interleaved coded modulation (BICM) and to provide a general information-theoretical framework for this concept. This analysis also yields tools for evaluating the performance of BICM (with bounds to error probabilities tighter than those previously known) as well as guidelines for its design.

Definitions and channel model are first introduced (Section II). Next, the information-theoretical foundations of BICM are laid in Section III by evaluating in general the capacity and the cutoff rate of bit-interleaved channels. Section IV is devoted to error analysis: various approximations and bounds are introduced, and the effect of the choice of signal labeling is discussed. In particular, the asymptotic optimality of Gray labeling in conjunction with BICM is showed. Design criteria are pointed out in Section V, where a number of examples are also shown. Conclusions are summarized in Section VI, where the main themes of this paper are reprised.

## II. System Model

In this section we recall the baseline model of coded modulation (CM) and introduce the model of BICM.

Before proceeding further, let us stipulate a terminological convention: hereafter we shall use the term "bit" to denote a binary digit, and "information bit" to denote the binary information unit. Thus for example, a string of $m$ bits is a sequence of symbols 0's and 1's. A random vector with $m$ uniform independent and identically distributed (i.i.d.) components has an entropy of $m$ information bits. Moreover, lower case symbols denote scalar quantities, boldface symbols denote vectors, and underlined boldface symbols (e.g., $\underline{\boldsymbol{c}}$) denote sequences of scalars or of vectors.

The CM and BICM models are represented by the block diagram of Fig. 1. The building blocks of both schemes are 1) an encoder (ENC); 2) an interleaver $\pi$; 3) a modulator, modeled by a labeling map $\mu$ and a signal set $\mathcal{X}$, i.e., a finite set of points in the complex $N$-dimensional Euclidean space $\mathbb{C}^N$; 4) a stationary finite-memory vector channel whose transition probability density function $p_{\boldsymbol{\theta}}(\boldsymbol{y} \mid \boldsymbol{x})$, $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^N$ may depend on a vector parameter $\boldsymbol{\theta}$; 5) a demodulator (DEM), which in the present scenario plays the role of a branch metric computer; 6)

Manuscript received August 11, 1996; revised June 1, 1997. This work was supported by the Italian Space Agency (ASI).

The authors are with Politecnico di Torino, I-10129 Torino, Italy.

Publisher Item Identifier S 0018-9448(98)02360-8.

0018–9448/98$10.00 © 1998 IEEE

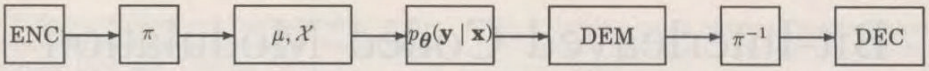



Fig. 1. Block diagram of transmission with coded modulation (CM) and bit-interleaved coded modulation (BICM). In the case of CM, $\pi$ denotes interleaving at the symbol level. In the case of BICM, $\pi$ denotes interleaving at the bit level.

a branch metric deinterleaver $\pi^{-1}$; 7) a decoder (DEC). The balance of this section is devoted to a detailed description of these blocks for CM and BICM.

### A. Vector Channel Model

We consider a vector channel characterized by a family of transition probability density functions (pdf)

$$\{p_{\boldsymbol{\theta}}(\boldsymbol{y} \mid \boldsymbol{x}) \ : \ \boldsymbol{\theta} \in \mathbb{C}^{N'}; \boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^{N}\} \tag{1}$$

parameterized by a complex-valued vector $\boldsymbol{\theta}$ representing the channel state. We assume that $\boldsymbol{\theta}$ is independent of the channel input $\boldsymbol{x}$, and that, conditionally on the sequence $\underline{\boldsymbol{\theta}}$, the channel is memoryless, i.e.,

$$p_{\underline{\boldsymbol{\theta}}}(\underline{\boldsymbol{y}} \mid \underline{\boldsymbol{x}}) = \prod_k p_{\boldsymbol{\theta}_k}(\boldsymbol{y}_k \mid \boldsymbol{x}_k). \tag{2}$$

Moreover, $\underline{\boldsymbol{\theta}}$ is assumed to be a stationary, finite memory random process. More precisely, we assume that there exists an integer $\nu \geq 0$ such that, for all $r$-tuples $\nu < k_1 < \cdots < k_r$ and for all $n$-tuples $j_1 < \cdots < j_n \leq 0$, the sequences $(\boldsymbol{\theta}_{k_1}, \cdots, \boldsymbol{\theta}_{k_r})$ and $(\boldsymbol{\theta}_{j_1}, \cdots, \boldsymbol{\theta}_{j_n})$ are statistically independent.

By this model we can represent a large number of typical communication channels such as the additive white Gaussian noise (AWGN) channel ($\boldsymbol{\theta}$ = constant), the AWGN channel with random phase ($\boldsymbol{\theta}$ is the residual phase due to imperfect carrier phase recovery) and frequency nonselective slow-fading channels [17] ($\boldsymbol{\theta}$ describes the multiplicative fading process). However, the model cannot take into account intersymbol interference (ISI), or frequency selectivity in fading channels since, in these cases, the channel state depends on the input sequence.

### B. Coded Modulation

A CM scheme can be viewed as the concatenation of an encoder ENC for a code $\mathcal{C}$ defined over a label alphabet $\mathcal{A}$ with an $N$-dimensional memoryless modulator over a signal set $\mathcal{X} \subseteq \mathbb{C}^N$ [18], [19]. We assume that $\mathcal{A}$ and $\mathcal{X}$ are finite and $|\mathcal{A}| = |\mathcal{X}| = M$. The transmitted code sequence $\underline{\boldsymbol{c}} \in \mathcal{C}$ is mapped onto the signal sequence $\underline{\boldsymbol{x}}$ by the (one-to-one) labeling map $\mu : \mathcal{A} \rightarrow \mathcal{X}$, acting componentwise.

If the code $\mathcal{C}$ is designed to correct random errors, it may be convenient to introduce a *symbol interleaver* $\pi$ between the encoder ENC and the mapper $\mu$. For most of this paper we shall assume that $\pi$ is an ideal interleaver (i.e., with infinite depth and completely random). Thus the interleaving/deinterleaving blocks can be excluded from consideration provided that the sequence $\underline{\boldsymbol{\theta}}$ is i.i.d., and hence characterized by its first-order pdf.

Let $\underline{\boldsymbol{y}} = (\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_k, \cdots)$ denote the channel output sequence resulting from the transmission of $\underline{\boldsymbol{x}}$. Optimum decoding depends on the channel state information (CSI) available at the receiver. Here we consider the two extreme cases of *perfect* CSI and no CSI.

*1) Detection with Perfect CSI:* By assuming that the receiver has perfect knowledge of the values taken on by each $\boldsymbol{\theta}_k$, the set of maximum-likelihood (ML) symbol metrics is given by

$$\{\log p_{\boldsymbol{\theta}_k}(\boldsymbol{y}_k \mid \boldsymbol{z})\}_{\boldsymbol{z} \in \mathcal{X}}. \tag{3}$$

After metric deinterleaving (denoted by $\pi^{-1}$ in Fig. 1), ML decision about the transmitted code sequence is made according to the rule

$$\hat{\underline{\boldsymbol{c}}} = \arg \max_{\underline{\boldsymbol{c}} \in \mathcal{C}} \sum_k \log p_{\boldsymbol{\theta}_k}(\boldsymbol{y}_k \mid \mu(c_k)) \tag{4}$$

where the summation runs over the whole sequence. Note that (4) is the ML decision rule also with finite-depth interleaving, or with no interleaving at all, since conditionally on $\underline{\boldsymbol{\theta}}$, the channel is memoryless (see (2)).

*2) Detection with No CSI:* The optimum receiver structure in the absence of CSI is different. A major difference is that, $\underline{\boldsymbol{\theta}}$ being unknown, the channel is not memoryless.[1] However, with ideal interleaving, for all finite-index sets $\mathcal{K}$

$$p(\underline{\boldsymbol{y}} \mid \underline{\boldsymbol{x}}) = E_{\underline{\boldsymbol{\theta}}}[p_{\underline{\boldsymbol{\theta}}}(\underline{\boldsymbol{y}} \mid \underline{\boldsymbol{x}})] = \prod_{k \in \mathcal{K}} p(\boldsymbol{y}_k \mid \boldsymbol{x}_k)$$

where we defined a new average transition pdf as

$$p(\boldsymbol{y} \mid \boldsymbol{x}) = E_{\boldsymbol{\theta}}[p_{\boldsymbol{\theta}}(\boldsymbol{y} \mid \boldsymbol{x})]. \tag{5}$$

Then, in the case of no CSI and ideal interleaving, the ML symbol metric and decision rule are obtained by using $p(\boldsymbol{y} \mid \boldsymbol{x})$ instead of $p_{\boldsymbol{\theta}}(\boldsymbol{y} \mid \boldsymbol{x})$ in (3) and in (4), respectively.

*Example 1: Ungerboeck Trellis-Coded Modulation (TCM):* The first and best known construction of CM [19] is obtained by choosing for $\mathcal{A}$ the set $\{0, 1\}^m$, for $\mathcal{C}$ a binary convolutional code of rate $k_c/m$, and for $\mu$ the *set partitioning* labeling, i.e., a binary labeling obtained by an $m$-level binary tree partition of $\mathcal{X}$ with nondecreasing minimum Euclidean distance of the signal subsets at subsequent partition levels [19], [5]. An example of set partitioning labeling for 16QAM is given in Fig. 2(a). We define the *spectral efficiency* of this construction to be $R = k_c/N$ information bits per complex dimension (bit/dim).

### C. Bit-Interleaved Coded Modulation

In general, BICM can be obtained by concatenating an encoder ENC for a binary code $\mathcal{C}$, with an $N$-dimensional memoryless modulator over a signal set $\mathcal{X} \subseteq \mathbb{C}^N$ of size $|\mathcal{X}| = M = 2^m$, through a bit interleaver $\pi$ and a one-to-one

[1] Noncoherent detection is a typical case: if the carrier phase is a slowly varying process, performance enhancement is achievable, even with no coding, by detecting blocks of signals [12].

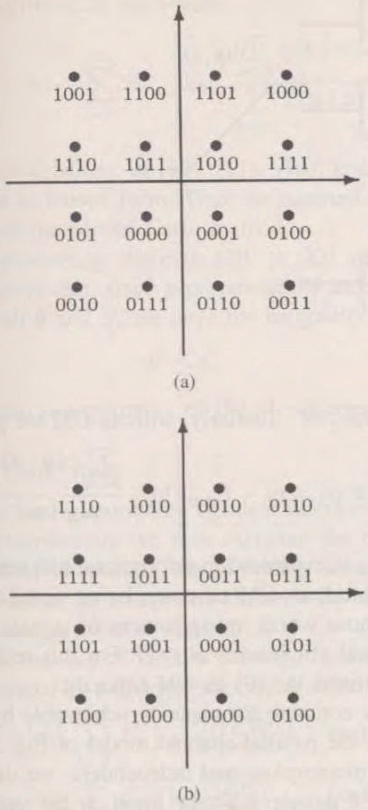



Fig. 2. 16QAM signal set with (a) a suitable set partitioning labeling and (b) Gray labeling.

binary labeling map $\mu : \{0,1\}^m \to \mathcal{X}$. The code sequence $\underline{c}$ is first interleaved by $\pi$. Next, the interleaved sequence $\pi(\underline{c})$ is broken into subsequences of $m$ bits each, which are mapped onto signals in $\mathcal{X}$. Finally, the resulting signal sequence $\underline{x}$ is transmitted over the vector channel. The bit interleaver can be seen as a one-to-one correspondence $\pi : k \to (k', i)$, where $k$ denotes the time ordering of the coded bits $c_k$, $k'$ denotes the time ordering of the signals $\boldsymbol{x}_{k'}$ transmitted over the vector channel, and $i$ indicates the position of the bit $c_k$ in the label of $\boldsymbol{x}_{k'}$. Here we assume ideal bit interleaving.

Let $\ell^i(\boldsymbol{x})$ denote the $i$th bit of the label of $\boldsymbol{x}$, and $\mathcal{X}_b^i$ the subset of all the signals $\boldsymbol{x} \in \mathcal{X}$ whose label has the value $b \in \{0,1\}$ in position $i$ (i.e., such that $\ell^i(\boldsymbol{x}) = b$.) Further, let $\bar{b}$ denote the complement of $b$ and $\boldsymbol{y}$ the channel output resulting from the transmission of $\boldsymbol{x}$.

The conditional pdf of $\boldsymbol{y}$, given $\ell^i(\boldsymbol{x}) = b \in \{0,1\}$ and $\boldsymbol{\theta}$, is

$$p_{\boldsymbol{\theta}}(\boldsymbol{y} \mid \ell^i(\boldsymbol{x}) = b) = \sum_{\boldsymbol{x} \in \mathcal{X}} p_{\boldsymbol{\theta}}(\boldsymbol{y} \mid \boldsymbol{x}) P(\boldsymbol{x} \mid \ell^i(\boldsymbol{x}) = b)$$
$$= 2^{-(m-1)} \sum_{\boldsymbol{x} \in \mathcal{X}_b^i} p_{\boldsymbol{\theta}}(\boldsymbol{y} \mid \boldsymbol{x}). \quad (6)$$

The last expression is obtained by noting that

$$p_{\boldsymbol{\theta}}(\boldsymbol{y} \mid \boldsymbol{x}, \ell^i(\boldsymbol{x}) = b) = p_{\boldsymbol{\theta}}(\boldsymbol{y} \mid \boldsymbol{x})$$

and that

$$P(\boldsymbol{x} \mid \ell^i(\boldsymbol{x}) = b) = \begin{cases} 0, & \text{if } \ell^i(\boldsymbol{x}) = \bar{b} \\ 2^{-(m-1)}, & \text{if } \ell^i(\boldsymbol{x}) = b \end{cases}$$

where we assume uniform input distribution

$$P(b = 0) = P(b = 1) = 1/2.$$

Again, the optimum receiver depends on the available CSI. For both cases of ideal CSI and no CSI with ideal bit interleaving, as we consider in this paper, at each time $k'$ the demodulator DEM produces the set of ML *bit* metrics

$$\lambda^i(\boldsymbol{y}_{k'}, b) = \begin{cases} \log \sum_{\boldsymbol{x} \in \mathcal{X}_b^i} p_{\boldsymbol{\theta}_{k'}}(\boldsymbol{y}_{k'} \mid \boldsymbol{x}), & \text{perfect CSI} \\ \log \sum_{\boldsymbol{x} \in \mathcal{X}_b^i} p(\boldsymbol{y}_{k'} \mid \boldsymbol{x}), & \text{no CSI} \end{cases} \quad (7)$$

for $b \in \{0,1\}$ and $i = 1, \cdots, m$. Finally, the ML decoder DEC makes decisions according to the rule

$$\hat{\underline{c}} = \arg \max_{\underline{c} \in \mathcal{C}} \sum_k \lambda^i(\boldsymbol{y}_{k'}, c_k). \quad (8)$$

The decision rule (8) is completely analogous to (4), and can be implemented by a Viterbi decoder.

*1) Simplified Branch Metrics:* The branch metrics (7) may be computationally too complex for implementation. Suboptimal simplified branch metrics [26] can be obtained by the log-sum approximation $\log \sum_j z_j \simeq \max_j \log z_j$, which is good as long as the sum in the left-hand side is dominated by a single term, as typically occurs in channels with high signal-to-noise ratio (SNR). This yields

$$\tilde{\lambda}^i(\boldsymbol{y}_{k'}, b) = \begin{cases} \max_{\boldsymbol{x} \in \mathcal{X}_b^i} \log p_{\boldsymbol{\theta}_{k'}}(\boldsymbol{y}_{k'} \mid \boldsymbol{x}), & \text{perfect CSI} \\ \max_{\boldsymbol{x} \in \mathcal{X}_b^i} \log p(\boldsymbol{y}_{k'} \mid \boldsymbol{x}), & \text{no CSI} \end{cases} \quad (9)$$

Branch metrics (9) are obtained from the set of symbol branch metrics (3) by selecting the maximum over the subsets $\mathcal{X}_b^i$. BICM branch metric computation can be implemented by the same demodulator for CM followed by some post-processor (namely, by a parallel bank of comparators). Hence, BICM metric computation has the same latency of CM, but requires $m(2^m - 2)$ additional binary comparisons.

*2) Gray Labeling:* Gray labeling plays a key role in BICM theory. For convenience, we restate here its definition in terms of BICM notation.

*Definition. Gray Labeling:* Let $\mathcal{X}$ denote a signal set of size $M = 2^m$, with minimum Euclidean distance $d_{\min}$. A binary map $\mu : \{0,1\}^m \to \mathcal{X}$ is a Gray labeling for $\mathcal{X}$ if, for all $i = 1, \cdots, m$ and $b \in \{0,1\}$, each $\boldsymbol{x} \in \mathcal{X}_b^i$ has at most one $\boldsymbol{z} \in \mathcal{X}_b^i$ at distance $d_{\min}$.

Fig. 2(b) shows a Gray-labeled 16QAM signal set.

*3) Equivalent Channel Model:* Fig. 3 shows the model of a BICM communication channel with ideal interleaving. The system consists of a set of $m$ parallel *independent* and *memoryless* binary input channels connected to the encoder output by a random switch, which models ideal interleaving. Each channel corresponds to a position in the label of the signals of $\mathcal{X}$. For every symbol $c_k$ of a coded sequence $\underline{c}$, the switch selects randomly (and independently of other selections) a position index $i \in \{1, \cdots, m\}$, and transmits $c_k$ on the $i$th channel. The decoder, which knows the sequence of channels used for

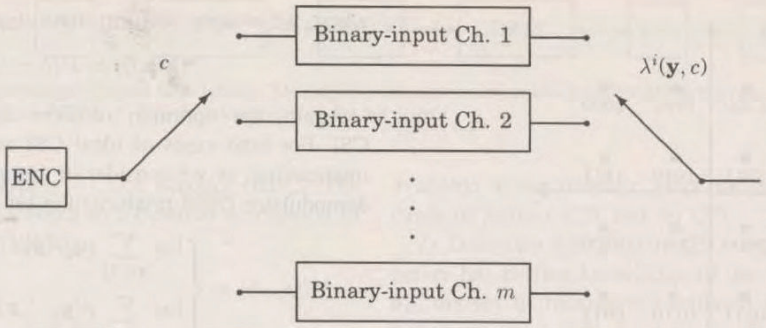



Fig. 3. Equivalent parallel channel model for BICM in the case of ideal interleaving.

the transmission of $\underline{c}$, computes the branch metrics (7) (or (9)) and selects $\hat{\underline{c}}$ according to (8).

*Example 2. Zehavi's 8PSK Code:* The effectiveness of BICM schemes over the Rayleigh fading channel was first pointed out by Zehavi in [26]. He picked for $\mathcal{C}$ the best 8-state, rate-$2/3$ binary convolutional code [11], and for $\mu$ a Gray-labeling map. Unlike with our BICM model, the output bits from the convolutional encoder were passed through three separate ideal interleavers, so that there was a fixed correspondence between the output bits of the encoder and the label positions. Apparently, there are no reasons justifying this fixed correspondence which, on the other hand, limits the flexibility of BICM and complicates the analysis. Moreover, a fixed correspondence between coded bits and label positions introduces unequal error protection, usually an undesired feature, and suboptimal performance when $\mathcal{C}$ is chosen at random. This motivates our BICM model: here all the coded bits output by the encoder are fed into a single-bit interleaver, and thus being on an equal footing, generate a maximum of symmetry.

## III. An Information-Theoretical View of BICM

In this section we compute the capacity and the cutoff rate of CM and BICM with ideal interleaving, under the constraint of uniform input probabilities. As usual, we assume ideal interleaving, so that the sequence of channel state parameters $\boldsymbol{\theta}$ is i.i.d.

### A. Channel Capacity

Consider the memoryless discrete-input, continuous-output channel with input $\boldsymbol{x}$, output $\boldsymbol{y}$, and transition distribution $p_{\boldsymbol{\theta}}(\boldsymbol{y} \mid \boldsymbol{x})$. Under the assumption $I(\boldsymbol{x}; \boldsymbol{\theta}) = 0$, from the "chain rule" of mutual information [10] we obtain the inequality $I(\boldsymbol{x}; \boldsymbol{y}) \leq I(\boldsymbol{x}; \boldsymbol{y} \mid \boldsymbol{\theta})$. The right-hand- and left-hand-side terms of this inequality are attained when the receiver has perfect CSI and no CSI, respectively. Then, the capacity under uniform inputs constraint and perfect CSI is given by the conditional average mutual information (AMI)

$$\tilde{C} = I(\boldsymbol{x}; \boldsymbol{y} \mid \boldsymbol{\theta}) = m - E_{\boldsymbol{x},\boldsymbol{y},\boldsymbol{\theta}} \left[ \log_2 \frac{\sum_{\boldsymbol{z} \in \mathcal{X}} p_{\boldsymbol{\theta}}(\boldsymbol{y} \mid \boldsymbol{z})}{p_{\boldsymbol{\theta}}(\boldsymbol{y} \mid \boldsymbol{x})} \right] \tag{10}$$

where $m = \log_2 M$. Similarly, without CSI we get

$$\tilde{C} = I(\boldsymbol{x}; \boldsymbol{y}) = m - E_{\boldsymbol{x},\boldsymbol{y}} \left[ \log_2 \frac{\sum_{\boldsymbol{z} \in \mathcal{X}} p(\boldsymbol{y} \mid \boldsymbol{z})}{p(\boldsymbol{y} \mid \boldsymbol{x})} \right]. \tag{11}$$

Here, capacity is expressed in information bits per $N$ complex dimensions (bit/dim). CM schemes, being signal space codes, i.e., codes whose words are sequences of signals $\boldsymbol{x} \in \mathcal{X}$, can achieve spectral efficiencies $R \leq \tilde{C}$. For this reason, we will refer to $\tilde{C}$ defined in (10) as *CM capacity*.

Let us now compute the capacity achievable by BICM. To do it, we use the parallel-channel model of Fig. 3. Since the channels are memoryless and independent, we drop the time index $k$. Let $b$ denote a binary input, $\boldsymbol{y}$ the vector channel output, and $S$ the random variable whose outcome determines the switch position (for the above, $S$ is i.i.d., uniformly distributed over $\{1, \cdots, m\}$, and known to the receiver). Since $S$, $\boldsymbol{\theta}$, and $b$ are independent, we can show that the AMI of $b$ and $\boldsymbol{y}$ satisfies the inequality $I(b; \boldsymbol{y} \mid S) \leq I(b; \boldsymbol{y} \mid \boldsymbol{\theta}, S)$. Again, these mutual informations are attained with no CSI and with perfect CSI, respectively. The conditional mutual information of $b$ and $\boldsymbol{y}$, given $S = i$, is then given by

$$I(b; \boldsymbol{y} \mid \boldsymbol{\theta}, S = i) = 1 - E_{b,\boldsymbol{y},\boldsymbol{\theta}} \left[ \log_2 \frac{\sum_{\boldsymbol{z} \in \mathcal{X}} p_{\boldsymbol{\theta}}(\boldsymbol{y} \mid \boldsymbol{z})}{\sum_{\boldsymbol{z} \in \mathcal{X}_b^i} p_{\boldsymbol{\theta}}(\boldsymbol{y} \mid \boldsymbol{z})} \right] \tag{12}$$

where $b$, $\boldsymbol{y}$, and $\boldsymbol{\theta}$ are conditionally jointly distributed as

$$p(b, \boldsymbol{y}, \boldsymbol{\theta} \mid S = i) = 2^{-m} \sum_{\boldsymbol{z} \in \mathcal{X}_b^i} p_{\boldsymbol{\theta}}(\boldsymbol{y} \mid \boldsymbol{z}) p_{\boldsymbol{\theta}}(\boldsymbol{\theta}). \tag{13}$$

The conditional AMI is obtained by averaging (12) with respect to $S$, so that

$$I(b; \boldsymbol{y} \mid \boldsymbol{\theta}, S) = \frac{1}{m} \sum_{i=1}^{m} I(b; \boldsymbol{y} \mid \boldsymbol{\theta}, S = i).$$

Finally, since there are $m$ parallel independent channels, the BICM capacity, with perfect CSI and uniform inputs, is given by

$$\begin{aligned} \hat{C} &= m I(b; \boldsymbol{y} \mid \boldsymbol{\theta}, S) \\ &= m - \sum_{i=1}^{m} E_{b,\boldsymbol{y},\boldsymbol{\theta}} \left[ \log_2 \frac{\sum_{\boldsymbol{z} \in \mathcal{X}} p_{\boldsymbol{\theta}}(\boldsymbol{y} \mid \boldsymbol{z})}{\sum_{\boldsymbol{z} \in \mathcal{X}_b^i} p_{\boldsymbol{\theta}}(\boldsymbol{y} \mid \boldsymbol{z})} \right]. \end{aligned} \tag{14}$$

Similarly, with no CSI we obtain

$$\hat{C} = m - \sum_{i=1}^{m} E_{b,\boldsymbol{y}}\left[\log_2 \frac{\sum_{\boldsymbol{z}\in\mathcal{X}} p(\boldsymbol{y}\mid\boldsymbol{z})}{\sum_{\boldsymbol{z}\in\mathcal{X}_b^i} p(\boldsymbol{y}\mid\boldsymbol{z})}\right]. \tag{15}$$

In general, expectations in (10), (11), (14), and (15) cannot be calculated in closed form. Thus we resorted to numerical integration via the Monte Carlo method.

The data-processing theorem [10, p. 32] can finally be invoked to prove that, since $\boldsymbol{y}$ is statistically independent of $b$ given $\boldsymbol{x}$ for all $\boldsymbol{\theta}$ and $S$, we have the inequality

$$\hat{C} \le \tilde{C}. \tag{16}$$

This shows the suboptimality of BICM schemes.

### B. Channel Cutoff Rate

Under the usual assumption of ideal interleaving and uniform input distribution, we now compute the cutoff rate of CM and BICM channels. This is widely accepted to be a sensible parameter for comparing channels for which a finite-complexity coding scheme is used.

The cutoff rate $\tilde{R}_0$ of the discrete-input continuous-output channel generated by a CM scheme can be written as [23], [14]

$$\tilde{R}_0 = \begin{cases} m - \log_2 \sum_{\boldsymbol{z}\in\mathcal{X}} E_{\boldsymbol{x},\boldsymbol{y},\boldsymbol{\theta}}\left[\sqrt{\frac{p_{\boldsymbol{\theta}}(\boldsymbol{y}\mid\boldsymbol{z})}{p_{\boldsymbol{\theta}}(\boldsymbol{y}\mid\boldsymbol{x})}}\right], & \text{perfect CSI} \\ m - \log_2 \sum_{\boldsymbol{z}\in\mathcal{X}} E_{\boldsymbol{x},\boldsymbol{y}}\left[\sqrt{\frac{p(\boldsymbol{y}\mid\boldsymbol{z})}{p(\boldsymbol{y}\mid\boldsymbol{x})}}\right], & \text{no CSI.} \end{cases} \tag{17}$$

The cutoff rate of a BICM scheme can be obtained from the Bhattacharyya bound [23] on the average bit-error probability $P_b$ of the parallel channel model of Fig. 3 in the absence of coding. By considering the ML bit metrics with perfect CSI given in (7), we obtain

$$P_b \le \frac{1}{m}\sum_{i=1}^{m} E_{b,\boldsymbol{y},\boldsymbol{\theta}}\left[\sqrt{\frac{p_{\boldsymbol{\theta}}(\boldsymbol{y}\mid\bar{b},i)}{p_{\boldsymbol{\theta}}(\boldsymbol{y}\mid b,i)}}\right] = B \tag{18}$$

where $B$ denotes the average Bhattacharyya factor of the BICM channel, with perfect CSI. Hence, we obtain from (6)

$$B \triangleq \frac{1}{m}\sum_{i=1}^{m} E_{b,\boldsymbol{y},\boldsymbol{\theta}}\left[\sqrt{\frac{\sum_{\boldsymbol{z}\in\mathcal{X}_{\bar{b}}^i} p_{\boldsymbol{\theta}}(\boldsymbol{y}\mid\boldsymbol{z})}{\sum_{\boldsymbol{z}\in\mathcal{X}_b^i} p_{\boldsymbol{\theta}}(\boldsymbol{y}\mid\boldsymbol{z})}}\right]. \tag{19}$$

Similarly, by considering the ML bit metrics without CSI given in (7), we get the average Bhattacharyya factor

$$B \triangleq \frac{1}{m}\sum_{i=1}^{m} E_{b,\boldsymbol{y}}\left[\sqrt{\frac{\sum_{\boldsymbol{z}\in\mathcal{X}_{\bar{b}}^i} p(\boldsymbol{y}\mid\boldsymbol{z})}{\sum_{\boldsymbol{z}\in\mathcal{X}_b^i} p(\boldsymbol{y}\mid\boldsymbol{z})}}\right]. \tag{20}$$

Finally, note that a channel use of the BICM channel is equivalent to $m$ channel uses of a binary-input channel with average Bhattacharyya factor $B$. Hence, the cutoff rate $\hat{R}_0$ of BICM can be written as

$$\hat{R}_0 = m(1 - \log_2(B+1)) \tag{21}$$

where $B$ is given by (19) with perfect CSI or by (20) with no CSI. In spite of the fact that $\tilde{R}_0$ can be evaluated analytically in many cases [14], $B$, and hence $\hat{R}_0$, cannot. Again, as we did for capacity, we resorted to Monte Carlo numerical integration to calculate $\hat{R}_0$.

### C. Numerical Calculations for Fading Channels

While the results obtained in the previous section hold in general, hereafter we restrict our attention the slow frequency nonselective Rician fading channels [14], which encompass Rayleigh and AWGN as special cases. With this model, the received signal after the transmission of $\boldsymbol{x} \in \mathcal{X}$ can be written as

$$\boldsymbol{y} = g\boldsymbol{x}e^{j\phi} + \boldsymbol{n} \tag{22}$$

where: i) $\boldsymbol{n}$ is a complex zero-mean Gaussian i.i.d. random vector with covariance $\frac{1}{2}\mathrm{E}[\boldsymbol{n}\boldsymbol{n}^\dagger] = \frac{1}{2}\sigma^2\boldsymbol{I}_N$ ($\boldsymbol{I}_N$ is the $N \times N$ identity matrix, and a dagger $\dagger$ denotes Hermitian transpose); ii) $g$ is a scalar complex fading gain;[2] iii) $\phi$ is the carrier phase, independent of $\boldsymbol{x}$ and $g$ and uniformly distributed over $[-\pi, \pi]$.

With Rician fading, $g$ is complex Gaussian with mean $\sqrt{K/(K+1)}$ and variance $1/(2(K+1))$ (where $K \ge 0$ is the *Rician parameter*, defined as the ratio of direct to diffuse power in the received signal [14]). We assume normalized signal energy $\frac{1}{2}E[|\boldsymbol{x}|^2] = 1/2$, so that the average SNR is $1/\sigma^2$. With this model $\boldsymbol{\theta} = (g, \phi)$ and $p_{\boldsymbol{\theta}}(\boldsymbol{y}\mid\boldsymbol{x})$ is given by

$$p_{\boldsymbol{\theta}}(\boldsymbol{y}\mid\boldsymbol{x}) = \frac{1}{(\pi\sigma^2)^N}\exp\left(-|\boldsymbol{y} - g\boldsymbol{x}e^{j\phi}|^2/\sigma^2\right). \tag{23}$$

By averaging with respect to $g$ and $\phi$, we obtain [12]

$$\begin{aligned} p(\boldsymbol{y}\mid\boldsymbol{x}) &= E_{\boldsymbol{\theta}}[p_{\boldsymbol{\theta}}(\boldsymbol{y}\mid\boldsymbol{x})] \\ &= \frac{1}{\pi^N|\Sigma_y|}\exp\left(-\boldsymbol{y}^\dagger\Sigma_y^{-1}\boldsymbol{y} - \bar{\boldsymbol{y}}^\dagger\Sigma_y^{-1}\bar{\boldsymbol{y}}\right) I_0\left(2\left|\bar{\boldsymbol{y}}^\dagger\Sigma_y^{-1}\boldsymbol{y}\right|\right) \end{aligned} \tag{24}$$

where we have defined

$$\bar{\boldsymbol{y}} = \sqrt{\frac{K}{K+1}}\boldsymbol{x}, \quad \text{and} \quad \Sigma_y = \frac{1}{K+1}\boldsymbol{x}\boldsymbol{x}^\dagger + \sigma^2\boldsymbol{I}_N$$

and where $I_0(\cdot)$ is the modified Bessel function of the first kind and order zero [17].

Numerical results can be obtained by using (23) and (24) in the general expressions of $\tilde{C}$, $\hat{C}$, $\tilde{R}_0$, and $\hat{R}_0$, for any choice of signal set and labeling map. Here we restrict ourselves to the limiting cases of AWGN and Rayleigh fading channels ($K \to \infty$ and $K = 0$, respectively). While looking at the following results, the reader should keep in mind that *bit/dim* indicates information bits per *complex dimension*.

*1) PSK and QAM with Perfect CSI:* Figs. 4 and 5 show CM and BICM capacity versus SNR for 4PSK, 8PSK, and 16QAM over the AWGN and Rayleigh fading channels, respectively. BICM capacity is computed with Gray and set partitioning (SP) labelings. The performance of BICM with Gray labeling is very close to that of CM (actually,

[2] This model does not apply when component interleaving is used in addition to bit interleaving, since in this case the $N$ components of $\boldsymbol{x}$ are affected by different fading coefficients [6], [1].

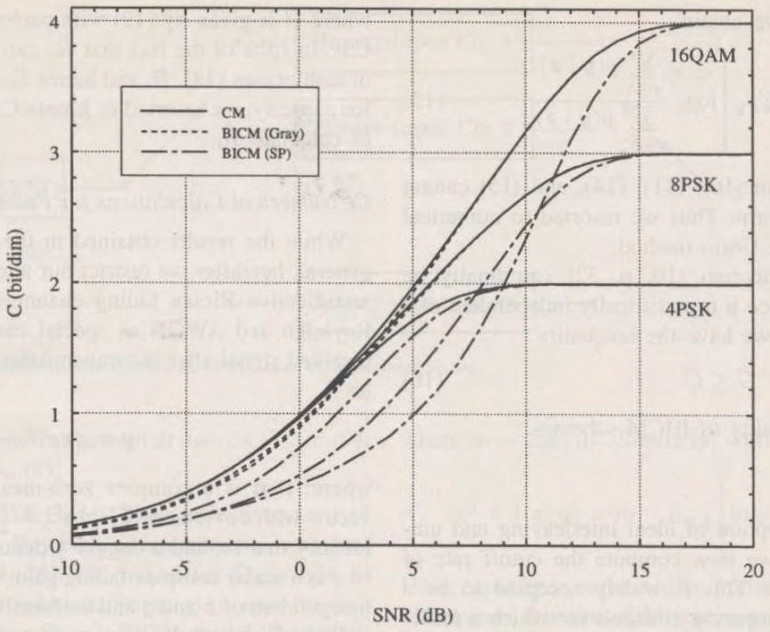



Fig. 4. BICM and CM capacity versus. SNR for 4PSK, 8PSK, and 16QAM over AWGN with coherent detection (SP denotesset-partitioning labeling).

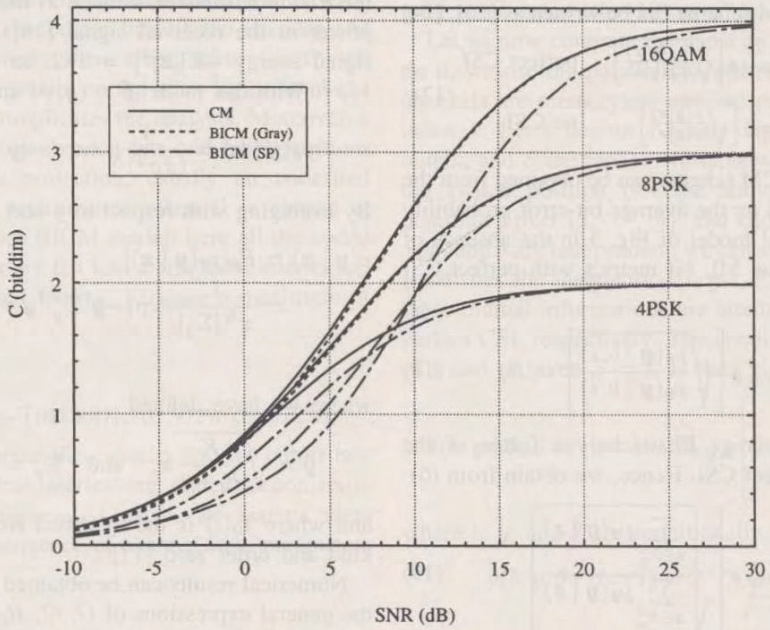



Fig. 5. BICM and CM capacity versus SNR for 4PSK, 8PSK, and 16QAM over Rayleigh fading with coherent detection and perfect CSI (SP denotes set-partitioning labeling).

we conjecture that Gray labeling maximizes BICM capacity) while the performance of BICM with SP labeling is several decibels worse. Similar differences between Gray and SP labelings can also be observed from the cutoff rate of the Rayleigh fading channel.

Our next results are based on cutoff rate. This parameter appears to be more suitable than capacity to compare BICM and CM, possibly because there is no fixed relation between $\hat{R}_0$ and $\tilde{R}_0$ whereas $\hat{C} \leq \tilde{C}$, as we know from (16). Figs. 6 and 7 show BICM and CM cutoff rate versus SNR for QAM signal sets, over the AWGN and Rayleigh fading channels, respectively. For 4, 16, 64, and 256QAM signal sets we used Gray labeling. For 8, 32, and 128QAM—for which Gray labeling is not possible—we used a *quasi*-Gray labeling, i.e., a labeling minimizing the number of signals for which the Gray condition (of having at most one nearest neighbor in the complement subset) is not satisfied. For AWGN, CM outperforms BICM at all SNR. The performance gap, which is large for large $M$ and low-rate codes, is reduced for high-rate codes. For example, CM256 QAM gains more than 3

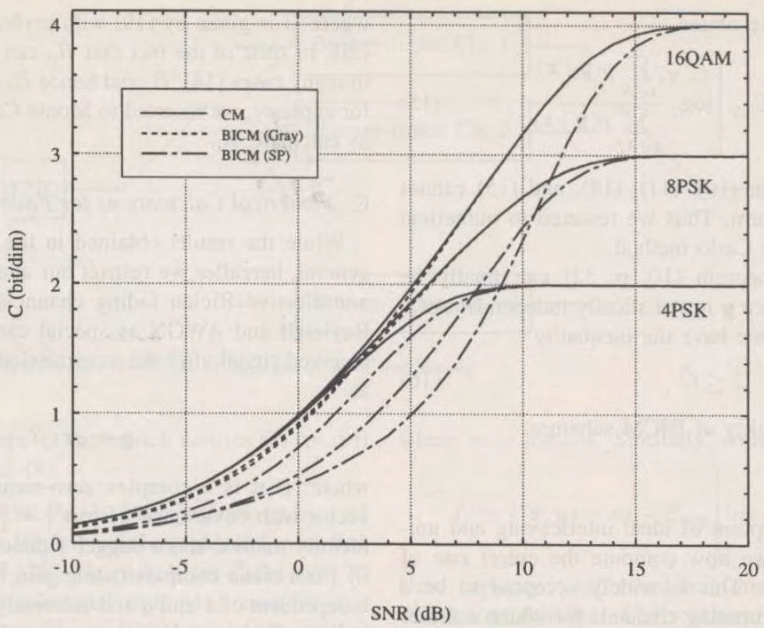



Fig. 4. BICM and CM capacity versus. SNR for 4PSK, 8PSK, and 16QAM over AWGN with coherent detection (SP denotesset-partitioning labeling).

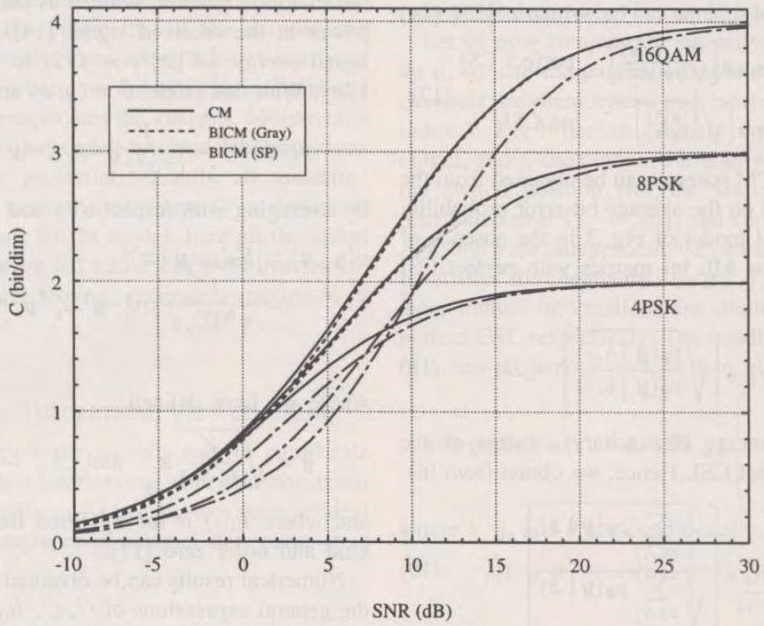



Fig. 5. BICM and CM capacity versus SNR for 4PSK, 8PSK, and 16QAM over Rayleigh fading with coherent detection and perfect CSI (SP denotes set-partitioning labeling).

we conjecture that Gray labeling maximizes BICM capacity) while the performance of BICM with SP labeling is several decibels worse. Similar differences between Gray and SP labelings can also be observed from the cutoff rate of the Rayleigh fading channel.

Our next results are based on cutoff rate. This parameter appears to be more suitable than capacity to compare BICM and CM, possibly because there is no fixed relation between $\hat{R}_0$ and $\tilde{R}_0$ whereas $\hat{C} \leq \tilde{C}$, as we know from (16). Figs. 6 and 7 show BICM and CM cutoff rate versus SNR for QAM signal sets, over the AWGN and Rayleigh fading channels, respectively. For 4, 16, 64, and 256QAM signal sets we used Gray labeling. For 8, 32, and 128QAM—for which Gray labeling is not possible—we used a *quasi*-Gray labeling, i.e., a labeling minimizing the number of signals for which the Gray condition (of having at most one nearest neighbor in the complement subset) is not satisfied. For AWGN, CM outperforms BICM at all SNR. The performance gap, which is large for large $M$ and low-rate codes, is reduced for high-rate codes. For example, CM256 QAM gains more than 3

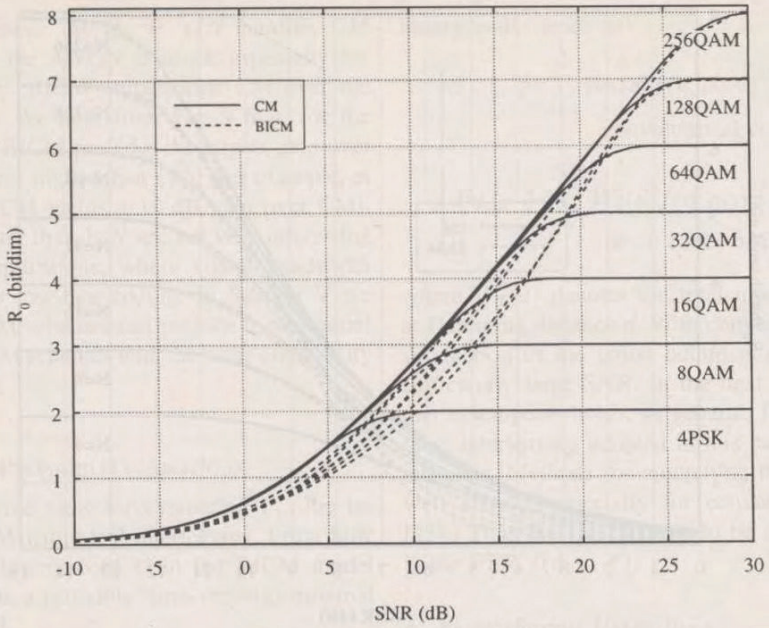



Fig. 6. BICM and CM cutoff rate versus SNR for QAM signal sets with Gray (or quasi-Gray) labeling over AWGN with coherent detection.

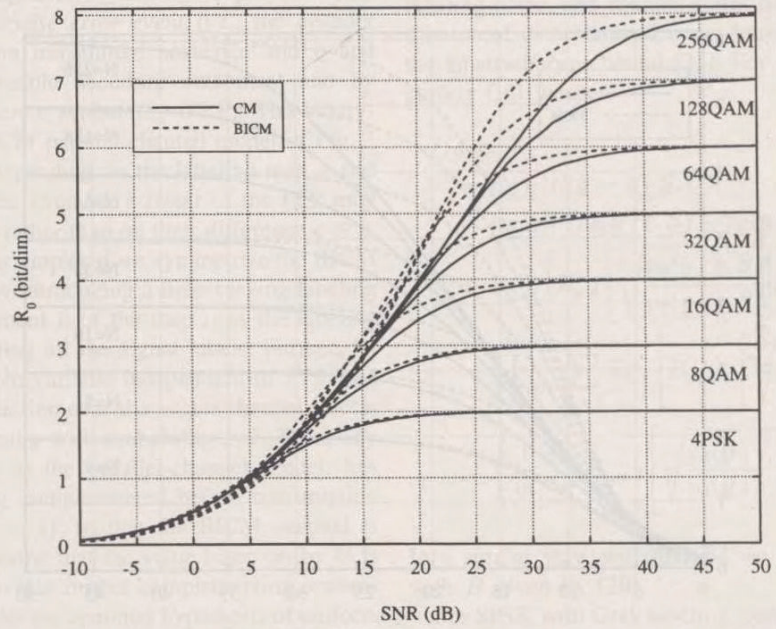



Fig. 7. BICM and CM cutoff rate versus SNR for QAM signal sets with Gray (or quasi-Gray) labeling over Rayleigh fading with coherent detection and perfect CSI.

dB over BICM 256QAM at $R_0 = 3$ bit/dim, but this gap is reduced to less than 0.5 dB at $R_0 = 7$ bit/dim. The situation is different for the Rayleigh channel. We observe that in this case, BICM generally outperforms CM for $R_0 > 1$ bit/dim. This difference in performance is especially apparent for high-rate codes: for example, BICM 256QAM gains about 4 dB over CM 256QAM at $R_0 = 7$ bit/dim. The above fact can be intuitively explained as follows. For low rates it is possible to design "practical" CM codes with large $M$-ary Hamming distance, so that their performance over the Rayleigh fading channel is good. For high rates, the complexity required by CM in order to obtain large $M$-ary Hamming distance is much larger than the complexity required by BICM, for which Hamming distance is given by the binary Hamming distance of the underlying binary code. Hence, for a given complexity, BICM compares favorably with respect to CM on the Rayleigh fading channel, especially for high rates.

Capacity and cutoff rate curves provide guidelines for code design. In particular, over the AWGN channel CM appears more suitable, although for high rates the loss of optimality

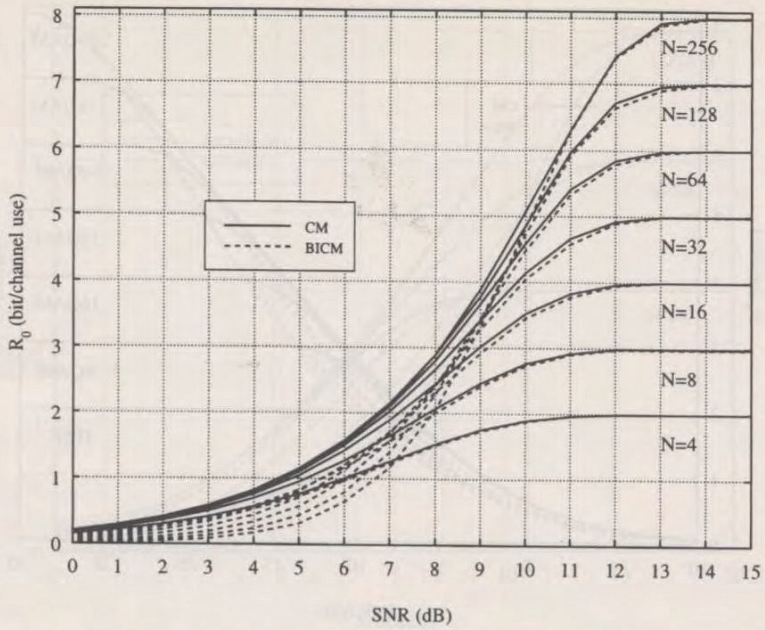



Fig. 8. BICM and CM cutoff rate versus. SNR for orthogonal signal sets over AWGN with noncoherent detection. The cutoff rate is expressed in information bit per channel use, where a channel use corresponds to $N$ complex dimensions.

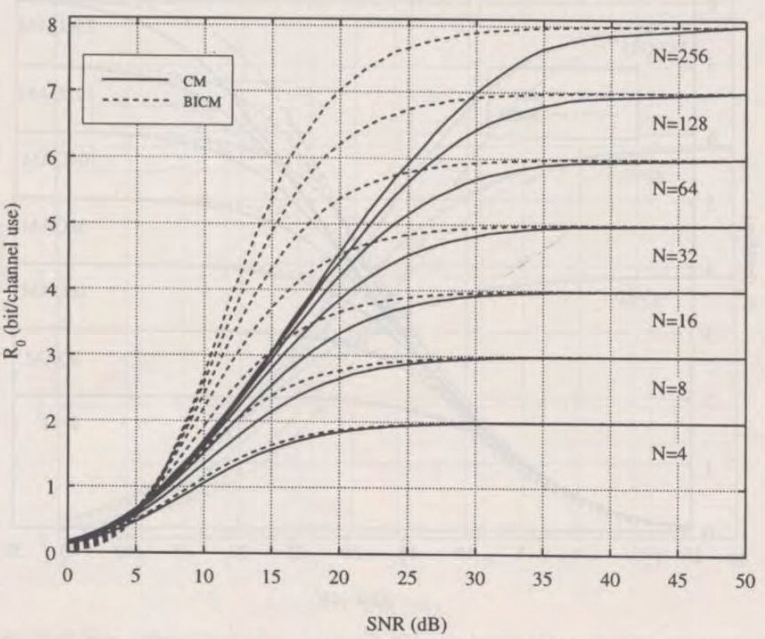



Fig. 9. BICM and CM cutoff rate versus SNR for orthogonal signal sets over Rayleigh fading with noncoherent detection and no CSI. The cutoff rate is expressed in information bit per channel use, where a channel use corresponds to $N$ complex dimensions.

introduced by BICM is marginal (thus leaving room for "pragmatic" approaches [24]). On the contrary, BICM is much more appropriate for the Rayleigh fading channel. As a consequence, if the channel model—as is the case for example for mobile radio—fluctuates in time between the extremes of Rayleigh and AWGN, BICM proves to be a more robust choice than CM.

*2) $N$-ary Orthogonal Signals with No CSI:* We consider a unit-energy $N$-ary orthogonal signal set $\mathcal{X}$ with $N = 2^m$ (the sequences $\underline{x} \in \mathcal{X}$ can be obtained, for example, as unit-energy Hadamard sequences [22]). Detection is noncoherent with no CSI. Given the symmetry of orthogonal signals, $\hat{C}$ and $\hat{R}_0$ do not depend on the labeling $\mu$, which can be arbitrarily chosen.

Again, we use the cutoff rate to compare BICM and CM. Figs. 8 and 9 show BICM and CM cutoff rates for $N = 4, 8, 16, 32, 64, 128$, and 256 over AWGN and Rayleigh fading channels, respectively. An application might be coded DS/SSMA, for which orthogonal signals are used to obtain

very low rate CM schemes.[3] At $R_0 = 1/N$ bit/dim, CM outperforms BICM over the AWGN channel, especially for large $N$. On the contrary, BICM outperforms CM over the Rayleigh fading channel. We note from Fig. 9 that, for the Rayleigh fading channel, BICM provides the largest gain over CM for spectral efficiencies higher than $1/N$ (for example, at $R_0 = 7/256$ bit/dim, BICM yields a 10-dB gain over CM). However, efficiencies better than $1/N$ are not very interesting for coded DS/SSMA applications, where signal bandwidth expansion is obtained by low-rate coding. In Section V we shall see that actual BICM schemes can provide a substantial coding gain over actual CM schemes with the same complexity also at rate $1/N$ bit/dim.

## IV. Error-Probability Analysis

In this section we derive tools for evaluating $P_b$, the bit error rate (BER) of BICM with ideal interleaving. From now on we assume that the binary code $\mathcal{C}$ in the BICM model is linear, so that it admits a (possibly time-varying) minimal trellis representation [18].

Let $\underline{c}$ and $\hat{\underline{c}}$ denote two distinct sequences stemming from the same state and merging after $\ell \geq 1$ trellis steps. Denote by $\{\underline{c} \to \hat{\underline{c}}\}$ a decoder *pairwise error event* (i.e., the decoder chooses $\hat{\underline{c}}$ when $\underline{c}$ is the transmitted sequence and $\underline{c}$ and $\hat{\underline{c}}$ are the only two possible decoding outcomes) and by $P(\underline{c} \to \hat{\underline{c}})$ the *pairwise error probability* (PEP). The binary-input channels of the BICM parallel channel model of Fig. 3 may be nonsymmetric, depending on the labeling map $\mu$ and on the signal set $\mathcal{X}$. Then, even for a linear $\mathcal{C}$, the PEP may depend on the pair $(\underline{c}, \hat{\underline{c}})$ rather than on their difference $\underline{c} \oplus \hat{\underline{c}}$. The analysis can be made simpler if we symmetrize the BICM parallel-channel model by introducing a time-varying labeling map. Define the complement $\bar{\mu}$ of the map $\mu$ as the labeling obtained by complementing all the signal labels. Further, let $U$ denote a binary random variable independent of $S$, $b$, and $\underline{\theta}$, which for each transmission over the vector channel selects either $\mu$ or $\bar{\mu}$, independently with probability $1/2$. Then, any binary symbol $b$, input to the parallel-channel model, has probability $1/2$ of being complemented before transmission (which occurs when $U = 1$), so that the BICM channel is made symmetric. We assume that the value taken on by $U$ is known to the receiver, so that further complementing restores the actual value of $b$. Under the common hypothesis of uniform input bits, performance is not degraded by symmetrization. The PEP of symmetrized BICM, with ideal interleaving, is a function of the Hamming distance $d$ between $\underline{c}$ and $\hat{\underline{c}}$, of the labeling $\mu$ and of the signal set $\mathcal{X}$. Therefore, we can write

$$P(\underline{c} \to \hat{\underline{c}}) = f(d, \mu, \mathcal{X}) \tag{25}$$

where the function $f(\cdot, \cdot, \cdot)$ depends on the vector channel and on the type of detection. Without loss of generality, in the following we assume that $\underline{c}$ and $\hat{\underline{c}}$ differ in the $d$ consecutive positions $1, \cdots, d$. Finally, the usual union bound for linear binary codes leads to

$$\begin{cases} P_b \leq \frac{1}{k_c} \sum_{d=1}^{\infty} W_I(d) f(d, \mu, \mathcal{X}) \\ \qquad\qquad \text{convolutional codes of rate } k_c/n_c \\ P_b \leq \frac{1}{K_c} \sum_{d=1}^{N_c} W_I(d) f(d, \mu, \mathcal{X}) \\ \qquad\qquad \text{block codes of rate } K_c/N_c \end{cases} \tag{26}$$

where $W_I(d)$ denotes the total input weight of error events at Hamming distance $d$. With convolutional codes, the infinite summation of the union bound is assumed to converge, for sufficiently large SNR. In the next section we shall see that this assumption holds, in general, for the BICM model with ideal interleaving adopted in this paper and for well-behaved channels. Methods for computing the coefficients $W_I(d)$ are well known, especially for convolutional codes (see, e.g., [23].) Thus the key quantity to be evaluated in the following is the PEP $f(d, \mu, \mathcal{X})$.

### A. Bhattacharyya Union Bound

Let $\underline{S}$, $\underline{U}$, and $\underline{\theta}$ denote the sequences of label positions, labeling maps, and channel states occurring during the transmission of $\underline{c}$. A standard upper bound on the PEP is through the Bhattacharyya bound [23]. For an ML decoder (8) with perfect CSI we get

$$\begin{aligned} & f(d, \mu, \mathcal{X}) \\ & \quad = E_{\underline{S},\underline{U}}[P(\underline{c} \to \hat{\underline{c}} \mid \underline{S}, \underline{U})] \\ & \quad = E_{\underline{S},\underline{U}}[P(p_{\underline{\theta}}(\underline{y} \mid \underline{c}, \underline{S}) \leq p_{\underline{\theta}}(\underline{y} \mid \hat{\underline{c}}, \underline{S}) \mid \underline{S}, \underline{U})] \\ & \quad \leq E_{\underline{S},\underline{U}}\left[ E_{\underline{y},\underline{\theta}}\left[ \sqrt{\frac{p_{\underline{\theta}}(\underline{y} \mid \hat{\underline{c}}, \underline{S})}{p_{\underline{\theta}}(\underline{y} \mid \underline{c}, \underline{S})}} \,\middle|\, \underline{S}, \underline{U} \right] \right] \\ & \quad = \prod_{k=1}^{d} \frac{1}{m} \sum_{i_k=1}^{m} E_{U_k}\left[ E_{\boldsymbol{y}_k, \boldsymbol{\theta}_k}\left[ \sqrt{\frac{p_{\boldsymbol{\theta}_k}(\boldsymbol{y}_k \mid \hat{c}_k, i_k)}{p_{\boldsymbol{\theta}_k}(\boldsymbol{y}_k \mid c_k, i_k)}} \,\middle|\, U_k \right] \right] \\ & \quad = \left( \frac{1}{m} \sum_{i=1}^{m} E_{b,\boldsymbol{y},\boldsymbol{\theta}}\left[ \sqrt{\frac{p_{\boldsymbol{\theta}}(\boldsymbol{y} \mid \bar{b}, i)}{p_{\boldsymbol{\theta}}(\boldsymbol{y} \mid b, i)}} \right] \right)^d = B^d. \end{aligned} \tag{27}$$

In a similar way, without CSI we obtain $f(d, \mu, \mathcal{X}) \leq B^d$, with $B$ given by (20).

For 8PSK with Gray labeling, perfect CSI, and the modified metric (9), a Chernoff bound on the PEP was obtained by Zehavi [26, Appendix A]. Although the latter can be expressed in closed form, its derivation relies upon the specific features of 8PSK and Gray labeling, and cannot be applied directly to other labelings and/or other signal sets. By using (27) in (26), on the contrary, we obtain bounds on $P_b$ for any labeling, signal set, and vector channel, with and without CSI. The drawback here is that we must evaluate numerically $B$: however, numerical evaluation is much faster than simulation of the whole BICM system, and allows evaluation of $P_b$ for values usually not reachable by simulation. Moreover, $B$ is independent of $\mathcal{C}$, and hence it can be evaluated for a given $\mu$ and $\mathcal{X}$, and used to compare the performances of different codes on any given channel.

[3] See [22]. Viterbi's orthogonal convolutional codes are in fact CM codes, as explained in Section V.

*Convergence of the Union Bound:* Under the assumptions of Section II-A, with ideal interleaving the channel is memoryless and stationary. Hence, the random coding bound based on the Bhattacharyya union bound for (time-varying) convolutional codes given in [23, Ch. 5] can be applied to our case. In particular, for $R < R_0$ the bound converges. For actual codes, by replacing $f(d, \mu, \mathcal{X})$ with $B^d$ in the infinite sum of (26), we easily see that there exists a $B_0$ such that if $B < B_0$ the bound converges. Then, as long as $B$ is a decreasing function of the channel SNR, we can safely assume that the union bound converges over a range of SNR values above a certain threshold.

### B. BICM Union Bound

In this subsection we derive a union bound which does not rely on Bhattacharyya (or Chernoff) upper bounds. This turns out to be quite loose, but we shall use it only as a starting point for the tight bounds and approximations which will be developed in the next section. Since the bit metrics (7) or (9) are difficult to deal with, our technique consists of going back to the original vector channel and finding expressions in terms of PEP's of pairs of signal sequences, input to the vector channel underlying the BICM channel.

In the following we consider the modified metrics (9) and perfect CSI. Assume that $\underline{c}$ and $\underline{\hat{c}}$ have Hamming distance $d$, and let

$$\begin{aligned} \mathcal{X}_{\underline{c}}^{\underline{S}} &= \mathcal{X}_{c_1}^{i_1} \times \cdots \times \mathcal{X}_{c_k}^{i_k} \times \cdots \times \mathcal{X}_{c_d}^{i_d} \\ \mathcal{X}_{\underline{\hat{c}}}^{\underline{S}} &= \mathcal{X}_{\hat{c}_1}^{i_1} \times \cdots \times \mathcal{X}_{\hat{c}_k}^{i_k} \times \cdots \times \mathcal{X}_{\hat{c}_d}^{i_d} \end{aligned} \tag{28}$$

denote the sequence (Cartesian product) of signal subsets selected by the bits $c_k$ of $\underline{c}$ (correct subsets) and $\hat{c}_k$ of $\underline{\hat{c}}$ (complement subsets) given the label positions $S_k = i_k$ and the labeling map selected by $U_k$. For given $\underline{S}$, $\underline{U}$, and $\underline{\theta}$, the path metric difference corresponding to the sequences $\underline{c}$ and $\underline{\hat{c}}$ can be written as

$$\begin{aligned} \delta &= \sum_{k=1}^{d} \max_{z_k \in \mathcal{X}_{c_k}^{i_k}} \log p_{\theta_k}(y_k \mid z_k) - \sum_{k=1}^{d} \max_{z_k \in \mathcal{X}_{\hat{c}_k}^{i_k}} \log p_{\theta_k}(y_k \mid z_k) \\ &= \max_{\underline{z} \in \mathcal{X}_{\underline{c}}^{\underline{S}}} \log p_{\underline{\theta}}(\underline{y} \mid \underline{z}) - \max_{\underline{z} \in \mathcal{X}_{\underline{\hat{c}}}^{\underline{S}}} \log p_{\underline{\theta}}(\underline{y} \mid \underline{z}). \end{aligned} \tag{29}$$

This yields

$$\begin{aligned} &P(\underline{c} \to \underline{\hat{c}} \mid \underline{S}, \underline{U}) \\ &\quad = P(\delta \le 0 \mid \underline{S}, \underline{U}) \\ &\quad = P\left( \max_{\underline{z} \in \mathcal{X}_{\underline{c}}^{\underline{S}}} p_{\underline{\theta}}(\underline{y} \mid \underline{z}) \le \max_{\underline{z} \in \mathcal{X}_{\underline{\hat{c}}}^{\underline{S}}} p_{\underline{\theta}}(\underline{y} \mid \underline{z}) \mid \underline{S}, \underline{U} \right) \\ &\quad \overset{a)}{=} E_{\underline{x}}\left[ P\left( \max_{\underline{z} \in \mathcal{X}_{\underline{c}}^{\underline{S}}} p_{\underline{\theta}}(\underline{y} \mid \underline{z}) \le \max_{\underline{z} \in \mathcal{X}_{\underline{\hat{c}}}^{\underline{S}}} p_{\underline{\theta}}(\underline{y} \mid \underline{z}) \mid \underline{S}, \underline{U}, \underline{x} \right) \right] \\ &\quad \overset{b)}{\le} E_{\underline{x}}\left[ P\left( p_{\underline{\theta}}(\underline{y} \mid \underline{x}) \le \max_{\underline{z} \in \mathcal{X}_{\underline{\hat{c}}}^{\underline{S}}} p_{\underline{\theta}}(\underline{y} \mid \underline{z}) \mid \underline{S}, \underline{U}, \underline{x} \right) \right] \\ &\quad \overset{c)}{\le} E_{\underline{x}}\left[ \sum_{\underline{z} \in \mathcal{X}_{\underline{\hat{c}}}^{\underline{S}}} P(p_{\underline{\theta}}(\underline{y} \mid \underline{x}) \le p_{\underline{\theta}}(\underline{y} \mid \underline{z}) \mid \underline{x}) \right] \\ &\quad \overset{d)}{=} 2^{-d(m-1)} \sum_{\underline{x} \in \mathcal{X}_{\underline{c}}^{\underline{S}}} \sum_{\underline{z} \in \mathcal{X}_{\underline{\hat{c}}}^{\underline{S}}} P(\underline{x} \to \underline{z}) \end{aligned} \tag{30}$$

where a) is obtained by conditioning and averaging with respect to the transmitted signal sequence $\underline{x} \in \mathcal{X}_{\underline{c}}^{\underline{S}}$; b) is obtained from the obvious relation

$$p_{\underline{\theta}}(\underline{y} \mid \underline{x}) \le \max_{\underline{z} \in \mathcal{X}_{\underline{\hat{c}}}^{\underline{S}}} p_{\underline{\theta}}(\underline{y} \mid \underline{z});$$

c) is obtained by the union bound over all the sequences $\underline{z} \in \mathcal{X}_{\underline{\hat{c}}}^{\underline{S}}$; and d) is obtained by assuming $\underline{x}$ uniformly distributed over $\mathcal{X}_{\underline{c}}^{\underline{S}}$ and by noting that

$$P(p_{\underline{\theta}}(\underline{y} \mid \underline{x}) \le p_{\underline{\theta}}(\underline{y} \mid \underline{z}) \mid \underline{x}) = P(\underline{x} \to \underline{z}) \tag{31}$$

is the PEP of the two signal sequences $\underline{x}$ and $\underline{z}$, transmitted over the vector channel.

We are now ready to bound the PEP of BICM: it suffices to average (30) with respect to $\underline{S}$ and $\underline{U}$. We obtain

$$\begin{aligned} f(d, \mu, \mathcal{X}) &\le m^{-d} \sum_{\underline{S}} 2^{-d} \sum_{\underline{U}} 2^{-d(m-1)} \sum_{\underline{x} \in \mathcal{X}_{\underline{c}}^{\underline{S}}} \sum_{\underline{z} \in \mathcal{X}_{\underline{\hat{c}}}^{\underline{S}}} P(\underline{x} \to \underline{z}) \\ &\triangleq f_{\mathrm{ub}}(d, \mu, \mathcal{X}) \end{aligned} \tag{32}$$

where $\underline{S}$ and $\underline{U}$ run over all the possible $m^d$ and $2^d$ sequences of position and labeling maps, respectively. We will refer to (32) as to the *BICM Union Bound* (UB).

The large number of terms in the right-hand-side summation makes (32) impractical, even for moderate values of $d$ and $m$. For example, if $\mathcal{X}$ is 16QAM and $d = 10$ (as for the free distance of the best 64-state rate-$1/2$ binary convolutional code [17, pp. 466–471]) we should evaluate and sum up $4^{10} \times 2^{10} \times 2^{10(4-1)} \times 2^{10(4-1)} = 2^{90}$ terms. In order to get a computationally efficient form for the BICM UB, let

$$\Delta(\boldsymbol{x}, \boldsymbol{z}) = \log p_{\boldsymbol{\theta}}(\boldsymbol{y} \mid \boldsymbol{x}) - \log p_{\boldsymbol{\theta}}(\boldsymbol{y} \mid \boldsymbol{z}) \tag{33}$$

denote the metric difference relative to the components $\boldsymbol{x}$ and $\boldsymbol{z}$ of the sequences $\underline{x}$ and $\underline{z}$, and

$$\Delta = \sum_{k=1}^{d} \Delta(\boldsymbol{x}_k, \boldsymbol{z}_k)$$

denote the total path metric difference corresponding to $\underline{x}$ and $\underline{z}$ (note that $\boldsymbol{x}_k \neq \boldsymbol{z}_k$ for all $k = 1, \cdots, d$, since $\mathcal{X}_{c_k}^{i_k} \cap \mathcal{X}_{\hat{c}_k}^{i_k} = \emptyset$). Then, we can write [17]

$$P(\underline{x} \to \underline{z}) = P(\Delta \le 0) = \frac{1}{2\pi j} \int_{\alpha - j\infty}^{\alpha + j\infty} \Phi_{\Delta}(s) \frac{ds}{s} \tag{34}$$

where $\Phi_{\Delta}(s)$ denotes the Laplace transform of the pdf of $\Delta$, given by

$$\Phi_{\Delta}(s) = E[e^{-s\Delta}] = \prod_{k=1}^{d} E[e^{-s\Delta(\boldsymbol{x}_k, \boldsymbol{z}_k)}] = \prod_{k=1}^{d} \Phi_{\Delta(\boldsymbol{x}_k, \boldsymbol{z}_k)}(s) \tag{35}$$

and where $\alpha \in (0, \alpha_0)$, the intersection of the region of convergence of $\Phi_\Delta(s)$ with the positive real line. Finally, by using (34) into (32) and by exchanging the sums and the integral, we get the BICM UB in the form

$$f_{\mathrm{ub}}(d, \mu, \mathcal{X}) = \frac{1}{2\pi j} \int_{\alpha - j\infty}^{\alpha + j\infty} [\psi_{\mathrm{ub}}(s)]^d \, \frac{ds}{s} \tag{36}$$

where we have defined

$$\psi_{\mathrm{ub}}(s) \triangleq \frac{1}{m 2^m} \sum_{i=1}^{m} \sum_{b=0}^{1} \sum_{\boldsymbol{x} \in \mathcal{X}_b^i} \sum_{\boldsymbol{z} \in \mathcal{X}_{\bar{b}}^i} \Phi_{\Delta(\boldsymbol{x},\boldsymbol{z})}(s) \tag{37}$$

and $\alpha$ belongs to the intersection of the regions of convergence of the $\Phi_{\Delta(\boldsymbol{x},\boldsymbol{z})}(s)$'s (for all pairs $\boldsymbol{x}, \boldsymbol{z}$) with the real positive line. Use of this technique is made easy by the fact that $\Phi_{\Delta(\boldsymbol{x},\boldsymbol{z})}(s)$ (and its convergence region) can be expressed in closed form in many cases of interest [3]. The integral over the vertical path $(\alpha - j\infty, \alpha + j\infty)$ can be efficiently computed by the method developed by the authors in [4] and based on a change of variable and Gauss–Chebyshev quadratures. Alternatively, one can use the looser bound described below (see [3] and references therein), which we state here without proof.

*1) Chernoff Bound:* Let $\Delta$ be a continuous real random variable which takes on values over the whole real line and let $\Phi_\Delta(s) = E[e^{-s\Delta}]$ denote the Laplace transform of the pdf of $\Delta$. Then: i) The convergence region of $\Phi_\Delta(s)$ is the vertical strip $\alpha_1 < \mathrm{Re}\,(s) < \alpha_2$ where $\alpha_1 < 0 < \alpha_2$; ii) The following inequality holds:

$$P(\Delta \le 0) \le \min_{0 < \alpha < \alpha_2} \Phi_\Delta(\alpha). \tag{38}$$

Then, by noting that $[\psi_{\mathrm{ub}}(s)]^d$ is a weighted sum of Laplace transforms of the form (35), we obtain the Chernoff-type bound

$$f_{\mathrm{ub}}(d, \mu, \mathcal{X}) \le \min_{0 < \alpha < \alpha_0} [\psi_{\mathrm{ub}}(\alpha)]^d. \tag{39}$$

We hasten to say that, unless the minimization with respect to $\alpha$ can be obtained in closed form, the computation of the right-hand side of (39) and (36) have roughly the same (low) complexity. Thus there is no real benefit in using the looser bound (39) instead of (36).

The BICM UB is very general, and can be easily extended to the case of no CSI by defining the appropriate path metric difference between signal sequences $\underline{\boldsymbol{x}}$ and $\underline{\boldsymbol{z}}$ (we omit here this trivial derivation). Yet, it is rather loose, due to steps b) and c) of (30). Step b) is a strict inequality if

$$\underline{\boldsymbol{x}} \neq \arg \max_{\underline{\boldsymbol{z}} \in \mathcal{X}_{\underline{\boldsymbol{c}}}^{\underline{\boldsymbol{S}}}} \log p_{\underline{\boldsymbol{\theta}}}(\underline{\boldsymbol{y}} \mid \underline{\boldsymbol{z}})$$

while step c) is a union bound which generates a large number of terms corresponding to almost overlapping pairwise-error events.

*Example 3. Uncoded 4PSK in AWGN:* Consider uncoded transmission of 4PSK with Gray labeling in AWGN with ideal coherent detection. Since 4PSK with Gray labeling provides two binary antipodal parallel independent channels, the presence of a bit interleaver is irrelevant, so that the exact BER is given by [17] $P_{\mathrm{b}} = f(1, \mu, \mathcal{X}) = Q(1/\sigma)$. By applying the BICM UB we get

$$P_{\mathrm{b}} \le f_{\mathrm{ub}}(1, \mu, \mathcal{X}) = Q(1/\sigma) + Q(\sqrt{2}/\sigma). \tag{40}$$

The two terms on the right-hand side correspond to pairwise errors at distance $\sqrt{2}$ and 2, respectively. Hence, the BICM UB takes into account the contribution of events at distance 2, even if they are irrelevant to the BER. This simple observation calls for some "expurgation" technique in order to exclude the irrelevant terms from the BICM UB. This approach will be formally stated in the following section for AWGN and fading channels with perfect CSI.

### C. Bounds for Fading Channels with Perfect CSI

In this subsection we derive a tight upper bound to the PEP of BICM for Rician fading channels with perfect CSI. Here $p_{\boldsymbol{\theta}}(\underline{\boldsymbol{y}} \mid \boldsymbol{x})$ takes the form (23), where, without loss of generality, we assume zero carrier phase. The bound is first derived for AWGN ($K \to \infty$), and then extended to the fading case ($K < \infty$).

With AWGN, maximizing $p_{\underline{\boldsymbol{\theta}}}(\underline{\boldsymbol{y}} \mid \underline{\boldsymbol{x}})$ over the possible signal sequences is tantamount to minimizing the squared Euclidean distance $|\underline{\boldsymbol{y}} - \underline{\boldsymbol{x}}|^2$. We exploit the signal space geometry to expurgate the BICM union bound. Let

$$\Gamma_{\underline{\boldsymbol{x}},\underline{\boldsymbol{z}}} = \{\underline{\boldsymbol{y}} : |\underline{\boldsymbol{y}} - \underline{\boldsymbol{x}}|^2 \ge |\underline{\boldsymbol{y}} - \underline{\boldsymbol{z}}|^2\} \tag{41}$$

denote the pairwise error region of the two signal sequences $\underline{\boldsymbol{x}}$ and $\underline{\boldsymbol{z}}$, so that $P(\underline{\boldsymbol{x}} \to \underline{\boldsymbol{z}}) = P(\Gamma_{\underline{\boldsymbol{x}},\underline{\boldsymbol{z}}})$. Given $\underline{\boldsymbol{S}}$ and $\underline{\boldsymbol{U}}$, let $\hat{\underline{\boldsymbol{x}}} \in \mathcal{X}_{\underline{\hat{\boldsymbol{c}}}}^{\underline{\boldsymbol{S}}}$ and $\hat{\underline{\boldsymbol{z}}} \in \mathcal{X}_{\underline{\hat{\bar{\boldsymbol{c}}}}}^{\underline{\boldsymbol{S}}}$ be two particular signal sequences corresponding to $\underline{\hat{\boldsymbol{c}}}$ and $\underline{\hat{\bar{\boldsymbol{c}}}}$, respectively. Then, we have

*Theorem 1. Irrelevant Error Events (AWGN):* The term $P(\hat{\underline{\boldsymbol{x}}} \to \hat{\underline{\boldsymbol{z}}})$ can be neglected in (30) without altering the inequality if there exist $\ell$ sequences $\underline{\boldsymbol{z}}^{(i)}$ different from $\hat{\underline{\boldsymbol{x}}}$ and $\hat{\underline{\boldsymbol{z}}}$, for all $i = 1, \cdots, \ell$, either in the complement subsets sequence $\mathcal{X}_{\underline{\hat{\boldsymbol{c}}}}^{\underline{\boldsymbol{S}}}$ or in the correct subsets sequence $\mathcal{X}_{\underline{\hat{\bar{\boldsymbol{c}}}}}^{\underline{\boldsymbol{S}}}$, such that

$$\Gamma_{\hat{\underline{\boldsymbol{x}}},\hat{\underline{\boldsymbol{z}}}} \subseteq \bigcup_{i=1}^{\ell} \Gamma_{\hat{\underline{\boldsymbol{x}}},\underline{\boldsymbol{z}}^{(i)}} \tag{42}$$

*Proof:* From (42) we write

$$\Gamma_{\hat{\underline{\boldsymbol{x}}},\hat{\underline{\boldsymbol{z}}}} = \bigcup_{i=1}^{\ell} \{\Gamma_{\hat{\underline{\boldsymbol{x}}},\hat{\underline{\boldsymbol{z}}}} \cap \Gamma_{\hat{\underline{\boldsymbol{x}}},\underline{\boldsymbol{z}}^{(i)}}\}.$$

Consider separately each subregion $\{\Gamma_{\hat{\underline{\boldsymbol{x}}},\hat{\underline{\boldsymbol{z}}}} \cap \Gamma_{\hat{\underline{\boldsymbol{x}}},\underline{\boldsymbol{z}}^{(i)}}\}$. Trivially, if $\underline{\boldsymbol{z}}^{(i)} \in \mathcal{X}_{\underline{\hat{\boldsymbol{c}}}}^{\underline{\boldsymbol{S}}}$, $\{\Gamma_{\hat{\underline{\boldsymbol{x}}},\hat{\underline{\boldsymbol{z}}}} \cap \Gamma_{\hat{\underline{\boldsymbol{x}}},\underline{\boldsymbol{z}}^{(i)}}\}$ is already counted in the pairwise error event $\{\hat{\underline{\boldsymbol{x}}} \to \underline{\boldsymbol{z}}^{(i)}\}$. The case $\underline{\boldsymbol{z}}^{(i)} \in \mathcal{X}_{\underline{\hat{\bar{\boldsymbol{c}}}}}^{\underline{\boldsymbol{S}}}$ must be considered more carefully. In this case, for $\underline{\boldsymbol{y}} \in \{\Gamma_{\hat{\underline{\boldsymbol{x}}},\hat{\underline{\boldsymbol{z}}}} \cap \Gamma_{\hat{\underline{\boldsymbol{x}}},\underline{\boldsymbol{z}}^{(i)}}\}$

$$\hat{\underline{\boldsymbol{x}}} \neq \arg \max_{\underline{\boldsymbol{x}} \in \mathcal{X}_{\underline{\hat{\bar{\boldsymbol{c}}}}}^{\underline{\boldsymbol{S}}}} \log p_{\underline{\boldsymbol{\theta}}}(\underline{\boldsymbol{y}} \mid \underline{\boldsymbol{x}}).$$

Hence, the BICM decoder path metric for $\underline{\hat{\bar{\boldsymbol{c}}}}$ is not the metric corresponding to $\hat{\underline{\boldsymbol{x}}}$. Therefore, the subregion $\{\Gamma_{\hat{\underline{\boldsymbol{x}}},\hat{\underline{\boldsymbol{z}}}} \cap \Gamma_{\hat{\underline{\boldsymbol{x}}},\underline{\boldsymbol{z}}^{(i)}}\}$

does not belong to the pairwise error region of the BICM decoder, and can be neglected. We conclude that $\Gamma_{\underline{\hat{x}},\underline{\hat{z}}}$ is irrelevant for the bound (30). □

Condition (42) is generally difficult to check. A much simpler condition is expressed by the following Theorem.

*Theorem 2. Decomposability (AWGN):* The term $P(\underline{\hat{x}} \to \underline{\hat{z}})$ can be neglected in (30) without altering the inequality if there exist two sequences $\underline{z}'$ and $\underline{z}''$ different from $\underline{\hat{x}}$ and $\underline{\hat{z}}$, either in $\mathcal{X}_{\underline{\hat{c}}}^{\underline{S}}$ or in $\mathcal{X}_{\underline{\hat{c}}}^{\underline{S}}$, such that

$$\begin{aligned} &1) \quad \mathrm{Re}\,[(\underline{z}' - \underline{\hat{x}})^{\dagger}(\underline{z}'' - \underline{\hat{x}})] \geq 0 \\ &2) \quad \underline{\hat{z}} = \underline{z}' + \underline{z}'' - \underline{\hat{x}}. \end{aligned} \tag{43}$$

*Proof:* Since (43) is translation-invariant, without loss of generality we can choose $\underline{\hat{x}} = \underline{0}$. Setting $\underline{y} \in \Gamma_{\underline{0},\underline{\hat{z}}}$, $\underline{y} \notin \Gamma_{\underline{0},\underline{z}'}$, and $\underline{y} \notin \Gamma_{\underline{0},\underline{z}''}$, we obtain

$$|\underline{y}|^2 \leq |\underline{y} - \underline{z}'|^2 \quad \text{and} \quad |\underline{y}|^2 \leq |\underline{y} - \underline{z}''|^2$$

or, equivalently,

$$2\,\mathrm{Re}\,(\underline{y}^{\dagger}\underline{z}') \leq |\underline{z}'|^2 \quad \text{and} \quad 2\,\mathrm{Re}\,(\underline{y}^{\dagger}\underline{z}'') \leq |\underline{z}''|^2.$$

Since $\underline{\hat{z}} = \underline{z}' + \underline{z}''$ and $\mathrm{Re}\,((\underline{z}')^{\dagger}\underline{z}'') \geq 0$, summing up the above inequalities we obtain

$$2\,\mathrm{Re}\,(\underline{y}^{\dagger}\underline{\hat{z}}) \leq |\underline{z}'|^2 + |\underline{z}''|^2 \leq |\underline{\hat{z}}|^2$$

contrary to the assumption that $\underline{y} \in \Gamma_{\underline{0},\underline{\hat{z}}}$. We conclude that $\underline{y}$ must be either in $\Gamma_{\underline{0},\underline{z}'}$ or $\Gamma_{\underline{0},\underline{z}''}$, and hence

$$\Gamma_{\underline{0},\underline{\hat{z}}} \subseteq \{\Gamma_{\underline{0},\underline{z}'} \cup \Gamma_{\underline{0},\underline{z}''}\}$$

The proof follows from Theorem 1. □

*Remark:* Theorem 2 is a generalization of the expurgation technique based on "decomposability" of error sequences, used by Verdú to bound the minimum BER of an uncoded multiuser system with antipodal binary modulation in AWGN [21].

*Corollary 1. Extension to Fading Channels ($K < \infty$):* The term $P(\underline{\hat{x}} \to \underline{\hat{z}})$ can be neglected in (30) without changing the inequality if there exist two sequences $\underline{z}'$ and $\underline{z}''$, different from $\underline{\hat{x}}$ and $\underline{\hat{z}}$, either in $\mathcal{X}_{\underline{\hat{c}}}^{\underline{S}}$ or in $\mathcal{X}_{\underline{\hat{c}}}^{\underline{S}}$, such that, for all components $\hat{x}_k$, $\hat{z}_k$, $z'_k$, and $z''_k$

$$\begin{aligned} &1) \quad \mathrm{Re}\,[(z'_k - \hat{x}_k)^{\dagger}(z''_k - \hat{x}_k)] \geq 0 \\ &2) \quad \hat{z}_k = z'_k + z''_k - \hat{x}_k. \end{aligned} \tag{44}$$

*Proof:* Condition (44) is the componentwise version of (43). Let $g_k$ denote the (scalar) fading gain affecting the $k$th component of the received signal. If (44) holds, we get trivially that

$$\begin{aligned} &1) \quad \mathrm{Re}\,[(g_k z'_k - g_k \hat{x}_k)^{\dagger}(g_k z''_k - g_k \hat{x}_k)] \geq 0 \\ &2) \quad g_k \hat{z}_k = g_k z'_k + g_k z''_k - g_k \hat{x}_k. \end{aligned} \tag{45}$$

Then, let $\boldsymbol{G} = \mathrm{diag}\,(g_k \boldsymbol{I}_N)$, and consider the new signal sequences $\boldsymbol{G}\underline{\hat{x}}$, $\boldsymbol{G}\underline{\hat{z}}$, $\boldsymbol{G}\underline{z}'$, and $\boldsymbol{G}\underline{z}''$. We can extend (45) to sequences, which yields

$$\begin{aligned} &1) \quad \mathrm{Re}\,[(\boldsymbol{G}\underline{z}' - \boldsymbol{G}\underline{\hat{x}})^{\dagger}(\boldsymbol{G}\underline{z}'' - \boldsymbol{G}\underline{\hat{x}})] \geq 0 \\ &2) \quad \boldsymbol{G}\underline{\hat{z}} = \boldsymbol{G}\underline{z}' + \boldsymbol{G}\underline{z}'' - \boldsymbol{G}\underline{\hat{x}}. \end{aligned} \tag{46}$$

Hence, Theorem 2 can be applied for any fixed $\boldsymbol{G}$. Finally, we can eliminate the event $\{\underline{\hat{x}} \to \underline{\hat{z}}\}$ from the bound by: i) conditioning with respect to $\boldsymbol{G}$; ii) applying Theorem 2 and eliminating the region $\Gamma_{\boldsymbol{G}\underline{\hat{x}},\boldsymbol{G}\underline{\hat{z}}}$ for $\boldsymbol{G}$ given; iii) averaging the expurgated bound with respect to $\boldsymbol{G}$. □

*1) Application to QAM Signal Sets with Gray Labeling:* Application of Corollary 1 (and of Theorem 2) is particularly simple when $\mathcal{X}$ is a square $M$-QAM signal set with Gray labeling (see, for example, the 16QAM signal set represented in Fig. 2(b)). It is easy to verify that, with Gray labeling, each $\boldsymbol{x} \in \mathcal{X}_b^i$ has exactly a single nearest neighbor $\hat{\boldsymbol{z}} \in \mathcal{X}_b^i$ (not necessarily at distance $d_{\min}$), for all $i = 1, \cdots, m$ and $b = 0, 1$. Moreover, for each $\boldsymbol{x} \in \mathcal{X}_b^i$, all the pairs $(\boldsymbol{x}, \boldsymbol{z})$ with $\boldsymbol{z} \in \mathcal{X}_b^i$ and $\boldsymbol{z} \neq \hat{\boldsymbol{z}}$, satisfy condition (44) of Corollary 1. For example, if $\boldsymbol{x} \in \mathcal{X}_1^1$ corresponds to the point labeled 1011 in Fig. 2(b), $\hat{\boldsymbol{z}} \in \mathcal{X}_0^1$ to the point labeled 0011, and $\boldsymbol{z} \in \mathcal{X}_0^1$ to the point labeled 0010, then $\boldsymbol{z}'$ and $\boldsymbol{z}''$ correspond to the points labeled 0011 and 1010, respectively. For a given integer $d$, $\mathcal{X}^d$ is a hypercube carved from a cubic lattice of dimension $d$. Hence, condition (44) can be immediately extended to sequences in $\mathcal{X}^d$. We conclude that for all $\underline{\boldsymbol{x}} \in \mathcal{X}_{\underline{\hat{c}}}^{\underline{S}}$ there exists only one relevant error event $\{\underline{\boldsymbol{x}} \to \underline{\hat{\boldsymbol{z}}}\}$, where $\underline{\hat{\boldsymbol{z}}}$ is the unique nearest neighbor of $\underline{\boldsymbol{x}}$ in $\mathcal{X}_{\underline{\hat{c}}}^{\underline{S}}$. By expurgating the BICM UB (32) we get the new bound

$$\begin{aligned} f(d, \mu, \mathcal{X}) &\leq m^{-d} \sum_{\underline{S}} 2^{-d} \sum_{\underline{U}} 2^{-d(m-1)} \sum_{\underline{\boldsymbol{x}} \in \mathcal{X}_{\underline{\hat{c}}}^{\underline{S}}} P(\underline{\boldsymbol{x}} \to \underline{\hat{\boldsymbol{z}}}) \\ &\triangleq f_{\mathrm{ex}}(d, \mu, \mathcal{X}). \end{aligned} \tag{47}$$

This will be referred to as *BICM Expurgated Bound* (EX). In order to obtain a computationally efficient form of it, we use once again a Laplace-transform approach. Let $\Delta(\boldsymbol{x}, \hat{\boldsymbol{z}})$ denote the metric difference

$$\Delta(\boldsymbol{x}, \hat{\boldsymbol{z}}) = \log p_{\boldsymbol{\theta}}(\boldsymbol{y} \mid \boldsymbol{x}) - \log p_{\boldsymbol{\theta}}(\boldsymbol{y} \mid \hat{\boldsymbol{z}})$$

between the branch metric of $\boldsymbol{x}$ and the branch metric of its (unique) nearest neighbor $\hat{\boldsymbol{z}} \in \mathcal{X}_b^i$. By using (34) and exchanging the sums and the integral, we get

$$f_{\mathrm{ex}}(d, \mu, \mathcal{X}) = \frac{1}{2\pi j} \int_{\alpha - j\infty}^{\alpha + j\infty} \psi_{\mathrm{ex}}(s)^d \, \frac{ds}{s} \tag{48}$$

where we have defined the function $\psi_{\mathrm{ex}}(s)$ as

$$\psi_{\mathrm{ex}}(s) \triangleq \frac{1}{m 2^m} \sum_{i=1}^{m} \sum_{b=0}^{1} \sum_{\boldsymbol{x} \in \mathcal{X}_b^i} \Phi_{\Delta(\boldsymbol{x}, \hat{\boldsymbol{z}})}(s). \tag{49}$$

The BICM EX is formally similar to the BICM UB, but it includes only one error event $\{\underline{\boldsymbol{x}} \to \underline{\hat{\boldsymbol{z}}}\}$ for each possible transmitted sequence $\underline{\boldsymbol{x}}$, rather than all the possible $2^{d(m-1)}$ error events $\{\underline{\boldsymbol{x}} \to \underline{\boldsymbol{z}}\}$, for all $\underline{\boldsymbol{z}} \in \mathcal{X}_{\underline{\hat{c}}}^{\underline{S}}$. Computational experience has shown that (48), when used with other signal sets, often provides a very tight approximation to the PEP. Actually, as we prove in the following subsection, with Gray labeling and AWGN the approximation obtained from BICM is asymptotically tight. Numerical evaluation of (48) can be efficiently performed by using the Gauss–Chebyshev method of [4].

*Example 3 (Continued):* If we apply the BICM EX to the case of uncoded 4PSK with Gray labeling over AWGN we obtain

$$P_{\text{bin}} \le f_{\text{ex}}(1,\mu,\mathcal{X}) = Q(1/\sigma)$$

which is tight (see Example 3 of Section IV-B).

*2) Asymptotic Behavior of the PEP:* Here we establish the behavior of $f(d,\mu,\mathcal{X})$ as $\sigma \to 0$ with fading channels with perfect CSI. This provides some insight into the asymptotic performance of BICM for high SNR. For Rician fading channels with perfect CSI we get [20]

$$\Phi_{\Delta(\boldsymbol{x},\boldsymbol{z})}(s) = \frac{\exp\left(-\frac{s(1-s\sigma^2)K|\boldsymbol{x}-\boldsymbol{z}|^2/(K+1)}{1+s(1-s\sigma^2)|\boldsymbol{x}-\boldsymbol{z}|^2/(K+1)}\right)}{1+s(1-s\sigma^2)|\boldsymbol{x}-\boldsymbol{z}|^2/(K+1)}. \tag{50}$$

For AWGN $f_{\text{ub}}(d,\mu,\mathcal{X})$ and $f_{\text{ex}}(d,\mu,\mathcal{X})$ can be written as

$$f_{\text{ub}}(d,\mu,\mathcal{X}) = m^{-d}\sum_{\underline{S}} 2^{-d}\sum_{\underline{U}} 2^{-d(m-1)} \times \sum_{\underline{\boldsymbol{x}}\in\mathcal{X}_{\underline{\hat{c}}}^{\underline{S}}}\sum_{\underline{\boldsymbol{z}}\in\mathcal{X}_{\underline{\hat{c}}}^{\underline{S}}} Q\left(\frac{|\underline{\boldsymbol{x}}-\underline{\boldsymbol{z}}|}{\sqrt{2}\sigma}\right) \tag{51}$$

$$f_{\text{ex}}(d,\mu,\mathcal{X}) = m^{-d}\sum_{\underline{S}} 2^{-d}\sum_{\underline{U}} 2^{-d(m-1)} \times \sum_{\underline{\boldsymbol{x}}\in\mathcal{X}_{\underline{\hat{c}}}^{\underline{S}}} Q\left(\frac{|\underline{\boldsymbol{x}}-\underline{\hat{\boldsymbol{z}}}|}{\sqrt{2}\sigma}\right). \tag{52}$$

The minimum possible Euclidean distance between any $\underline{\boldsymbol{x}} \in \mathcal{X}_{\underline{\hat{c}}}^{\underline{S}}$ and $\underline{\boldsymbol{z}} \in \mathcal{X}_{\underline{\hat{c}}}^{\underline{S}}$, for all possible sequences of positions $\underline{S}$, is $\sqrt{d}d_{\min}$, where $d_{\min}$ denotes the minimum Euclidean distance of $\mathcal{X}$. From (51) we get

$$\begin{aligned}\lim_{\sigma\to 0}\frac{f_{\text{ub}}(d,\mu,\mathcal{X})}{Q(\sqrt{d}d_{\min}/(\sqrt{2}\sigma))} &= m^{-d}\sum_{\underline{S}} 2^{-d}\sum_{\underline{U}} 2^{-d(m-1)}\sum_{\underline{\boldsymbol{x}}\in\mathcal{X}_{\underline{\hat{c}}}^{\underline{S}}} N_{\min}(d,\underline{\boldsymbol{x}}) \\ &= N_{\min}(d)\end{aligned} \tag{53}$$

where $N_{\min}(d,\underline{\boldsymbol{x}})$ is the number of neighbors $\underline{\boldsymbol{z}} \in \mathcal{X}_{\underline{\hat{c}}}^{\underline{S}}$ at distance $\sqrt{d}d_{\min}$ from $\underline{\boldsymbol{x}}$ and $N_{\min}(d)$ is the expected number of nearest neighbors of an error event of Hamming distance $d$, averaged over $\underline{\boldsymbol{x}} \in \mathcal{X}_{\underline{\hat{c}}}^{\underline{S}}$, $\underline{S}$, and $\underline{U}$. In order to compute $N_{\min}(d)$, we write

$$N_{\min}(d,\underline{\boldsymbol{x}}) = \prod_{k=1}^{d} N_{\min}(1,\boldsymbol{x}_k).$$

Hence, from the right-hand side of (51) we get

$$N_{\min}(d) = [N_{\min}(1)]^d \tag{54}$$

where

$$N_{\min}(1) \triangleq \frac{1}{m2^m}\sum_{i=1}^{m}\sum_{b=0}^{1}\sum_{\boldsymbol{x}\in\mathcal{X}_b^i} N_{\min}(1,\boldsymbol{x}) \tag{55}$$

is the average number of signals at distance $d_{\min}$ in the complement subset of a signal $\boldsymbol{x}$. Note that $N_{\min}(1)$ depends only on $\mathcal{X}$ and on the labeling $\mu$. From (54) we may observe the impact of the value of $N_{\min}(1)$ on the performance of BICM: in fact, if this is greater than 1 then $N_{\min}(d)$ increases exponentially with the Hamming distance $d$, thus impairing the performance of the coding scheme at moderate SNR. If, on the contrary, $N_{\min}(1) < 1$, increasing $d$ does decrease the value of $N_{\min}(d)$, and hence improves the performance. This fact is further discussed in the following, where we show how the choice of the labeling map $\mu$ affects the value of $N_{\min}(1)$.

Let now $\mathcal{D}$ denote the event that a signal sequence $\underline{\boldsymbol{x}} \in \mathcal{X}_{\underline{\hat{c}}}^{\underline{S}}$ has a nearest neighbor at distance $\sqrt{d}d_{\min}$ in $\mathcal{X}_{\underline{\hat{c}}}^{\underline{S}}$, and let $\mathcal{I}_{\mathcal{D}}(\underline{\boldsymbol{x}})$ denote the indicator function of $\mathcal{D}$. Then, from (51) we get

$$\begin{aligned}\lim_{\sigma\to 0}\frac{f_{\text{ex}}(d,\mu,\mathcal{X})}{Q(\sqrt{d}d_{\min}/(\sqrt{2}\sigma))} &= m^{-d}\sum_{\underline{S}} 2^{-d}\sum_{\underline{U}} 2^{-d(m-1)}\sum_{\underline{\boldsymbol{x}}\in\mathcal{X}_{\underline{\hat{c}}}^{\underline{S}}} \mathcal{I}_{\mathcal{D}}(\underline{\boldsymbol{x}}) \\ &= P(\mathcal{D}) \le 1.\end{aligned} \tag{56}$$

In the case of Gray labeling we have

*Theorem 3. Asymptotically Tight Bound and Approximation (AWGN):* If $\mu$ is a Gray labeling for the signal set $\mathcal{X}$, then $N_{\min}(d) \le 1$, and

$$\lim_{\sigma\to 0}\frac{f_{\text{ub}}(d,\mu,\mathcal{X})}{f(d,\mu,\mathcal{X})} = \lim_{\sigma\to 0}\frac{f_{\text{ex}}(d,\mu,\mathcal{X})}{f(d,\mu,\mathcal{X})} = 1. \tag{57}$$

*Proof:* If $\mu$ is a Gray labeling for $\mathcal{X}$, by definition $N_{\min}(d,\underline{\boldsymbol{x}})$ is either 1 or 0, hence $N_{\min}(d,\underline{\boldsymbol{x}}) = \mathcal{I}_{\mathcal{D}}(\underline{\boldsymbol{x}})$ and

$$N_{\min}(d) = P(\mathcal{D}) \le 1. \tag{58}$$

In order to prove (57), we need a lower bound on $f(d,\mu,\mathcal{X})$. We follow the technique of [21] and consider a *genie* decoder with additional side information. In particular, suppose that, for each transmitted signal sequence $\underline{\boldsymbol{x}} \in \mathcal{X}_{\underline{\hat{c}}}^{\underline{S}}$ the genie decoder knows that the only possible concurrent signal sequence is $\underline{\hat{\boldsymbol{z}}}$, i.e., the unique sequence at minimum distance from $\underline{\boldsymbol{x}}$ in the complement subset $\mathcal{X}_{\underline{\hat{c}}}^{\underline{S}}$. If $|\underline{\hat{\boldsymbol{z}}}-\underline{\boldsymbol{x}}| > \sqrt{d}d_{\min}$, the PEP of the genie decoder is trivially bounded from below by zero. Hence, we can write

$$f(d,\mu,\mathcal{X}) \ge P(\mathcal{D})Q\left(\frac{\sqrt{d}d_{\min}}{\sqrt{2}\sigma}\right). \tag{59}$$

If $\mathcal{X}$ is a squared QAM signal set, we can write

$$\begin{aligned}P(\mathcal{D})Q\left(\frac{\sqrt{d}d_{\min}}{\sqrt{2}\sigma}\right) &\le f(d,\mu,\mathcal{X}) \le f_{\text{ex}}(d,\mu,\mathcal{X}) \\ &\le f_{\text{ub}}(d,\mu,\mathcal{X}).\end{aligned} \tag{60}$$

By taking the limit for $\sigma \to 0$ and by using (53), (56), and (58) we obtain (57). If $\mathcal{X}$ is not a squared QAM signal set, the order of the middle inequality in (60) may be reversed (i.e., it may be $f_{\text{ex}}(d,\mu,\mathcal{X}) \le f(d,\mu,\mathcal{X})$). However, it is immediate to check that the left and the right inequalities still hold, thus giving (57) as $\sigma \to 0$. □

For any other labeling, the asymptotic ratio between the PEP upper and lower bounds is $N_{\min}(d)/P(\mathcal{D}) > 1$. Hence, Theorem 3 shows a sort of asymptotic optimality of Gray labeling for BICM.

For $K < \infty$, the asymptotic analysis is more complicated, because a closed-form expression for the PEP $P(\underline{\boldsymbol{x}} \to \underline{\boldsymbol{z}})$ for arbitrary $\underline{\boldsymbol{x}}$, $\underline{\boldsymbol{z}}$, and $K$ is missing, and because, if $K < \infty$, $P(\underline{\boldsymbol{x}} \to \underline{\boldsymbol{z}})$ does not decrease exponentially with SNR. For these reasons, we restrict our analysis to the BICM EX and we use the Chernoff bound (see (38)). From (38) and (50) we obtain

$$\min_{0<\alpha<\alpha_2} \Phi_{\Delta(\underline{\boldsymbol{x}},\underline{\hat{\boldsymbol{z}}})}(\alpha) = \prod_{k=1}^{d} \frac{\exp\left(-\frac{K|\boldsymbol{x}_k-\hat{\boldsymbol{z}}_k|^2/(4\sigma^2(K+1))}{1+|\boldsymbol{x}_k-\hat{\boldsymbol{z}}_k|^2/(4\sigma^2(K+1))}\right)}{1+|\boldsymbol{x}_k-\hat{\boldsymbol{z}}_k|^2/(4\sigma^2(K+1))}. \tag{61}$$

By assuming that $\mathcal{X}$ is a square QAM constellation with Gray labeling, by substituting (61) into the BICM EX (47), and by choosing $\sigma$ so small that $1 \ll |\boldsymbol{x}_k - \hat{\boldsymbol{z}}_k|^2/(4\sigma^2(K+1))$ for all $k = 1, \cdots, d$, we get

$$f_{\text{ex}}(d, \mu, \mathcal{X}) \sim \left(\frac{4(K+1)e^{-K}}{d_{\text{h}}^2/\sigma^2}\right)^d \tag{62}$$

where

$$d_{\text{h}}^2 = \left(\frac{1}{m2^m}\sum_{i=1}^{m}\sum_{b=0}^{1}\sum_{\boldsymbol{x}\in\mathcal{X}_b^i}\frac{1}{|\boldsymbol{x}-\hat{\boldsymbol{z}}|^2}\right)^{-1}. \tag{63}$$

Hence, for any given $K < \infty$, as $\sigma \to 0$ the function $f_{\text{ex}}(d, \mu, \mathcal{X})$ depends on the inverse $d$th power of SNR. In the fading case there is no single dominant term, and therefore the union bound may not be tight as $\sigma \to 0$. Results of Section V show that the BICM EX is very close to the real PEP (obtained by simulation) also in the fading case, while the BICM UB is extremely loose at all SNR's.

The coefficient $d_{\text{h}}^2$ can be interpreted as the *harmonic mean* of the minimum squared Euclidean distance between complementary subsets of $\mathcal{X}$. By comparing the left-hand of (62) with the Chernoff bound on the PEP of a CM scheme in Rician fading (see, for example, [14, p. 162]), we observe that $(d_{\text{h}}^2)^d$ plays the role of the so-called *squared product distance* of the error event.

The role of the two basic parameters $N_{\min}(1)$ and $d_{\text{h}}^2$ in the design of BICM will be emphasized in next section.

## V. Design Guidelines

The analysis of previous sections shows that the performance of BICM with ideal interleaving depends mainly on the following parameters. i) The minimum binary Hamming distance of $\mathcal{C}$, denoted in the following by $d_2(\mathcal{C})$. ii) The minimum Euclidean distance $d_{\min}$ of $\mathcal{X}$. iii) The average number of nearest neighbors $N_{\min}(1)$, defined in (55). iv) The harmonic mean of the minimum squared distance $d_{\text{h}}^2$, defined in (63).

On the AWGN channel with coherent detection (perfect CSI), the performance of BICM is asymptotically dominated by the error event at minimum Euclidean distance

$$d_{\text{E}} = \sqrt{d_2(\mathcal{C})}\, d_{\min}. \tag{64}$$

From the union bound (26), the error coefficient corresponding to the distance $d_{\text{E}}$ is $W_I(d_2(\mathcal{C}))[N_{\min}(1)]^{d_2(\mathcal{C})}$ divided by $k_c$ or by $K_c$, in the case of convolutional or block codes, respectively.

In the case of independent (or ideally interleaved) Rayleigh fading channel with perfect CSI, the asymptotic error performance of any code is proportional to the inverse of SNR raised to a power equal to the *code diversity*, i.e., the minimum Hamming distance between signal sequences produced by the encoder/modulator [14]. We showed that in BICM this code diversity equals $d_2(\mathcal{C})$. Recalling the definition of the spectral efficiency $R$, with PSK/QAM and Gray labeling, and letting $\sigma \to 0$ and $\text{SNR} = 1/\sigma^2 = RE_b/N_0$, from (62) we can write

$$\log_{10} P_{\text{b}} \simeq -d_2(\mathcal{C})\left[\left(Rd_{\text{h}}^2\right)_{\text{dB}} + \left(\frac{E_b}{N_0}\right)_{\text{dB}}\right] \Big/ 10 + \text{const.} \tag{65}$$

Thus for high values of $E_b/N_0$, the BER curve on a logarithmic scale approaches a straight line with slope proportional to $-d_2(\mathcal{C})$, translated horizontally by the offset $(Rd_{\text{h}}^2)_{\text{dB}}$. As we shall see in an example, this rule of thumb turns out to be useful for a quick comparison between different BICM schemes over the Rayleigh fading channel.

Consider next the case where $\mathcal{C}$ is a convolutional code. The decoder complexity can be measured by the *trellis complexity* $\tau = N_S N_B$, where $N_S = 2^v$ denotes the number of states in the minimal encoder for $\mathcal{C}$ ($v$ is the encoder memory), and $N_B = 2^{k_c}$ is the number of branches merging into each state at each trellis step ($k_c$ is the number of encoder inputs). The complexity of punctured convolutional codes is the same $\tau$ of the low-rate code; they are derived from [7].

The design of BICM can be summarized as follows:

1) Given a signal set $\mathcal{X}$ of cardinality $M = 2^m$ and dimension $N$, a desired decoder complexity $\tau$, and a required spectral efficiency $R$, choose the best binary convolutional code of rate $r_c = k_c/n_c$ such that $R = r_c m/N$, with complexity $\leq \tau$.
2) Depending on the application (AWGN, Rayleigh fading, etc.) and on the detection technique (perfect CSI, no CSI, etc.), choose the labeling map $\mu$ according to some optimality criterion. We showed that, for perfect CSI, practical and sensible criteria are the minimization of $N_{\min}(1)$ and the maximization of $d_{\text{h}}^2$. Table I lists $N_{\min}(1)$ and $d_{\text{h}}^2$ for some PSK and QAM signal sets and some labelings (observe that Gray and quasi-Gray labelings provide values of $N_{\min}(1)$ never exceeding 1, unlike other labelings).

In this way, the binary code is picked independently of $\mu$ and $\mathcal{X}$. Since the best convolutional codes are known for many rates and complexities (see [17], [11], [7], and references therein) normally no search for new binary codes is needed. We mention in passing that BICM easily allows the concatenation of any given signal set with the powerful

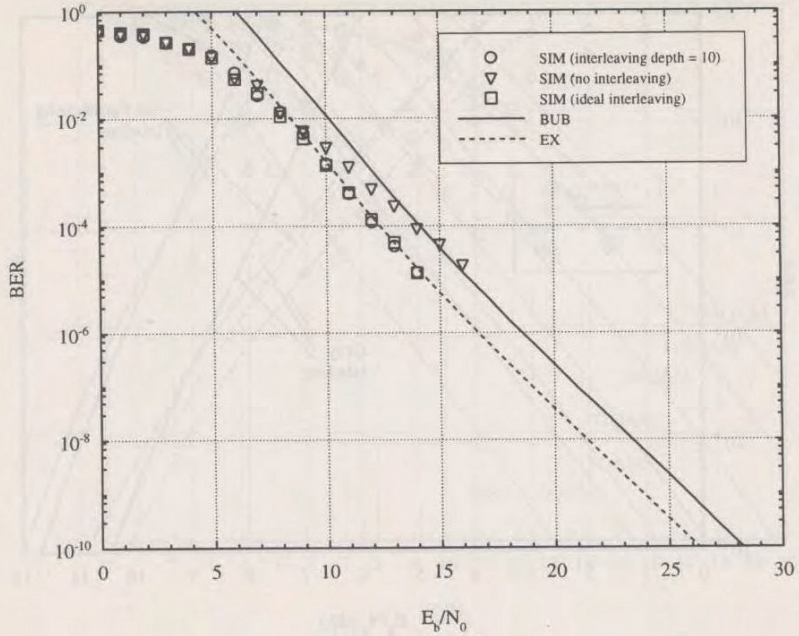



Fig. 10. BER of BICM 8PSK code obtained from the optimal 8-state, rate-2/3 code and Gray labeling. Rayleigh fading with perfect CSI.

TABLE I
VALUES OF $N_{\min}(1)$ AND $d_h^2$ FOR SOME PSK AND QAM SIGNAL SETS. (THE SIGNAL SET AVERAGE ENERGY IS NORMALIZED TO 1. LABELS "$\sim G$" AND "$W$" DENOTE QUASI-G RAY AND THE LABELING OF 32QAM PROPOSED BY WEI IN [25] FOR DESIGNING UNEQUAL ERROR PROTECTION CM)

| $\mathcal{X}$ | $\mu$ | $N_{\min}(1)$ | $d_h^2$ |
|---|---|---|---|
| 4PSK | Gray | 1 | 2 |
| | SP | 1.5 | 2 |
| 8PSK | Gray | 0.666 | 0.7664 |
| | SP | 1.166 | 0.664 |
| 8QAM | $\sim$ G | 0.9167 | 0.96 |
| | SP | 1.333 | 0.96 |
| 16QAM | Gray | 0.75 | 0.4923 |
| | SP | 1.75 | 0.4414 |
| 32QAM | $\sim$ G | 0.75 | 0.2788 |
| | W | 1.2 | 0.2133 |
| 64QAM | Gray | 0.5833 | 0.1442 |
| 128QAM | $\sim$ G | 0.6339 | 0.077 |
| 256QAM | Gray | 0.4688 | 0.043 |

class of binary parallel concatenated codes [15] (also known as "turbo codes").

### A. Numerical Results

In this subsection we show a selection of numerical results aimed at illustrating some features and applications of BICM. In the figures, curves marked by BUB denote union-Bhattacharyya bound, UB the BICM union bound, EX the BICM expurgated bound (or approximation), and SIM computer simulation. All the simulation results presented hereafter were obtained by using the suboptimal branch metric (9).

*1) Effect of Finite-Depth Interleaving:* Here we prove that interleaving is indeed necessary, although it need not be very deep if the channel has a short memory. Fig. 10 shows the BER of BICM over independent Rayleigh fading with perfect CSI, where $\mathcal{X}$ is 8PSK, $\mu$ is Gray labeling, and $\mathcal{C}$ is the optimal 8-state, rate-2/3 code used by Zehavi [26]. The only difference between this and Zehavi's BICM scheme is that here we use a single bit interleaver instead of separate interleavers for the three encoder outputs. Simulation in the case of ideal interleaving shows excellent agreement with the BICM EX approximation. The Bhattacharyya union bound is about 2 dB away from the true BER, as typical of fading channels. Zehavi's analysis based on the Chernoff bound [26] shows about the same gap from simulation. Hence, the usefulness of the tighter bounds developed in this paper is apparent. Simulations are also shown in the case of no interleaving and with interleaving depth equal to 10. Note that, since the Rayleigh fading channel used here is memoryless, the only effect of interleaving is to break the correlation introduced by the modulation, which carries three bits in a single transmitted signal. Absence of interleaving degrades the BER: however, a relatively short interleaver is sufficient to approach the ultimate performance.

*2) Gray Labeling:* Here we prove that SP labeling is poor for BICM. Fig. 11 shows the BER of BICM over AWGN with perfect CSI, where $\mathcal{X}$ is 16QAM and $\mathcal{C}$ is the *de facto* standard 64-state, rate-1/2 binary convolutional code with (octal) generators (133, 171) [17, pp. 466–471]. Gray and SP labelings are compared. Note that BICM EX is a true

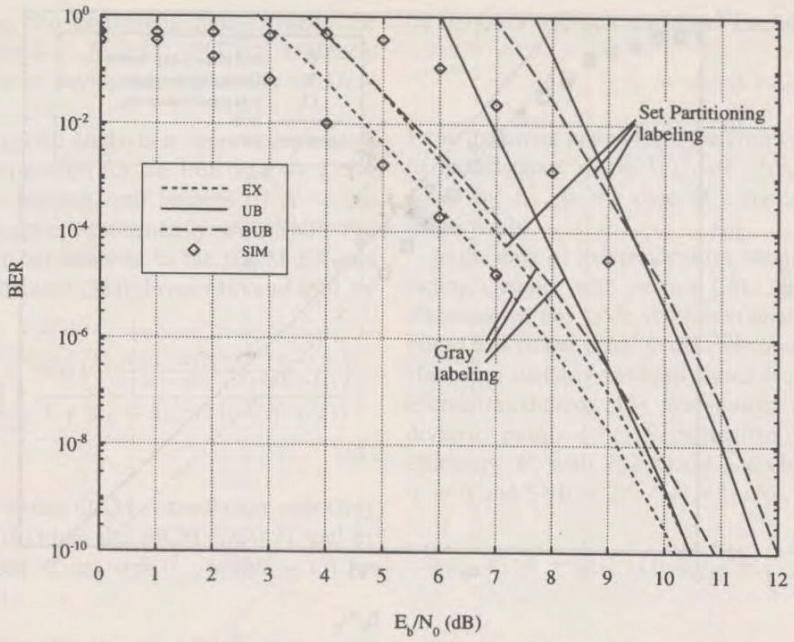



Fig. 11. BER of BICM 16QAM code obtained from the optimal 64-states rate-1/2 code. AWGN channel.

(very tight) bound only for Gray labeling. With SP labeling, the BICM EX curve underestimates the actual BER. This is because, without Gray labeling, there exist many more nearest neighbors for each transmitted signal sequence, while the BICM EX counts just one neighbor. Hence, if the labeling is not Gray, the BICM EX curve yields too optimistic values. However, the Bhattacharyya union bound and the BICM UB always provides an upper bound (also for the SP labeling), although the latter is rather loose in the range of BER values of practical interest.

*3) BICM PSK/QAM Codes for the Fading Channel:* Fig. 12 shows the BER over Rayleigh fading channels of BICM codes obtained by concatenating the same rate-1/2 code of the previous example with different signal sets. Gray labeling is used whenever possible. When $\mathcal{X}$ does not admit Gray labeling, a quasi-Gray labeling is chosen (i.e., a labeling that minimizes the number of points with more than one nearest neighbor whose label differ by more than one bit). Since the code is the same, all the BER curves have asymptotically the same slope. The $E_b/N_0$ (in decibels) gap between them can be evaluated from (65). The differences between the values of $(Rd_h^2)_{dB}$ and those pertaining to 4PSK are reported in Table II. By comparing these values with the curves in Fig. 12, we note how (65) gives a quite accurate prediction. Note also that BICM EX gives true upper bounds for Gray labeling, but only an approximation in the quasi-Gray case.

A method for increasing code diversity with BICM consists of concatenating an expanded signal set to a low-rate code [13]. For given $R$ and $\tau$, this leads to asymptotically steeper BER curves. However, signal set expansion reduces the value of $d_h^2$, so that this technique may not provide a better performance in the range of BER values of interest. The

TABLE II
SPECTRAL EFFICIENCIES AND DIFFERENCES IN THE VALUES OF $(Rd_h^2)_{dB}$ FOR THE BICM CODES OF FIG. 12

| $\mathcal{X}$ | $R$ (bit/dim) | $E_b/N_0$ gap (dB) |
|---|---|---|
| 4PSK | 1 | 0 |
| 8PSK | 1.5 | 2.4 |
| 16QAM | 2 | 3.07 |
| 32QAM | 2.5 | 4.57 |
| 64QAM | 3 | 6.65 |
| 128QAM | 3.5 | 8.7 |
| 256QAM | 4 | 10.65 |

crossover point between different BICM codes over Rayleigh fading can be coarsely estimated by finding the intersection between the straight lines defined by (65) (by neglecting the "const." terms). Fig. 13 shows two examples of BICM design based on signal set expansion. We considered two BICM schemes with $R = 2$ bit/dim and $\tau = 128$, one obtained by concatenating to 8PSK the best rate-2/3, 64-state punctured code [7] and the other obtained by concatenating to 16QAM the best rate-1/2, 64-state code of previous example (both with Gray labeling). From (65) we estimate the crossover at $E_b/N_0 = 2.95$ dB, which indicates that in this case signal set expansion yields a coding gain at almost any BER of interest. Next, we considered two BICM codes with $R = 1$ bit/dim and $\tau = 128$, one obtained by concatenating to 4PSK the same best rate-1/2, 64-state code and the other obtained by concatenating to 16QAM the best rate-1/4, 64-state code [17, pp. 466–471] (both with Gray labeling). From (65) we estimate the crossover at $E_b/N_0 = 9.16$ dB, which indicates that in this case signal

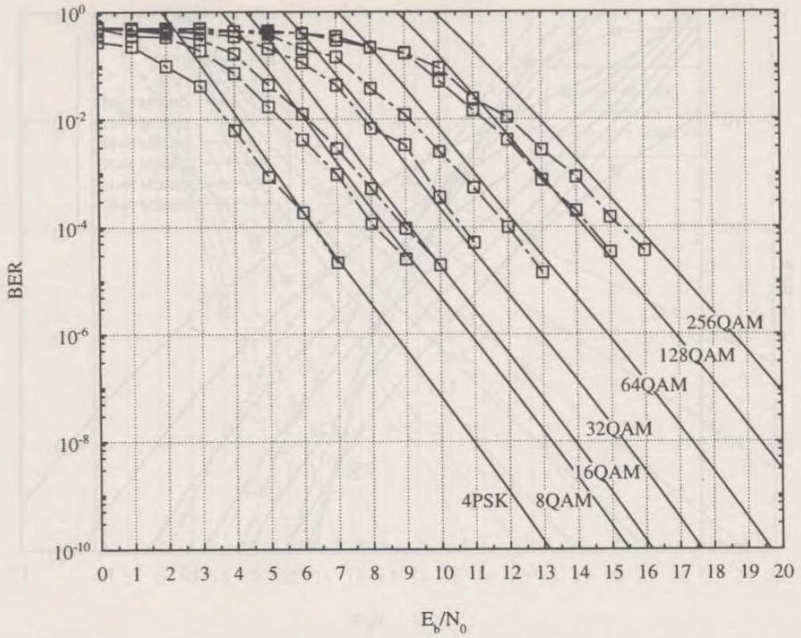



Fig. 12. BER of BICM obtained from the optimal 64-state rate-1/2 code and PSK/QAM signal sets. Rayleigh fading with perfect CSI.

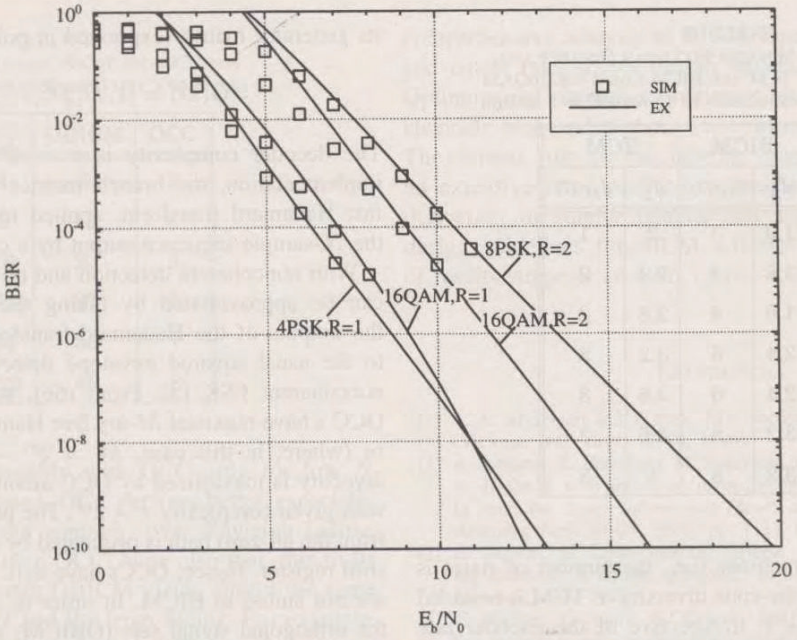



Fig. 13. BER of BICM for $R = 2$ and $R = 1$ with $\tau = 128$ (64 states). The codes are obtained by concatenating the best rate-2/3 punctured code with 8PSK, the best rate-1/2 code with 16QAM, the best rate-1/2 code with 4PSK and the best rate-1/4 code with 16QAM. Rayleigh fading with perfect CSI.

set expansion yields a coding gain only at very low BER values. The actual BER curves intersect at $E_b/N_0 = 11$ dB.

Finally, we can compare BICM and TCM by using bounds on the achievable minimum Euclidean distance and code diversity. For BICM, we can use Heller bound and its extension to rate $k_c/n_c$ codes as provided in [11]. For TCM, sphere-packing bounds on $d_E$ can be found in [5, Ch. 4] and references therein. In the case of Ungerboeck's TCM (i.e., TCM obtained from a binary encoder for $C$), the code diversity is given by the $M$-ary Hamming distance $d_M(C)$ of $C$ ($M = 2^m$), considered as a $k_c$-to-1 binary input $M$-ary output code. In this case, we can use the upper bound on the $M$-ary Hamming distance of [16]

$$d_M(C) \le \min\left\{\left\lfloor\frac{v}{k_c}\right\rfloor + 1, \min_{tk_c \ge m}\left\{\frac{M-1}{M}\frac{2^{tk_c}}{2^{tk_c}-1}\left(\left\lfloor\frac{v}{k_c}\right\rfloor + t\right)\right\}\right\} \quad (66)$$

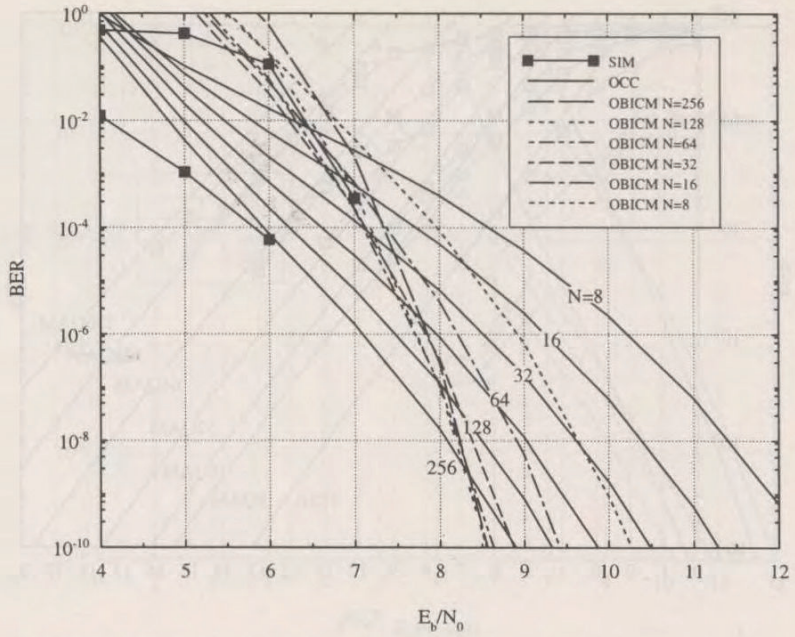



Fig. 14. OBICM and OCC over AWGN with noncoherent detection. Simulation results are shown for $N = 256$ only.

TABLE III
UPPER BOUNDS TO MINIMUM EUCLIDEAN DISTANCE AND CODE DIVERSITY FOR TCM AND BICM CODES FOR 16QAM (AVERAGE ENERGY NORMALIZED TO 1) WITH $R = 3$ bit/dim

| Encoder | BICM | | TCM | |
|---|---|---|---|---|
| memory | $d_E^2$ | $d_2(\mathcal{C})$ | $d_E^2$ | $d_M(\mathcal{C})$ |
| 2 | 1.2 | 3 | 2 | 1 |
| 3 | 1.6 | 4 | 2.4 | 2 |
| 4 | 1.6 | 4 | 2.8 | 2 |
| 5 | 2.4 | 6 | 3.2 | 2 |
| 6 | 2.4 | 6 | 3.6 | 3 |
| 7 | 3.2 | 8 | 3.6 | 3 |
| 8 | 3.2 | 8 | 4 | 3 |

where $v$ is the encoder memory (i.e., the number of states is $N_S = 2^v$). In particular, the code diversity of TCM is bounded from above by $\lfloor v/k_c \rfloor + 1$, irrespective of the encoder rate $k_c/m$. Table III shows the bounds on the minimum Euclidean distance and on the minimum code diversity of BICM and TCM, for 16QAM codes with $R = 3$ bit/dim. As expected from the discussions above, for the same complexity BICM achieves better code diversity, while TCM achieves better minimum Euclidean distance.

*BICM Codes for DS/SSMA with Noncoherent Detection:* Viterbi's orthogonal convolutional codes (OCC) have been proposed for coding and spreading in DS/SSMA. These codes can be seen as TCM schemes, where a binary encoder of rate $1/m$ and $2^{m-1}$ states is concatenated to a Hadamard modulator which generates an $N = 2^m$-dimensional orthogonal signal set. The binary encoder is a simple shift register, and its generator matrix (expressed in polynomial form) is

$$G(D) = [1, D, D^2, \cdots, D^{m-1}].$$

The decoder complexity is $\tau = 2^m = N$. In a practical implementation, the branch metrics can be computed by a fast Hadamard transform, applied to the received vector $\boldsymbol{y}$, the $N$-sample sequence output by a chip matched filter [22].

With noncoherent detection and no CSI, the branch metrics can be approximated by taking the squared magnitude of the outputs of the Hadamard transform [17], [12], similarly to the usual squared envelope detector used for orthogonal noncoherent FSK [8]. From (66), we get immediately that OCC's have maximal $M$-ary free Hamming distance $d_M(\mathcal{C}) = m$ (where, in this case, $M = 2^m = N$). Hence, the code diversity is maximized by OCC among all the TCM schemes with given complexity $\tau = 2^m$. The path at $M$-ary distance $m$ from the all-zero path is originated by a single "1" entering the shift register. Hence, OCCs have $d_2(\mathcal{C}) = d_M(\mathcal{C})$, so that they are not suited to BICM. In order to design a BICM scheme for orthogonal signal sets (OBICM) with the same decoding complexity $\tau = 2^m$ of OCC's (and the same demodulator described above), we pick the optimum binary convolutional codes of rate $1/m$ and $2^{m-1}$ states [17, pp. 466–471]. After ideal bit interleaving, $m$-bit labels are mapped onto the $N$-ary orthogonal signal set. The resulting spectral efficiency is still, as for OCC, $R = 1/N$ bit/dim, but now the code diversity is $d_2(\mathcal{C})$ (see Table IV). Figs. 14 and 15 show the BER of OCC and OBICM with noncoherent detection over AWGN and Rayleigh fading, respectively, for $N = 8, 16, 32, 64, 128, 256$. In the case of OCC, we can evaluate the union bound with exact PEP computations obtained by following [4] (these curves are labeled "TUB"). As for OBICM, we used our Bhattacharyya union bound, which is fairly tight in this case.

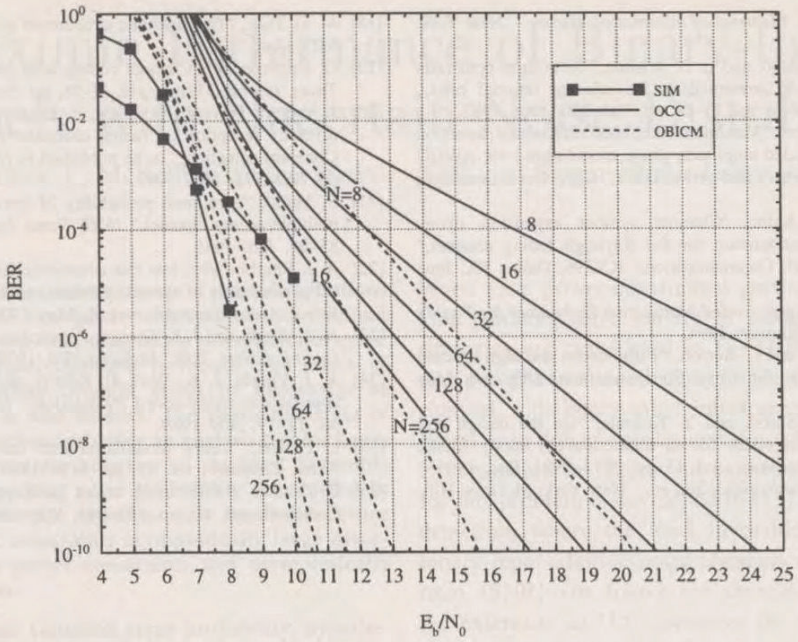



Fig. 15. OBICM and OCC over Rayleigh fading with noncoherent detection and no CSI. Simulation results are shown for $N = 256$ only.

TABLE IV
CODE DIVERSITY OF OBICM AND OCC FOR THE SAME DECODER COMPLEXITY AND RATE.

| $N$ | $N_S$ | OBICM | OCC |
|---|---|---|---|
| 8 | 4 | 8 | 3 |
| 16 | 8 | 13 | 4 |
| 32 | 16 | 20 | 5 |
| 64 | 32 | 27 | 6 |
| 128 | 64 | 36 | 7 |
| 256 | 128 | 45 | 8 |

OBICM compares favorably with OCC only for low $N$ over AWGN. As $N$ increases, OCC perform better, especially for BER $\geq 10^{-5}$. On the contrary, over Rayleigh fading, OBICM performs better than OCC. Note also that, due to its high code diversity, low-rate OBICM yields almost the same performance over AWGN and Rayleigh fading. For example, for $N = 256$ the performance loss at BER $= 10^{-5}$ due to Rayleigh fading with respect to AWGN is only 0.5 dB, while the corresponding loss of OCC with $N = 256$ is about 4.5 dB. This makes OBICM an interesting solution for coded DS/SSMA with noncoherent detection over channels where fading may range from AWGN to Rayleigh depending on the propagation environment.

## VI. CONCLUSION

The main theme of this paper is that on some channels the separation of demodulation and decoding might be beneficial, provided that the encoder output is interleaved bit-wise and a suitable soft-decision metric is used in the Viterbi decoder. A comprehensive analysis of BICM, based on channel capacity and cutoff rate, shows this in information-theoretical terms. Optimum and simpler, suboptimum bit metrics are derived for channels with and without state information at the receiver. The central role of the labeling map is pinpointed, while an extensive error probability analyisis, which includes the derivation of sundry bounds and approximations, leads to design guidelines for BICM schemes. A comprehensive set of results suggests an array of possible applications.

## REFERENCES

[1] S. A. Al-Semari and T. Fuja, "Bit interleaved I-Q TCM," in *ISITA'96* (Victoria, B.C., Sept. 17–20, 1996).

[2] A. Aoyama, T. Yamazato, M. Katayama, and A. Ogawa, "Performance of 16-QAM with increased diversity on Rayleigh fading channels," in *Proc. Int. Symp. Information Theory and Its Applications* (Sydney, Australia, Nov. 20–24, 1994), pp. 1133–1137.

[3] E. Biglieri, G. Caire, and G. Taricco, "Error probability over fading channels: A unified approach," to be published in *Europ. Trans. Commun.*, July 1997.

[4] E. Biglieri, G. Caire, G. Taricco, and J. Ventura, "Simple method for evaluating error probabilities," *Electron. Lett.*, vol. 32, no. 3, pp. 191–192, Feb. 1996.

[5] E. Biglieri, D. Divsalar, P. J. McLane, and M. K. Simon, *Introduction to Trellis-Coded Modulation with Applications*. New York: MacMillan, 1991.

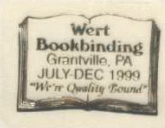
[6] J. Boutros, E. Viterbo, C. Rastello, and J.-C. Belfiore, "Good lattice constellations for both Rayleigh fading and Gaussian channels," *IEEE Trans. Inform. Theory*, vol. 42, pp. 502–518, Mar. 1996.

[7] J. B. Cain, G. C. Clark, and J. M. Geist, "Punctured convolutional codes of rate $(n-1)/n$ and simplified maximum likelihood decoding," *IEEE Trans. Inform. Theory*, vol. IT-25, pp. 97–100, Jan. 1979.

[8] G. Caire, J. Ventura, and E. Biglieri, "Coded and pragmatic-coded orthogonal modulation for the fading channel with non-coherent detection," in *Proc. IEEE Int. Conf. Communications (ICC'95)* (Seattle, WA, June 18–22), 1995.

[9] G. Caire, G. Taricco, and E. Biglieri, "Capacity and cut-off rate of bit-interleaved channels," in *Proc. Int. Symp. Information Theory and Its Applications ISITA'96* (Victoria, BC, Canada, Sept. 17–20, 1996).

[10] T. Cover and J. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.
[11] D. G. Daut, J. W. Modestino, and L. D. Wismer, "New short constraint length convolutional code constructions for selected rational rates," *IEEE Trans. Inform. Theory*, vol. IT-28, pp. 794–800, Sept. 1982.
[12] D. Divsalar and M. Simon, "Maximum-likelihood differential detection of uncoded and trellis coded amplitude phase modulation over AWGN and fading channels—Metrics and performance," *IEEE Trans Commun.*, vol. 42, Jan. 1994.
[13] U. Hansson and T. Aulin, "Channel symbol expansion diversity—Improved coded modulation for the Rayleigh fading channel," presented at the Int. Conf. Communications, ICC'96, Dallas TX, June 23–27, 1996.
[14] S. H. Jamali and T. Le-Ngoc, *Coded-Modulation Techniques for Fading Channels*. New York: Kluwer, 1994.
[15] S. Le Goff, A. Glavieux, and C. Berrou, "Turbo-codes and high spectral efficiency modulation," in *Int. Conf. Communications (ICC'94)*, May 1996, pp. 645–649.
[16] G. Kaplan, S. Shamai (Shitz), and Y. Kofman, "On the design and selection of convolutional codes for an uninterleaved, bursty Rician channel," *IEEE Trans. Commun.*, vol. 43, pp. 2914–2921, Dec. 1995.
[17] J. Proakis, *Digital Communications*, 2nd ed. New York: McGraw-Hill, 1989.
[18] M. D. Trott, "The algebraic structure of trellis codes," Ph.D. dissertation, MIT, Aug. 1992.
[19] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Trans. Inform. Theory*, vol. IT-28, pp. 56–67, Jan. 1982.
[20] J. Ventura-Traveset, G. Caire, E. Biglieri, and G. Taricco, "Impact of diversity reception on fading channels with coded modulation. Part I: Coherent detection," to be published in *IEEE Trans. Commun.*, vol. 45, pp. 563–572, May 1997.
[21] S. Verdú, "Minimum probability of error for asynchronous Gaussian multiple-access channel," *IEEE Trans. Inform. Theory*, vol. IT-32, pp. 85–96, Jan. 1986.
[22] A. J. Viterbi, "Very low rate convolutional codes for maximum theoretical performance of spread-spectrum multiple-access channels," *IEEE J. Select. Areas Commun.*, vol. 8, May 1990.
[23] A. J. Viterbi and J. K. Omura, *Principles of Digital Communication and Coding*. New York: McGraw-Hill, 1979.
[24] A. J. Viterbi, J. K. Wolf, E. Zehavi, and R. Padovani, "A pragmatic approach to trellis-coded modulation," *IEEE Commun. Mag.*, vol. 27, pp. 11–19, July 1989.
[25] L. F. Wei, "Coded modulation with unequal error protection," *IEEE Trans. Commun.*, vol. 41, pp. 1439–1449, Oct. 1993.
[26] E. Zehavi, "8-PSK trellis codes for a rayleigh channel," *IEEE Trans. Commun.*, vol. 40, pp. 873–884, May 1992.


Wert
Bookbinding
Grantville, PA
JULY-DEC 1999
"We're Quality Bound"

# APPENDIX DEGAUDENZI01

*screenshot-catalog.libraries.psu.edu-2022.07.27-15_09_17*
*https://catalog.libraries.psu.edu/catalog/38176353*
*27.07.2022*

Scheduled maintenance: Item availability, I Want It, and My Account management will be unavailable on July 13, 2022 from approximately 6:15-7am ET.

PennState University Libraries | Catalog

Bookmarks 0 Course Reserves My Account

Keyword Search... Search Advanced search Start Over

Share Bookmark Report an Issue

# MARC View

| | | | |
|---|---|---|---|
| LEADER 02425nas a2200493 a 4500 | | | |
| 001 38176353 | | | |
| 003 SIRSI | | | |
| 005 20220714090051.0 | | | |
| 006 m o d | | | |
| 007 cr n | | | |
| 008 021021c20039999enkbr pss 0 a0eng c | | | |
| 010 | | | a\| 2002215229 |
| 015 | | | a\| GBA600060 2\| bnb |
| 016 | 7 | | a\| 013345284 2\| Uk |
| 019 | | | a\| DO NOT EDIT this 360 MARC RECORD |
| 022 | | | a\| 1542-0981 l\| 1542-0973 y\| 1542-0973 |
| 035 | | | a\| ss360mi0118420 |
| 037 | | | b\| Journals Subscriptions Department, John Wiley & Sons, Ltd., 1 Oldlands Way, Bognor Regis, West Sussex PO22 9SA, UK |
| 040 | | | a\| NSD b\| eng c\| NSD d\| CUS d\| FUG d\| OCLCQ d\| NSD d\| IUL d\| MUQ d\| OCLCQ d\| UKMGB d\| OCLCQ d\| OCLCF d\| OCLCO d\| VT2 d\| OCLCO d\| OCLCQ d\| NLE d\| OCLCQ d\| EZC d\| OCLCQ d\| NJT d\| SFB d\| CNMTR d\| BWN d\| OCLCO d\| WaSeSS |
| 042 | | | a\| pcc a\| nsdp |
| 050 | | 4 | a\| TK5104 b\| .I5772 |
| 082 | 1 | 0 | a\| 621 2\| 13 |
| 130 | 0 | | a\| International journal of satellite communications and networking (Online) |
| 210 | 0 | | a\| Int. j. satell. commun. netw. b\| (Online) |
| 222 | | 0 | a\| International journal of satellite communications and networking b\| (Online) |
| 245 | 1 | 0 | a\| International journal of satellite communications and networking h\| [electronic resource]. |
| 246 | 3 | 0 | a\| Satellite communications and networking |
| 260 | | | a\| [Chichester, West Sussex] : b\| John Wiley & Sons |
| 310 | | | a\| Bimonthly |
| 362 | 0 | | a\| Vol. 21, issue 1 (Jan./Feb. 2003)- |
| 500 | | | a\| Title from table of contents (Wiley InterScience, viewed: Feb. 15, 2006). |
| 530 | | | a\| Also issued in print. |
| 538 | | | a\| Mode of access: World Wide Web. |
| 588 | | | a\| Latest issue consulted: Vol. 24, issue 1 (Jan./Feb. 2006); Wiley InterScience, viewed: Feb. 15, 2006. |
| 650 | | 0 | a\| Digital communications v\| Periodicals. |
| 650 | | 0 | a\| Artificial satellites in telecommunication v\| Periodicals. |
| 655 | | 0 | a\| Electronic journals. |
| 776 | 1 | | t\| International journal of satellite communications and networking x\| 1542-0973 w\| (DLC) 2002215228 w\| (OCoLC)50812246 |
| 780 | 0 | 0 | t\| International journal of satellite communications (Online) x\| 1099-1247 w\| (DLC) 2001212310 w\| (OCoLC)44069350 |
| 856 | 4 | 0 | u\| http://SK8ES4MC2L.search.serialssolutions.com/?V=1.0&L=SK8ES4MC2L&S=JCs&C=INTEJOUROFSAC&T=marc |
| 910 | | | a\| CONSER online |
| 853 | 2 | 0 | 8\| 1 a\| v. b\| issue u\| 6 v\| r i\| (year) j\| (month/month) w\| b x\| 01/02 |
| 863 | 4 | 1 | 8\| 1.1 a\| 21 b\| 1 i\| 2003 j\| 01/02 |
| 949 | | | a\| Electronic resource w\| ASIS c\| 1 i\| 38176353-1001 l\| ONLINE m\| ONLINE r\| Y s\| Y t\| ONLINE u\| 7/16/2022 o\| .STAFF. 360 MARC |

ASK A LIBRARIAN

ASK A LIBRARIAN

PennState

Copyright @2019 The Pennsylvania State University. All rights reserved. Except where otherwise noted, this work is subject to a Creative Commons Attribution 4.0 license. Details and exceptions.

Legal Statements | PSU Hotlines

PENN STATE UNIVERSITY LIBRARIES

Libraries Home (814) 865-6368
Libraries Intranet (Staff Only)
Accessibility Help
Website Feedback
Policies and Guidelines
Staff Directory

CONNECT WITH PENN STATE UNIVERSITY LIBRARIES

Facebook
Twitter
Instagram

screenshot-catalog.libraries.psu.edu-2022.07.27-15_09_17
https://catalog.libraries.psu.edu/catalog/38176353
27.07.2022

Scheduled maintenance: Item availability, I Want It, and My Account management will be unavailable on July

PennState University Libraries | Catalog

Keyword | Search... | Search | Advanc

# International journal of satellite co and networking [electronic resourc

Uniform Title: International journal of satellite communications an

Additional Titles: Int. j. satell. commun. netw. (Online) and Satellite co

Published: [Chichester, West Sussex] : John Wiley & Sons

Access Online

serialssolutions.com

**Continues:**
International journal of satellite communications (Online) 1099-1247

**Dates of Publication and/or Sequential Designation:**
Vol. 21, issue 1 (Jan./Feb. 2003)-

**Subject(s):**
Digital communications—Periodicals
Artificial satellites in telecommunication—Periodicals

**Genre(s):**
Electronic journals

**ISSN:**
1542-0981

**Publication Frequency:**
Bimonthly

**Note:**
Title from table of contents (Wiley InterScience, viewed: Feb. 15, 2006).

**Other Forms:**
Also issued in print.

**Technical Details:**
Mode of access: World Wide Web.

PennState

Copyright ©2019 The Pennsylvania State University. All rights reserved. Except where otherwise noted, this work is subject to a Creative Commons Attribution 4.0 license. Details and exceptions.

Legal Statements | PSU Hotlines

PENN STATE UNIVERSITY LIBRARIES

Libraries Home (814) 865-6

Libraries Intranet (Staff Only)

Accessibility Help


Website Feedback
Policies and Guidelines
Staff Directory

Advertisement

international journal of satellite communications and networking

Volume 24, Issue 4

Pages: 261-317
July/August 2006

< Previous Issue | Next Issue >

Export Citation(s)

# Research Articles

Full Access

## Turbo-coded APSK modulations design for satellite broadband communications

Riccardo De Gaudenzi, Albert Guillén i Fàbregas, Alfanso Martinez

Pages: 261-281 | First Published: 19 May 2006

Abstract | PDF | References | Request permissions

Full Access

## Review and comparison of tropospheric scintillation prediction models for satellite communications

P. Yu, I. A. Glover, P. A. Watson, O. T. Davies, S. Ventouras, C. Wrench

Pages: 283-302 | First Published: 18 May 2006

Abstract | PDF | References | Request permissions

Full Access

## Semi-random LDPC codes for CDMA communication over non-linear band-limited satellite channels

Mohamed Adnan Landolsi

Pages: 303-317 | First Published: 08 June 2006



Tools

Submit an Article

Browse free sample issue

Get content alerts

Subscribe to this journal

More from this journal

Special Issues
To Our Authors Newsletter
La Tex Class File

DIVERSITY
in Research Jobs

Please contact us to see your job listed here

**Senior Research Technician - College of Engineering and Physical Sciences**
**United Kingdom | £24,291 to £30,207**

College of Engineering and Physical Sciences Location: University of Birmingham, Edgbaston, Birmingham, UK Full time starting salary is normally in...

UNIVERSITY OF BIRMINGHAM

Employer: University of Birmingham

Apply for this job

**Senior Technician - School of Engineering Band 500**
**United Kingdom | £24,291 to £30,207**

School of Engineering in the College of Engineering and Physical Sciences Location: University of Birmingham, Edgbaston, Birmingham, UK Full time s...

UNIVERSITY OF BIRMINGHAM

Employer: University of Birmingham

Apply for this job

**Lecturer in Mechanical Engineering (Vehicle Dynamics)**
**Leeds, United Kingdom | £34,304 to £40,927 p.a.**

About Wiley Online Library

Privacy Policy
Terms of Use
About Cookies
Manage Cookies
Accessibility
Wiley Research DE&I Statement and Publishing Policies
Developing World Access

Help & Support

Contact Us
Training and Support
DMCA & Reporting Piracy

Opportunities

Subscription Agents
Advertisers & Corporate Partners

Connect with Wiley

The Wiley Network
Wiley Press Room

Copyright © 1999-2022 John Wiley & Sons, Inc. All rights reserved

# APPENDIX DEGAUDENZI02

INTERNATIONAL JOURNAL OF SATELLITE COMMUNICATIONS AND NETWORKING
*Int. J. Satell. Commun. Network.* 2006; **24**:261–281
Published online 19 May 2006 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/sat.841

# Turbo-coded APSK modulations design for satellite broadband communications

Riccardo De Gaudenzi[1,*,†], Albert Guillén i Fàbregas[2], and Alfanso Martinez[3]

[1] *European Space Agency (ESA-ESTEC), Noordwijk, The Netherlands*
[2] *Institute for Telecommunications Research, University of South Australia, Australia*
[3] *Technische Universitat Eindhoven, Eindhoven, The Netherlands*

## SUMMARY

This paper investigates the design of power and spectrally efficient coded modulations based on amplitude phase shift keying (APSK) modulation with application to satellite broadband communications. APSK represents an attractive modulation format for digital transmission over nonlinear satellite channels due to its power and spectral efficiency combined with its inherent robustness against nonlinear distortion. For these reasons APSK has been very recently introduced in the new standard for satellite Digital Video Broadcasting named DVB-S2. Assuming an ideal rectangular transmission pulse, for which no nonlinear inter-symbol interference is present and perfect pre-compensation of the nonlinearity, we optimize the APSK constellation. In addition to the minimum distance criterion, we introduce a new optimization based on the mutual information; this new method generates an optimum constellation for each spectral efficiency. To achieve power efficiency jointly with low bit error rate (BER) floor we adopt a powerful binary serially concatenated turbo-code coupled with optimal APSK modulations through bit-interleaved coded modulation. We derive tight approximations on the maximum-likelihood decoding error probability, and results are compared with computer simulations. The proposed coded modulation scheme is shown to provide a considerable performance advantage compared to current standards for satellite multimedia and broadcasting systems. Copyright © 2006 John Wiley & Sons, Ltd.

Received 1 May 2005; Revised 1 February 2006; Accepted 29 March 2006

KEY WORDS: turbo codes; amplitude-phase shift keying (APSK); modulation; bit-interleaved coded modulation (BICM); coded modulation; nonlinear channels; satellite communications

## 1. INTRODUCTION

A major strength of satellite communications systems lies on their ability to efficiently broadcast digital multi-media information over very large areas [1]. A notable example is the

*Correspondence to: Riccardo De Gaudenzi, European Space Agency (ESA-ESTEC), RF Payload Systems division, Keplerlaan 1, P.O. Box 299, AG Noordwijk, The Netherlands.
† E-mail: riccardo.de.gaudenzi@esa.int

Copyright © 2006 John Wiley & Sons, Ltd.

so-called direct-to-home (DTH) digital television broadcasting. Satellite systems also provide a unique way to complement the terrestrial telecommunication infrastructure in scarcely populated regions. The introduction of multi-beam satellite antennas with adaptive coding and modulation (ACM) schemes will allow an important efficiency increase for satellite systems operating at Ku or Ka-band [2]. Those technical enhancements require the exploitation of power- and spectrally efficient modulation schemes conceived to operate over the satellite nonlinear channel. In this paper, we will design high-efficiency 16- and 32-ary coded modulation schemes suited for nonlinear satellite channels. The analysis presented here is complemented in [3] with the effects related to satellite nonlinear distortion, band-limited transmission pulse, demodulator timing, amplitude and phase estimation errors.

To the authors' knowledge there are few references in the literature dealing with 16-ary constellation optimization over *nonlinear channels*, the typical environment for satellite channels. Previous work showed that 16-QAM does not compare favourably with either trellis-coded (TC) 16-PSK or uncoded 8-PSK in satellite nonlinear channels [4]. The concept of circular APSK modulation was already proposed 30 years ago by Thomas *et al*. [5], where several nonband-limited APSK sets were analysed by means of uncoded bit error rate bounds; the suitability of APSK for nonlinear channels was also made explicit, but concluded that for single carrier operation over nonlinear channel APSK performs worse than PSK schemes. In the current paper, we will revert the conclusion. It should be remarked that Reference [5] mentioned the possibility of modulator pre-compensation but did not provide performance results related to this technique. Foschini *et al*. [6] optimized QAM constellations using asymptotic uncoded probability of error under average power constraints, deriving optimal 16-ary constellation made of an almost equilateral lattice of triangles. This result is not applicable to satellite channels. In Reference [7] some comparison between squared QAM and circular APSK over linear channels was performed based on the computation of the error bound parameter, showing some minor potential advantage of APSK. Further work on mutual information for modulations with average and peak power constraints is reported in Reference [8], which proves the advantages of circular APSK constellations under those power constraints. Mutual information performance loss for APSK in peak power limited Gaussian complex channels is reported in Reference [9] and compared to classical QAM modulations; it is shown that under this assumption APSK considerably outperforms QAM in terms of mutual information, the gain particularly remarkable for 16- and 64-ary constellations.

Forward error correcting codes for our application must combine power efficiency and low BER floor with flexibility and simplicity to allow for high-speed implementation. The existence of practical, simple, and powerful such coding designs for binary modulations has been settled with the advent of turbo codes [10] and the recent re-discovery of low-density parity-check (LDPC) codes [11]. In parallel, the field of channel coding for nonbinary modulations has evolved significantly in the latest years. Starting with Ungerboeck's work on TC modulation (TCM) [12], the approach had been to consider channel code and modulation as a single entity, to be jointly designed and demodulated/decoded. Schemes have been published in the literature, where turbo codes are successfully merged with TCM [13]. Nevertheless, the elegance and simplicity of Ungerboeck's original approach gets somewhat lost in a series of *ad hoc* adaptations; in addition, the turbo-code should be jointly designed with a given modulation, a solution impractical for system supporting several constellations. A new pragmatic paradigm has crystallized under the name of bit-interleaved coded modulation (BICM) [13], where extremely good results are obtained with a standard nonoptimized, code. An additional

Copyright © 2006 John Wiley & Sons, Ltd.

DOI: 10.1002/sat

advantage of BICM is its inherent flexibility, as a single mother code can be used for several modulations, an appealing feature for broadband satellite communication systems where a large set of spectral efficiencies is needed.

This paper is organized as follows. Section 2 gives the system model under the ideal case of a rectangular transmission pulse.[‡] Section 3 gives a formal description of APSK signal sets, describes the maximum mutual information and maximum minimum distance optimization criteria and discusses some of the properties of the optimized constellations. Section 4 deals with code design issues, describes the BICM approach, provides some analytical considerations based on approximate maximum-likelihood (ML) decoding error probability bounds, and provides some numerical results. The conclusions are finally drawn in Section 5.

## 2. SYSTEM MODEL

The baseband equivalent of the transmitted signal at time $t$, $s_T(t)$, is given by

$$s_T(t) = \sqrt{P}\sum_{k=0}^{L-1} x(k)p_T(t - kT_s) \tag{1}$$

where $P$ is the signal power, $x(k)$ is the $k$th transmitted symbol, drawn from a complex-valued APSK signal constellation $\mathcal{X}$, with $|\mathcal{X}| = M$, $p_T$ is the transmission filter impulse response, and $T_s$ is the symbol duration (in seconds), corresponding to one channel use. Without loss of generality, we consider transmission of frames with $L$ symbols. The spectral efficiency $R$ is defined as the number of information bits conveyed at every channel use, and in measured in bits per second per Hertz (bps/Hz).

The signal $s_T(t)$ passes through a high-power amplifier (HPA) operated close to the saturation point. In this region, the HPA shows nonlinear characteristics that induce phase and amplitude distortions to the transmitted signal. The amplifier is modelled by a memoryless nonlinearity, with an output signal $s_A(t)$ at time $t$ given by

$$s_A(t) = F(|s_T(t)|)e^{j(\phi(s_T(t))+\Phi(|s_T(t)|))} \tag{2}$$

where we have implicitly defined $F(A)$ and $\Phi(A)$ as the AM/AM and AM/PM characteristics of the amplifier for a signal with instantaneous signal amplitude $A$. The signal amplitude is the instantaneous complex envelope, so that the baseband signal is decomposed as $s_T(t) = |s_T(t)|e^{j\phi(s_T(t))}$.

In this paper, we assume an (ideal) signal modulating a train of rectangular pulses. These pulses do not create inter-symbol interference when passed through an amplifier operated in the nonlinear region. Under these conditions, the channel reduces to an AWGN, where the modulation symbols are distorted following (2). Let $x_A$ denote the distorted symbol corresponding to $x = |x|e^{j\phi(x)} \in \mathcal{X}$, that is, $x_A = F(|x|)e^{j(\phi(x)+\Phi(|x|))}$. After matched filtering and sampling at time $kT_s$, the discrete-time received signal at time $k$, $y(k)$ is then given by

$$y(k) = \sqrt{E_s}x_A(k) + n(k), \quad k = 0, \ldots, L-1 \tag{3}$$

[‡]This assumption has been dropped in the paper [14].

Copyright © 2006 John Wiley & Sons, Ltd.

DOI: 10.1002/sat

with $E_s$ the symbol energy, given by $E_s = PT_s$, $x_A(k)$ is the symbol at the $k$th time instant, as defined above, and $n(k) \sim \mathcal{N}_{\mathbb{C}}(0, N_0)$ is the corresponding noise sample.

This simplified model suffices to describe the nonlinearity up to the nonlinear ISI effect, and allows us to easily design constellation and codes. In the paper [14], the impact of nonlinear ISI has been considered, as well as other realistic demodulation effects such as timing and phase recovery.

## 3. APSK CONSTELLATION DESIGN

In this section, we define the generic multiple-ring APSK constellation family. We propose new criteria for the design of digital QAM constellations of 16 and 32 points, with special emphasis on the behaviour on nonlinear channels.

### 3.1. Constellation description

$M$-APSK constellations are composed of $n_R$ concentric rings, each with uniformly spaced PSK points. The signal constellation points $x$ are complex numbers, drawn from a set $\mathcal{X}$ given by

$$\mathcal{X} = \begin{cases} r_1 e^{j((2\pi/n_1)i+\theta_1)}, & i = 0, \ldots, n_1 - 1 \quad \text{(ring 1)} \\ r_2 e^{j((2\pi/n_2)i+\theta_2)}, & i = 0, \ldots, n_2 - 1 \quad \text{(ring 2)} \\ \vdots & \\ r_{n_R} e^{j((2\pi/n_R)i+\theta_{n_R})}, & i = 0, \ldots, n_{n_R} - 1 \quad \text{(ring } n_R\text{)} \end{cases} \tag{4}$$

where we have defined $n_\ell$, $r_\ell$ and $\theta_\ell$ as the number of points, the radius and the relative phase shift for the $\ell$th ring. We will nickname such modulations as $n_1 + \cdots + n_{n_R}$-APSK. Figure 1 depicts the $4 + 12$- and $4 + 12 + 16$-APSK modulations with quasi-Gray mapping. In particular, for next generation broadband systems [2, 15], the constellation sizes of interest are $|\mathcal{X}| = 16$ and 32, with $n_R = 2$ and 3 rings, respectively. In general, we consider that $\mathcal{X}$ is normalized in energy, i.e. $E[|x|^2] = 1$, which implies that the radii $r_\ell$ are normalized such that $\sum_{\ell=1}^{n_R} n_\ell r_\ell^2 = 1$. Notice also that the radii $r_\ell$ are ordered, so that $r_1 < \cdots < r_{n_R}$.

Clearly, we can also define the phase shifts and the ring radii in relative terms rather than in absolute terms, as in (4); this removes one dimension in the optimization process, yielding a practical advantage. We let $\phi_\ell = \theta_\ell - \theta_1$ for $\ell = 1, \ldots, n_R$ be the phase shift of the $\ell$th ring with respect to the inner ring. We also define $\rho_\ell = r_\ell / r_1$ for $\ell = 1, \cdots, n_R$ as the relative radii of the $\ell$th ring with respect to $r_1$. In particular, $\phi_1 = 0$ and $\rho_1 = 1$.

### 3.2. Constellation optimization in AWGN

We are interested in finding an APSK constellation, defined by the parameters $\boldsymbol{\rho} = (\rho_1, \ldots, \rho_{n_R})$ and $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_{n_R})$, such that a given cost function $f(\mathcal{X})$ reaches a minimum. The simplest, and probably most natural, cost function is the minimum Euclidean distance between any two points in the constellation. Section 3.2.1 shows the results under this criterion. These results are extended in Section 3.2.2, where the cost function is replaced by the mutual information of the AWGN channel; it also shown that significant gains may be achieved for low and moderate values of signal-to-noise ratio (SNR) by fine-tuning the constellation.

Copyright © 2006 John Wiley & Sons, Ltd.

DOI: 10.1002/sat

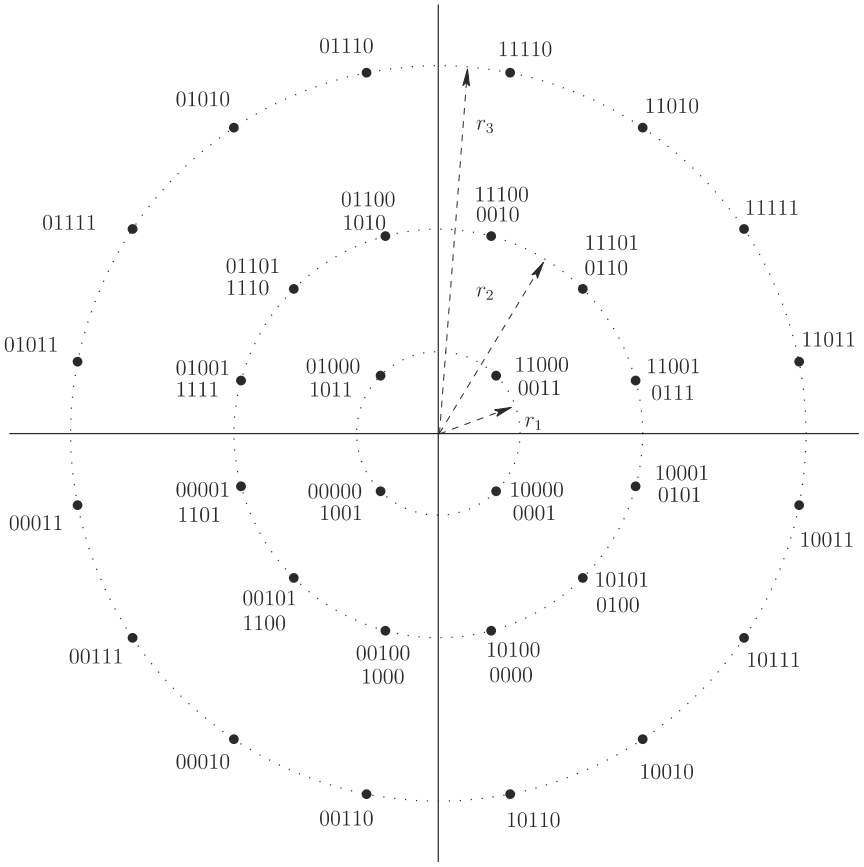



Figure 1. Parametric description and pseudo-Gray mapping of 16 and 32-APSK constellations with $n_1 = 4$, $n_2 = 12$, $\phi_2 = 0$ and $n_1 = 4$, $n_2 = 12$, $n_3 = 16$, $\phi_2 = 0$, $\phi_3 = \pi/16$. For the first two rings: mapping below corresponds to 4 + 12-APSK, mapping above to 4 + 12 + 16-APSK.

*3.2.1. Minimum Euclidean distance maximization*. The union bound on the uncoded symbol error probability [16] yields,

$$P_e \leqslant \frac{1}{M} \sum_{x \in \mathcal{X}} \sum_{\substack{x' \in \mathcal{X} \\ x' \neq x}} Q\left(\sqrt{\frac{E_s |x - x'|^2}{2N_0}}\right) \tag{5}$$

where $Q(x) = 1/\sqrt{2\pi} \int_x^\infty \mathrm{e}^{-(t^2/2)}\, \mathrm{d}t$ is the Gaussian tail function. At high SNR Equation (5) is dominated by the pairwise term at minimum squared Euclidean distance $\delta^2_{\min} = \min_{x,x' \in \mathcal{X}} |x - x'|^2$. Due to the monotonicity of the $Q$ function, it is clear that maximizing this distance optimizes the error performance estimated with the union bound at high SNR.

The minimum distance of the constellation depends on the number of rings $n_R$, the number of points in each ring $n_1, \ldots, n_{n_R}$, the radii $r_1, \ldots, r_{n_R}$, and the offset among the rings $\phi_1, \ldots, \phi_{n_R}$. The constellation geometry clearly indicates that the distances to consider are between points belonging to the same ring, or between points in adjacent rings. Simple calculations give the

Copyright © 2006 John Wiley & Sons, Ltd.

DOI: 10.1002/sat

following formula,

$$\delta^2_{\text{ring}\,i} = 2r_i^2\left[1-\cos\left(\frac{2\pi}{n_i}\right)\right] \tag{6}$$

for the distance between points in $i$th ring, with radius $r_i$ and $n_i$ points. For the adjacent rings the calculation is only slightly more complicated, and gives the following:

$$\delta^2_{\text{rings}\ i,i+1} = r_i^2 + r_{i+1}^2 - 2r_i r_{i+1}\cos\theta \tag{7}$$

where $\theta$ is the minimum relative offset between any pair of points of rings $i$ and $i+1$, respectively. As the phase of point $l_i$ in ring $i$ is given by $\phi_i + 2\pi l_i/n_i$, we easily obtain:

$$\theta = \min_{l_i,l_{i+1}}\left|(\phi_i-\phi_{i+1}) + 2\pi\left(\frac{l_i}{n_i}-\frac{l_{i+1}}{n_{i+1}}\right)\right| \tag{8}$$

The minimum distance of the constellation is given by taking the minimum of all these inter-ring and intra-ring values:

$$\delta^2_{\min} = \min_{\substack{i=1,\ldots,n_R\\ j=1,\ldots,n_R-1}} \{\delta^2_{\text{ring}\ i}, \delta^2_{\text{rings}\ j,j+1}\} \tag{9}$$

For the sake of space limitations, we concentrate on 16-ary constellations. Thanks to symmetry considerations, is clear that the best offset between rings happens when $\phi_2 = \pi/n_2$. Figure 2 shows the minimum distance for several candidates: $4+12$-, $6+10$-, $5+11$ and $1+5+10$-APSK. It may be observed that the highest minimum distance is achieved for approximately $\rho_2 = 2.0$, except for $4+12$-APSK, where $\rho_2 = 2.7$. The results for $\phi = 0$ are also plotted, and show that the corresponding minimum distance is smaller. We will see later in Section 3.2.2 how this effect translates into error rate performance.

*3.2.2. Mutual information maximization.* The mutual information (assuming equiprobable symbols) for a given signal set $\mathcal{X}$ provides the maximum transmission rate (in bits/channel use) at which error-free transmission is possible with such signal set, and is given by (e.g. Reference [13]),

$$f(\mathcal{X}) = C = \log_2 M - \mathbb{E}_{x,n}\left\{\log_2\left[\sum_{x'\in\mathcal{X}}\exp\left(-\frac{1}{N_0}\left|\sqrt{E_s}(x-x')+n\right|^2 - \left|n\right|^2\right)\right]\right\} \tag{10}$$

Interestingly, for a given SNR, or equivalently, for a given spectral efficiency $R$, an optimum constellation can be obtained, a procedure we apply in the following to 16- and 32-ary constellations.

In general, closed-form optimization of this expression is a daunting task, so we resort to numerical techniques. Expression (10) can be easily evaluated by using the Gauss–Hermite quadrature rules, making numerical evaluation very simple. Note, however, that it is possible to calculate a closed-form expression for the asymptotic case $E_s/N_0 \to +\infty$. First, note that the expectation in Equation (10) can be rewritten as

$$\lambda(\mathcal{X}) \triangleq \mathbb{E}_{x,n}\left\{\log_2\left[\sum_{x'\in\mathcal{X}}\exp\left(-\frac{1}{N_0}\left(E_s|x-x'|^2 + 2\,\text{Re}\left(\sqrt{E_s}(x-x')n\right)\right)\right)\right]\right\} \tag{11}$$

Using the dominated convergence theorem [17], the influence of the noise term vanishes asymptotically, since the limit can be pushed inside the expectation. Furthermore, the only

Copyright © 2006 John Wiley & Sons, Ltd.

DOI: 10.1002/sat

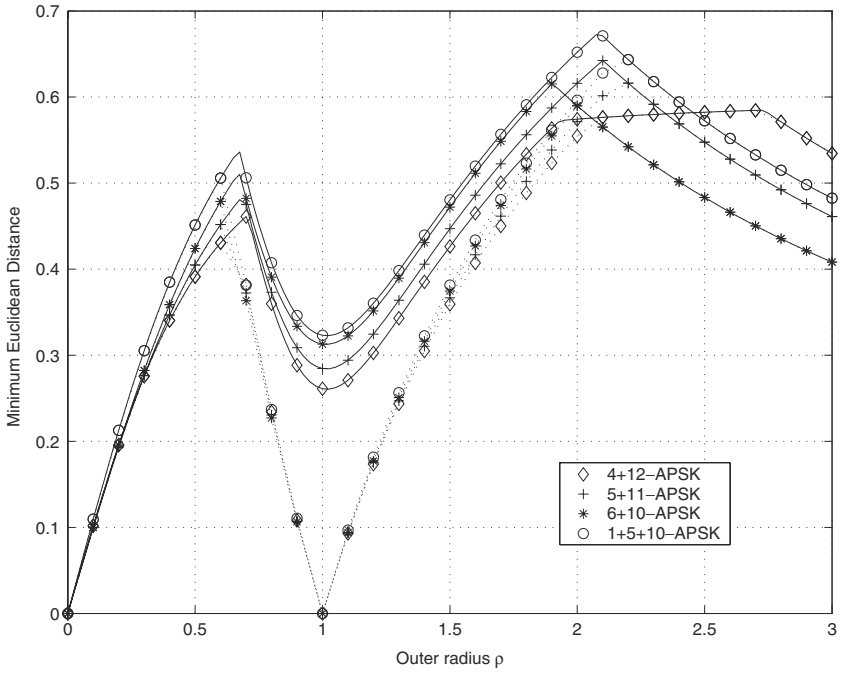



Figure 2. Minimum Euclidean distances for several 16-ary signal constellations. Solid lines correspond to $\phi = \pi/n_2$; dotted lines to $\phi = 0$.

remaining terms in the summation over $x' \in \mathcal{X}$ are $x' = x$ and those closest in Euclidean distance $\delta^2_{\min} = \min_{x' \in \mathcal{X}} |x - x'|^2$, of which there are $n_{\min}(x)$. Therefore the expectation becomes

$$\lambda(\mathcal{X}) \simeq \mathbb{E}_x \left\{ \log_2 \left[ 1 + n_{\min}(x) \exp\left( -\frac{1}{N_0} E_s \delta^2_{\min} \right) \right] \right\} \tag{12}$$

Noting that the exponential takes very small values, the approximation $\log_2(1 + x) \simeq x \log_2 e$ for $|x| \ll 1$ holds, thus by simplifying further the expectation we obtain:

$$\lambda(\mathcal{X}) \simeq \mathbb{E}_x \left\{ n_{\min}(x) \exp\left( -\frac{E_s}{N_0} \delta^2_{\min} \right) \log_2 e \right\} \simeq \alpha \exp\left( -\frac{E_s}{N_0} \delta^2_{\min} \right) \tag{13}$$

where $\alpha$ is a constant that does not depend on the constellation minimum distance $\delta_{\min}$ nor on SNR. Then the capacity at large SNR becomes:

$$f(\mathcal{X}) = \log_2 M - \alpha \exp\left( -\frac{E_s}{N_0} \delta^2_{\min} \right) \tag{14}$$

It appears then clear that the procedure corresponds to the maximization of the minimum Euclidean distance, as in Section 3.2.1.

Figure 3 shows the numerical evaluation of Equation (10) for a given range of values of $\rho_2$ and $\phi = \phi_2 - \phi_1$ for the 4 + 12-APSK constellation at $E_s/N_0 = 12$ dB. Surprisingly, there is no noticeable dependence on $\phi$. Therefore, the two-dimensional optimization can be done by simply finding the $\rho_2$ that maximizes mutual information. This result was found to hold

Copyright © 2006 John Wiley & Sons, Ltd.

DOI: 10.1002/sat

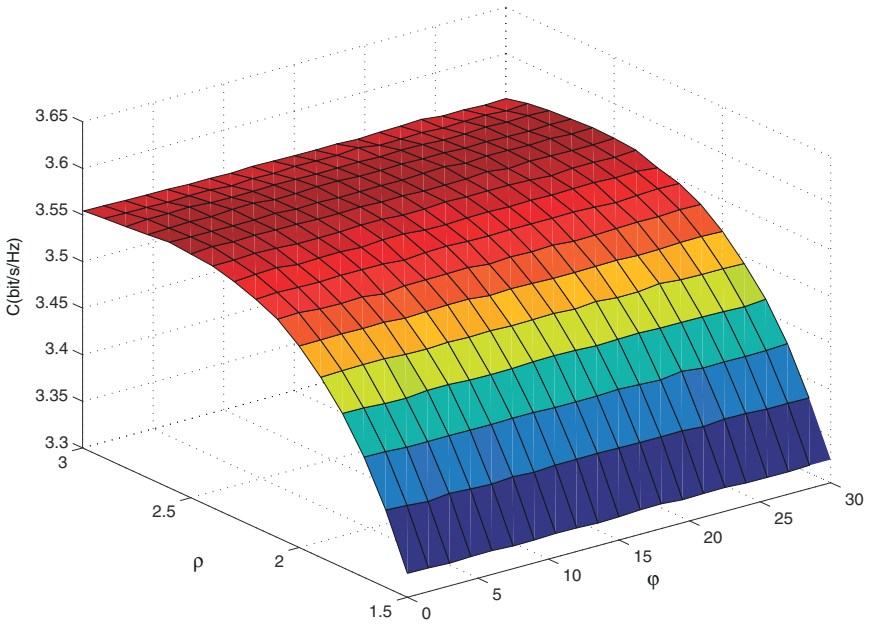



Figure 3. Capacity surface for the 16-APSK ($n_1 = 4$, $n_2 = 12$), with $E_s/N_0 = 12$ dB.

true also for the other constellations and hence, in the following, mutual information optimization results do not account for $\phi$. Figure 4 shows the union bound on the symbol error probability (5) for several 16-APSK modulations, and for the optimum value of $\rho_2$ at $R = 3$ bps/Hz (found with the mutual information analysis). Continuous lines indicate $\phi = 0$ while dotted lines refer to the maximum value of the relative phase shift, i.e. $\phi = \pi/n_2$, showing no dependence on $\phi$ at high SNR. This absence of dependency is justified by the fact that the optimum constellation separates the rings by a distance larger than the number of points in the ring itself, so that the relative phase $\phi$ has no significant impact in the distance spectrum of the constellation.

For 16-APSK it is also interesting to investigate the mutual information dependency on $n_1$ and $n_2$. Figure 5(a) depicts the mutual information curves for several configurations of optimized 16-APSK constellations and compared with classical 16-QAM and 16-PSK signal sets. As we can observe, mutual information curves are very close to each other, showing a slight advantage of $6 + 10$-APSK over the rest. In particular, note that there is a small gain, of about 0.1 dB, in using the optimized constellation for every $R$, rather than the calculated with the minimum distance (or high SNR). However, as discussed in Reference [14], $6 + 10$-APSK and $1 + 5 + 10$-APSK show other disadvantages compared to $4 + 12$-APSK for phase recovery and nonlinear channel behaviour.

Similarly, Figure 5(b) reports capacity of optimized $4 + 12 + 16$-APSK (with the corresponding optimal values of $\rho_2$ and $\rho_3$) compared to 32-QAM and 32-PSK. We observe slight capacity gain of 32-APSK over PSK and QAM constellations. Other 32-APSK constellations with different distribution of points in the three rings did not provide significantly better results.

Finally Table I provides the optimized 16- and 32-APSK parameters for various coding rates, giving an optimum constellation for each given spectral efficiency.

Copyright © 2006 John Wiley & Sons, Ltd.

DOI: 10.1002/sat

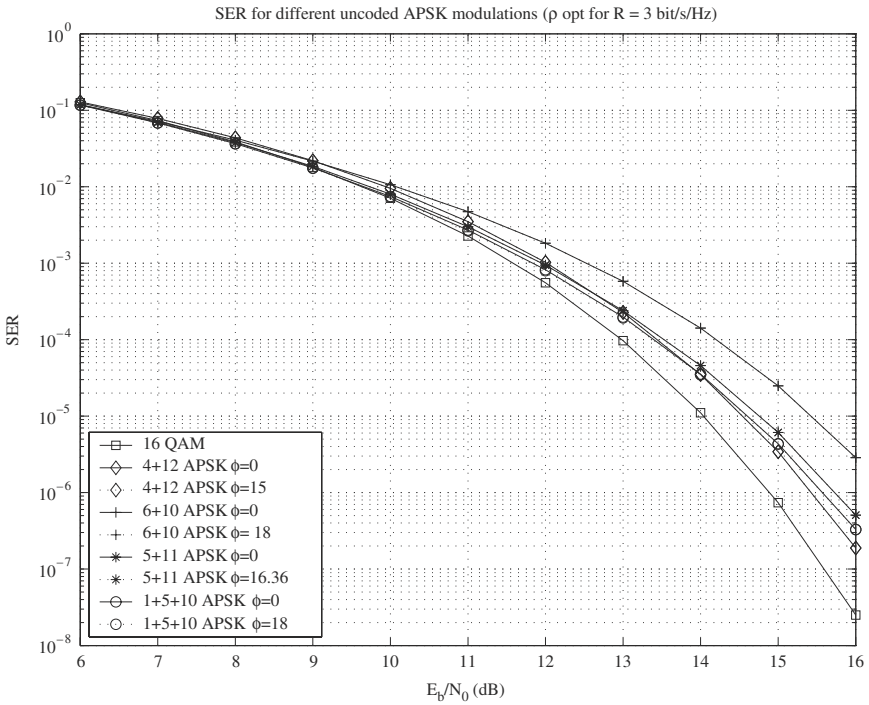



Figure 4. Union bound on the uncoded symbol error probability for several APSK modulations. Note that the continuous line and the dashed line are indistinguishable because they are superimposed.

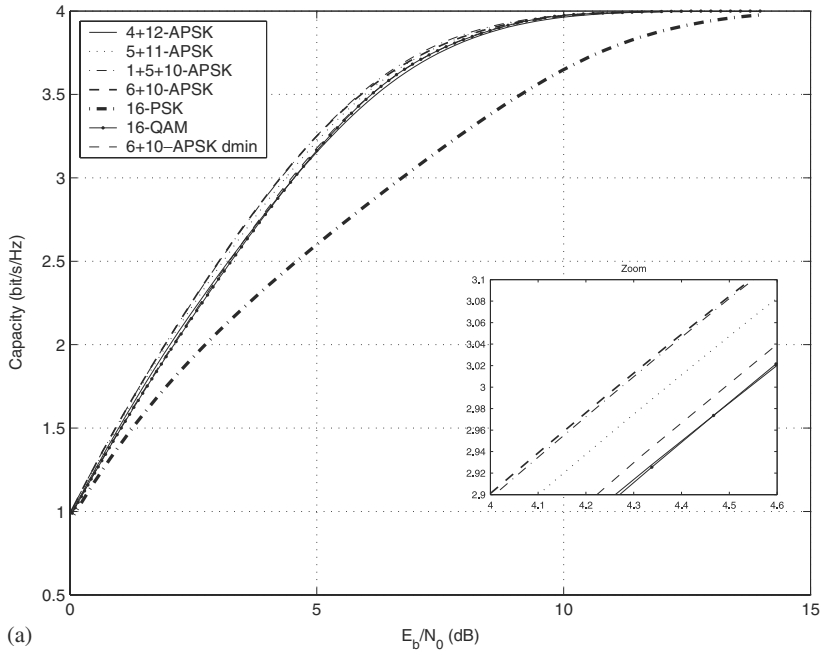
### *3.3. Constellation optimization for nonlinear channels*

*3.3.1. Peak-to-envelope considerations.* For nonlinear transmission over an amplifier, $4 + 12$-APSK is preferable to $6 + 10$-APSK because the presence of more points in the outer ring allows to maximize the HPA DC power conversion efficiency. It is better to reduce the number of inner points, as they are transmitted at a lower power, which corresponds a lower DC efficiency. It is known that the HPA power conversion efficiency is monotonic with the input power drive up to its saturation point. Figure 6 shows the distribution of the transmitted signal envelope for 16-QAM, $4 + 12$-APSK, $6 + 10$-APSK, $5 + 11$-APSK, and 16-PSK. In this case, the shaping filter is a square-root raised cosine (SRRC) with a roll-off factor $\alpha = 0.35$ as for the DVB-S2 standard [15]. As we observe, the $4 + 12$-APSK envelope is more concentrated around the outer ring amplitude than 16-QAM and $6 + 10$-PSK, being remarkably close to the 16-PSK case. This shows that the selected constellation represents a good trade-off between 16-QAM and 16-PSK, with error performance close to 16-QAM, and resilience to nonlinearity close to 16-PSK. Therefore, $4 + 12+$APSK is preferable to the rest of 16-ary modulations considered. Similar advantages have been observed for 32-APSK compared to 32-QAM.

*3.3.2. Static distortion compensation.* The simplest approach for counteracting the HPA nonlinear characteristic for the APSK signal, as already introduced in Section 2, is to modify the complex-valued constellation points at the modulator side. Thanks to the multiple-ring nature of the APSK constellation, pre-compensation is easily done by a simple modification of the

Copyright © 2006 John Wiley & Sons, Ltd.

DOI: 10.1002/sat



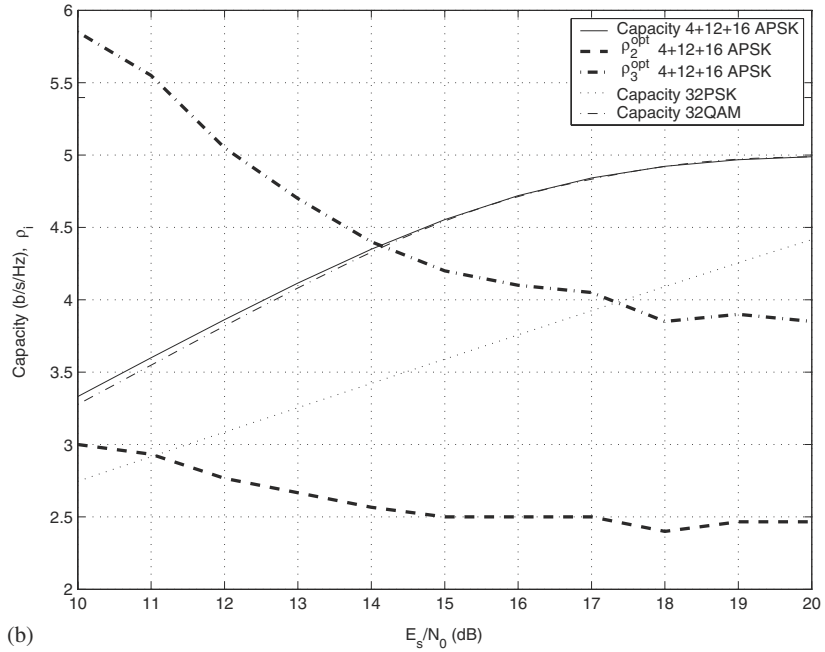




Figure 5. Capacity and $\rho^{\text{opt}}$ for the optimized APSK signal constellations vs QAM and PSK: (a) 16-ary constellations (zoom around 3 bps/Hz); and (b) 32-ary constellations.

parameters $\rho_\ell$, and $\phi_\ell$. The objective is to exploit the known AM/AM and AM/PM HPA characteristics in order to obtain a good replica of the desired signal constellation geometry after the HPA, as if it had not suffered any distortion. This can be simply obtained by artificially increasing the relative radii $\rho_\ell$ and modifying the relative phases $\phi_\ell$ at the modulator side. This

Copyright © 2006 John Wiley & Sons, Ltd.

DOI: 10.1002/sat

Table I. Optimized constellation parameters for 16- ary and 32-ary APSK.

| Modulation order | Coding rate, $r$ | Spectral eff. (bps/Hz) | $\rho_1^{opt}$ | $\rho_2^{opt}$ |
|---|---|---|---|---|
| 4 + 12-APSK | 2/3 | 2.67 | 3.15 | N/A |
| 4 + 12-APSK | 3/4 | 3.00 | 2.85 | N/A |
| 4 + 12-APSK | 4/5 | 3.20 | 2.75 | N/A |
| 4 + 12-APSK | 5/6 | 3.33 | 2.70 | N/A |
| 4 + 12-APSK | 8/9 | 3.56 | 2.60 | N/A |
| 4 + 12-APSK | 9/10 | 3.60 | 2.57 | N/A |
| 4 + 12 + 16-APSK | 3/4 | 3.75 | 2.84 | 5.27 |
| 4 + 12 + 16-APSK | 4/5 | 4.00 | 2.72 | 4.87 |
| 4 + 12 + 16-APSK | 5/6 | 4.17 | 2.64 | 4.64 |
| 4 + 12 + 16-APSK | 8/9 | 4.44 | 2.54 | 4.33 |
| 4 + 12 + 16-APSK | 9/10 | 4.50 | 2.53 | 4.30 |

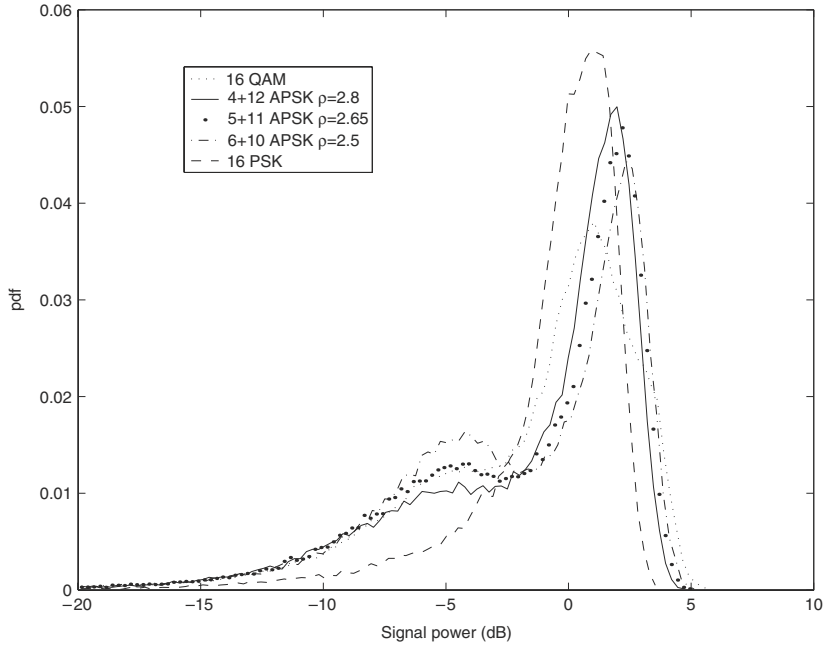



Figure 6. Simulated histogram of the transmitted signal envelope power for 16-ary constellations.

approach neglects nonlinear ISI effects at the matched filter output which are not present under the current assumption of rectangular symbols; ISI issues has been discussed in Reference [14].

In the 16-ary APSK case, the new constellation points $x'$ follow (4), with new radii $r'_1$, $r'_2$, such that $F(r'_1) = r_1$, and $F(r'_2) = r_2$. Concerning the phase, it is possible to pre-correct for the relative phase offset introduced by the HPA between inner and outer ring by simply changing the relative phase shift by $\phi'_2 = \phi_2 + \Delta\phi$, with $\Delta\phi = \phi(r'_2) - \phi(r'_1)$. These operations can be readily implemented in the digital modulator by simply modifying the reference constellation parameters $\rho'$, $\phi'$, *with no hardware complexity impact* or out-of-band emission increase at the

Copyright © 2006 John Wiley & Sons, Ltd.

DOI: 10.1002/sat

linear modulator output. On the other side, this allows to shift all the compensation effort into the modulator side allowing the use of an optimal demodulator/decoder for AWGN channels even when the amplifier is close to saturation. The pre-compensated signal expression at the modulator output is then

$$s_T^{\text{pre}} = \sqrt{P}\sum_{k=0}^{L-1} x'(k)p_T(t - kT_s) \tag{15}$$

where now $x'(k) \in \mathscr{X}'$ are the pre-distorted symbols with $r'_\ell$ and $\phi'_\ell$ for $\ell = 1, \ldots, n_R$.

## 4. FORWARD ERROR CORRECTION CODE DESIGN AND PERFORMANCE

In this section, we describe the coupling of turbo-codes and the APSK signal constellations through BICM and we discuss some of the properties of this approach.[§] As already mentioned in Section 1, such approach is a good candidate for flexible constellation format transmission. The main drivers for the selection of the FEC code have been flexibility, i.e. use a single mother code, independently of the modulation and code rates; complexity, i.e. have a code as compact and simple as possible; and good performance, i.e. approach Shannon's capacity bound as much as possible.

We consider throughout a coded modulation scheme for which the transmitted symbols $\mathbf{x} = (x_0, \ldots, x_{L-1})$ are obtained as follows: (1) The information bits sequence $\mathbf{a} = (a_0, \ldots, a_{K-1})$ is encoded with a binary code $\mathscr{C} \in \mathbb{F}_2^N$ of rate $r = K/N$; (2) the encoded sequence $\mathbf{c} = (c_0, \ldots, c_{N-1}) \in \mathscr{C}$ is bit-interleaved, with an index permutation $\boldsymbol{\pi} = (\pi_0, \ldots, \pi_{N-1})$; (3) the bit-interleaved sequence $\mathbf{c}_\pi$ is mapped to a sequence of modulation symbols $\mathbf{x}$ with a labelling rule $\mu : \mathbb{F}_2^M \to \mathscr{X}$, such that $\mu(a_1, \ldots, a_M) = x$. In addition to the description of the code, we also propose the use of some new heuristics to tune the final design of the BICM codes.

### 4.1. Code design

It was suggested in Reference [13] that the binary code $\mathscr{C}$ can be optimized for a binary channel (such as BPSK or QPSK with AWGN). We substantiate this claim with some further insights on the effect of the code minimum distance in the error performance. The Bhattachharyya union bound (BUB) on the frame error probability $P_e$ for a BICM modulation assuming that no iterations are performed at the demapper side is given by [13]:

$$P_e \leqslant \sum_d A(d)B(E_s/N_0)^d \tag{16}$$

where $A(d)$ is the number of codewords at a Hamming distance $d$, $d_{\min}$ is the minimum Hamming distance, with $B(E_s/N_0)$ denoting the Bhattachharyya factor, which is given by

$$B(E_s/N_0) = \frac{1}{M\log_2 M}\sum_{i=1}^{\log_2 M}\sum_{b=0}^{1}\sum_{x\in\mathscr{X}_{i=b}} \mathbb{E}_n\left\{\sqrt{\frac{\sum_{z\in\mathscr{X}_{i=\bar{b}}}\exp(-(1/N_0)|x-z+n|^2)}{\sum_{z\in\mathscr{X}_{i=b}}\exp(-(1/N_0)|x-z+n|^2)}}\right\} \tag{17}$$

[§]The optimization method based on the mutual information proposed in Section 3.2.2 can be easily extended to the case of the BICM mutual information [13] with almost identical results assuming Gray mapping. However, we use the proposed method in order to keep the discussion general and not dependent on the selected coding scheme.

Copyright © 2006 John Wiley & Sons, Ltd.

DOI: 10.1002/sat

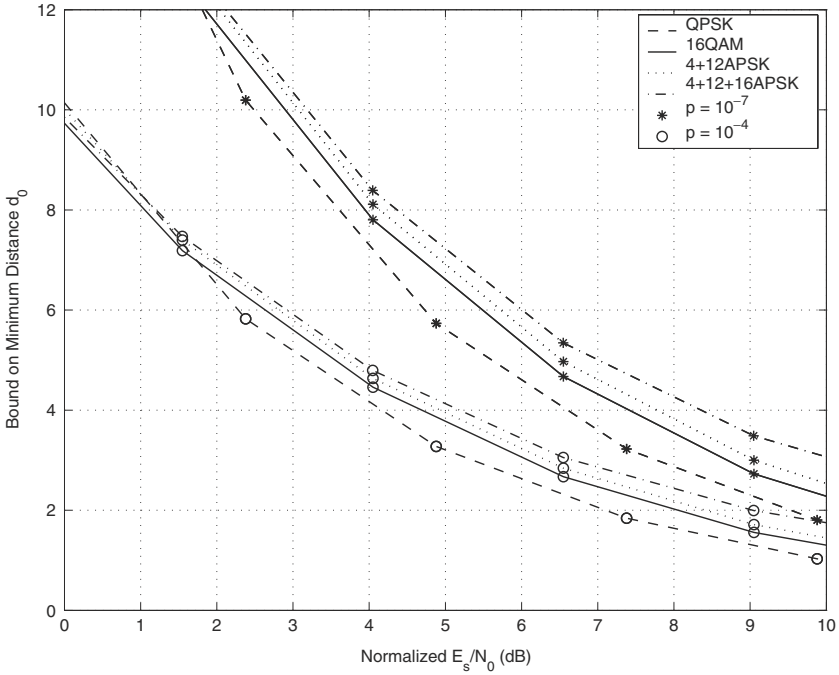



Figure 7. Lower bound $d_0$ vs normalized $E_s/N_0$ for target $P_e = 10^{-4}, 10^{-7}$, QPSK and 16- and 32-ary modulations and Gray labelling.

where $\mathcal{X}_{i=b} = \{x \in \mathcal{X} | \mu_i^{-1}(x) = b\}$, where $\mu_i^{-1}(x) = b$ denotes that the $i$th position of binary label $x$ is equal to $b$. Equation (17) can be evaluated very efficiently using the Gauss–Hermite quadrature rules. For sufficiently large $E_s/N_0$ the BUB in Equation (5) is dominated by the term at minimum distance, i.e. the error floor

$$P_e \simeq A_{d_{\min}} B(E_s/N_0)^{d_{\min}} \tag{18}$$

From this equation, we can derive an easy lower bound on the $d_0$ on the minimum distance of $\mathscr{C}$ for a given target error rate, modulation, and number of codewords at $d_{\min}$:

$$d_{\min} \geqslant \lceil d_0 \rceil \quad \text{where } d_0 = \frac{\log P_e - \log A_{d_{\min}}}{\log B} \tag{19}$$

where $\lceil x \rceil$ denotes the smallest integer greater or equal to $x$. Notice that the target error rate is fixed to be the error floor under ML decoding.¶ The lowest error probability floor is achieved by a code $\mathscr{C}$ with $A_{d_{\min}} = 1$. Figure 7 shows the lower bound $d_0$ with $A_{d_{\min}} = 1$, as a function of $E_s/N_0$ for target $P_e = 10^{-4}, 10^{-7}$, QPSK, 16-QAM, 16- and 32-APSK modulations and Gray labelling. In order to ease the comparison, a normalized SNR is used, defined as

$$\left.\frac{E_s}{N_0}\right|^{\text{norm}} = \frac{E_s}{N_0}\frac{1}{2^R - 1} \tag{20}$$

¶Although this does not necessarily hold under iterative decoding, it does still provide a useful guideline into the performance.

Copyright © 2006 John Wiley & Sons, Ltd.

DOI: 10.1002/sat

where $R$ is the spectral efficiency, and the normalization is thus to the channel capacity. The code rate has been taken $r = 3/4$ for all cases. Note that a capacity-achieving pair modulation-code would work at a normalized $E_s/N_0|^{\text{norm}} = 1$ or 0 dB.

A remarkable conclusion is that BICM with Gray mapping preserves the properties of $\mathscr{C}$ regardless of the modulation used, since we observe that the requirements for nonbinary modulations are similar to those for binary modulations (in the error-floor region). In order to work at about 3 dB from capacity, that is, a normalized $E_s/N_0|^{\text{norm}} = 3$ dB, the needed $d_0$ is about 5 and 10 for a frame error rate of $10^{-4}$ and $10^{-7}$, respectively.

We consider that $\mathscr{C}$ is a serial concatenatation of convolutional codes (SCCC) [18], with outer code $\mathscr{C}^{\text{O}}$ of length $L_{\text{O}}$ and rate $r_{\text{O}}$ and inner code $\mathscr{C}^{\text{I}}$ of length $L_{\text{I}}$ and rate $r_{\text{I}}$. Obviously, $L_{\text{I}} = N$ and $r_{\text{O}} r_{\text{I}} = r$. The resulting spectral efficiency is $R = r \log_2 M$. It provides two key advantages with respect to parallel turbo codes: lower error floor, possibly achieving the bit error rate requirements (BER $\leqslant 10^{-10}$) without any external code; and simpler constituent codes simpler than in turbo codes or in classical concatenated codes. In addition, with an SCCC the outer code is fully integrated into the decoding process, which includes iterations between decoding stages for the inner and outer codes. This avoids the need to use an additional external code, such as a Reed–Solomon (and its associated interleaver). In some sense, the outer code is already included in the SCCC code, thus saving one extra encoding/decoding step, and one memory level, therefore reducing the required complexity.

The best choice in terms of low error floor forces the lowest possible rate for the outer encoder, as this maximizes the interleaver gain, which increases exponentially with the outer code free distance [18]. We should then set the outer code rate equal to the total code rate, and the inner code rate to 1. Also, it turns out that the best choice for the inner encoder is the two-state differential encoder also known as 'accumulator'. It meets the requirements of simplicity, it is 'almost' systematic, in the sense that the dependency among the bits in its output sequence is very mild, and moreover, it is recursive as imposed by the design rules of SCCCs for the inner encoder. Last but not least, this choice leads to a very simple inner SISO, a highly desirable feature for a design working at high data rates.

In practice the maximum block length to be used shall be selected accounting for the maximum allowed end-to-end latency and decoder complexity. One recent finding [19] allows to split an arbitrary block interleaver in an arbitrary number of smaller nonoverlapping interleavers. This allows to greatly reduce the decoder complexity when parallel SISO units are used to achieve high-speed decoding as memory requirements does not increase with the degree of parallelism.

As an outer code, we have selected the standard binary 16-state convolutional code, rate 3/4 [20]. Its free distance is 4, large enough so that interleaving gain can be achieved, and the minimum distance of the concatenated code grows towards infinity with the blocklength [21]. Furthermore, and if required, the code may be punctured to higher rates [22], with no loss in the code distance. Further numerical results are presented in Section 4.3.

### *4.2. Demodulation*

Decoding of BICM consists of a concatenation of two steps, namely maximum-*a posteriori* (MAP) soft-input soft-output (SISO) demapper (symbol-to-bit likelihood computer), and a MAP SISO decoder of $\mathscr{C}$. These two steps exchange extrinsic information messages $m_{\mu \to \mathscr{C}}$ (from

Copyright © 2006 John Wiley & Sons, Ltd.

DOI: 10.1002/sat

the demapper to the SISO decoder of $\mathscr{C}$), and $m_{\mathscr{C}\to\mu}$ (from the SISO decoder of $\mathscr{C}$ to the demapper) through the iterations. Extrinsic information messages $m$ (or metrics) can be in the form of likelihood probabilities, log-likelihood ratios or some combination or approximation of them. When either $\mathscr{C}^{\mathrm{O}}$ or $\mathscr{C}^{\mathrm{I}}$, or both, are convolutional codes, MAP SISO decoding is efficiently computed by the BCJR algorithm [23]. For example, the extrinsic log-likelihood ratio corresponding to $m_{\mu\to\mathscr{C}}$ for the $i$th coded bit of the $k$th symbol and $l$th iteration is given by

$$\Lambda(c_{k,i}^{(l)}) = \log \frac{\sum_{x\in\mathscr{X}_{i=0}} p(y_k|x) \prod_{j\neq i} P_{\mathscr{C}\to\mu}^{(l-1)}(c_{k,j})}{\sum_{x\in\mathscr{X}_{i=1}} p(y_k|x) \prod_{j\neq i} P_{\mathscr{C}\to\mu}^{(l-1)}(c_{k,j})} \tag{21}$$

where $p(y_k|x) \propto \exp(-(1/N_0)|y_k - \sqrt{E_s}x|^2)$, $P_{\mathscr{C}\to\mu}^{(l)}(c)$ denotes the extrinsic probability message corresponding to $m_{\mathscr{C}\to\mu}$ on the coded bit $c$ at the $l$th iteration.

There is a marginal information loss in considering no iterations at the demodulator side when Gray mapping is used for transmitting high rates [13], i.e. $P_{\mathscr{C}\to\mu}^{(l-1)}(c_{k,j}) = 0.5$ implies almost no loss in spectral efficiency using Gray mapping. When demapper iterations are allowed, Gray mapping is known not to gain through the iterations [24]. Moreover, when other mapping rules are used, scheduling the operations for such decoder (a SCC with BICM) can be a very complicated task and has been solved only for $N \to \infty$ (see e.g. Reference [25] for recent results on the subject). For all these aforementioned reasons, we will assume Gray mapping and that information flows from demodulator to decoder only, with no feed-back.

### 4.3. Performance analysis

Density evolution (or approximations such as EXIT charts [24]) of such turbo-coded BICM is a very complicated task due to the concatenation of three elements exchanging extrinsic information messages through the iterations. Such techniques lead in general to three-dimensional surfaces which are difficult to deal with in practical decoding algorithms for finite length codes [25]. We will therefore resort to a mixture of computer simulations and bounds on maximum likelihood (ML) decoding error probability. Regarding convergence, simulations can accurately estimate the values of $E_b/N_0$ at which the decoding algorithm does not converge, as will be shown shortly.

We denote the binary-input channel between the modulator and demodulator as the equivalent binary-input BICM channel. It has been recently shown [26] that such channel can be very well approximated as AWGN[‖] with SNR $\gamma = -\log B(E_s/N_0)$. Therefore, standard bounds for binary-input channels can be successfully applied here. In particular, the standard union bound (UB) yields

$$P_e \lessapprox \sum_d A_d Q(\sqrt{-2d \log B(E_s/N_0)}) \tag{22}$$

At high SNR, (16) and (22) are dominated by the pairwise terms corresponding to the few codewords with low Hamming distance (error floor). When turbo-like codes are used, union bounding techniques are known not to provide good estimates of the error probability,

[‖] Notice that the Gaussian approximation (GA) is common practice in density evolution techniques [24].

Copyright © 2006 John Wiley & Sons, Ltd.

DOI: 10.1002/sat

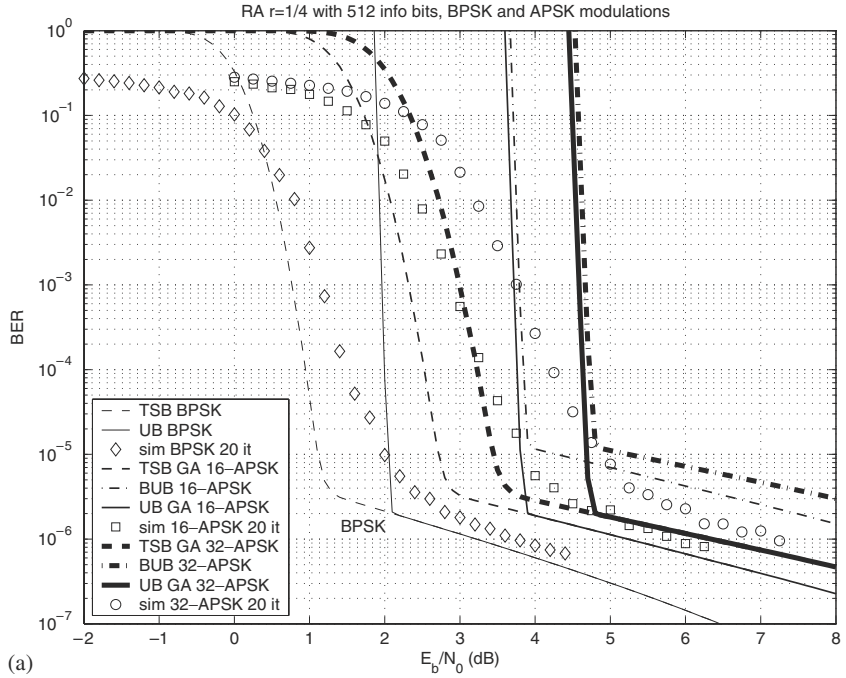



Figure 8. Bit-error probability bounds and simulations for BPSK and 16-APSK ($n_1 = 4$ and $n_2 = 12$) and 32-APSK ($n_1 = 4$, $n_2 = 12$ and $n_3 = 16$) with pseudo-Gray labelling: (a) RA code with $r = 1/4$ and $K = 512$ information bits per frame; and (b) SCCC with rate 3/4 16 states convolutional code as outer code and inner accumulator with $N = 5000$.

Copyright © 2006 John Wiley & Sons, Ltd.

DOI: 10.1002/sat

and one typically resorts to improved bounds such as the tangential-sphere bound (TSB) [27],

$$P_F \lessapprox \int_{-\infty}^{+\infty} \frac{dz_1}{\sqrt{2\pi\sigma^2}} e^{-z_1^2/2\sigma^2} \left\{ 1 - \overline{\Gamma}\left(\frac{N-1}{2}, \frac{r_{z_1}}{2\sigma^2}\right) + \sum_{d:\delta/2<\alpha_\delta} A_d \overline{\Gamma}\left(\frac{N-2}{2}, \frac{r_{z_1}^2 - \beta_\delta(z_1)^2}{2\sigma^2}\right) \left[ Q\left(\frac{\beta_\delta(z_1)}{\sigma}\right) - Q\left(\frac{r_{z_1}}{\sigma}\right) \right] \right\} \quad (23)$$

where $\overline{\Gamma}(a,x) = (1/\Gamma(a)) \int_0^x t^{a-1} e^{-t}\, dt$ is the normalized incomplete gamma function and $\Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t}\, dt$ is the gamma function,

$$\sigma^2 = (-2 \log B(E_s/N_0))^{-1} \quad (24)$$

$$r_{z_1} = r(1 - z_1/4R) \quad (25)$$

$$\beta_\delta(z_1) = (r_{z_1}/\sqrt{1 - \delta^2/4R^2})(\delta/2r) \quad (26)$$

$$\alpha_\delta = r\sqrt{1 - \delta^2/4R^2} \quad (27)$$

$R^2 = N$, $\delta = 2\sqrt{d}$ and $r$, the cone radius, is the solution of

$$\sum_{d:\delta/2<\alpha\delta} A_d \int_0^{\theta_k} \sin^{N-3} \phi\, d\phi = \sqrt{\pi}\Gamma((N-2)/2)/\Gamma((N-1)/2) \quad (28)$$

with

$$\theta_k = \cos^{-1}((\delta/2r)(1/\sqrt{1 - \delta^2/4R^2})) \quad (29)$$

Figure 8 shows the bit-error probability bounds using the Gaussian approximation, Equation (22) and simulations for BPSK, 4 + 12-APSK and 4 + 12 + 16-APSK with the pseudo-Gray labellings of Figure 1. In particular, in Figure 8(a) we use a repeat and accumulate (RA) code [28] with $r = 1/4$ and $K = 512$ information bits per frame. In this case, the weight enumerator can be computed in closed form [28]. We observe that as expected BICM preserves the properties of the underlying binary code $\mathscr{C}$, since both waterfall and floor occur at almost the same probability values. We also observe that the approximations in Equations (22) and (23) are very accurate and yield much better error probability estimates than the standard Bhattacharyya bound. Same conclusions apply to Figure 8(b) where we use a SCCC with the optimal 16 states $r = 3/4$ convolutional code as outer code and inner accumulator, with interleaver size $N = 5000$. Storage limitations prevent us from showing the curves for larger $N$. However, we observe that already with $N = 5000$, we have very low error floors. In particular, in the DVB-S2 application, the interleaver size is set to 16 200 or 64 000, which implies that almost-error-free transmission is possible with such code.

Copyright © 2006 John Wiley & Sons, Ltd.

DOI: 10.1002/sat

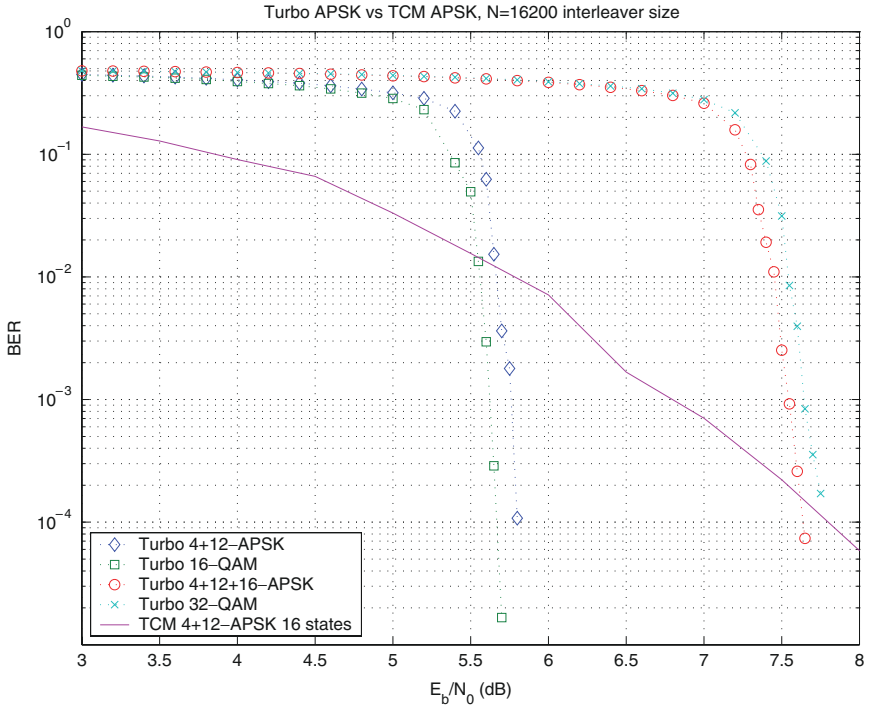



Figure 9. Turbo-coded APSK with $N = 16\,200$ vs $4 + 12$-TCM.

Finally, Figure 9, shows the simulated BER performance for the same SCCC with the optimal 16 states $r = 3/4$ convolutional code as outer code and inner accumulator, with interleaver size $N = 16\,200$[**] with $4 + 12$-APSK, 16-QAM, $4 + 12 + 16$-APSK and 32-QAM. For the sake of comparison, we also plot the BER for a $4 + 12$-APSK with a 16 states TCM, typical of satellite systems current standard [1]. As we observe, the SCCC codes yield a substantial performance improvement with respect to TCM. In the TCM case, one usually concatenates a Reed–Solomon code operating at an input BER of $10^{-4}$, which usually diminishes the spectral efficiency and increases the receiver complexity. Notice as well that 32-APSK achieves a better performance than 32-QAM, giving a further justification to the use of modulations in the APSK family instead of the classical QAM.

## 5. CONCLUSIONS

Extensive analysis and simulations for turbo-coded APSK modulations, with particular emphasis on its applicability to satellite broadband communications have been presented in this paper. In particular, we have investigated APSK constellation optimization under mutual information and minimum Euclidean distance criteria, under the simplified assumption of

[**] The selected FEC block size ensures that the FEC floor is well below the required BER of $10^{-10}$ for satellite broadcasting systems.

Copyright © 2006 John Wiley & Sons, Ltd.

DOI: 10.1002/sat

rectangularly shaped transmission pulses. We have shown that the degrees of freedom in the design of an APSK modulation can be exploited thanks to the mutual information maximization, and this has been applied to the design of 16- and 32-ary constellations. This technique has been shown to extend the standard minimum Euclidean distance maximization, yielding a small but significant improvement.

The pragmatic approach of BICM allows for a good coupling between such optimized APSK modulations with powerful binary turbo-codes, due to its inherent flexibility for multiple-rate transmission. Some new heuristics have been used to further justify the design of a single mother code to be used for all rates. A theoretical explanation of the the fact that the error floor typical of turbo codes remains at a constant distance from capacity has been presented. We have presented some new ML decoding error probability bounds for BICM APSK, and we have compared them with simulations findings. Numerical results based on simulation of bit-error rate probability for high rate transmission with turbo-coded APSK have been presented, showing large advantage of the presented scheme over standard TCM.

## REFERENCES


1. Cominetti M, Morello A. Digital video broadcasting over satellite (DVB-S): a system for broadcasting and contribution applications. *International Journal of Satellite Communications* 2000; **18**(6):393–410.
2. De Gaudenzi R, Rinaldo R. Adaptive coding and modulation for next generation broadband multimedia systems. *Proceedings of the 20th AIAA Satellite Communication Systems Conference*, Montreal, May 2002.
3. De Gaudenzi R, Guillén i Fàbregas A, Martinez A. Performance analysis of turbo-coded APSK modulations over nonlinear satellite channels. *IEEE Transactions on Wireless Communications* 2005, in press.
4. Biglieri E. High-level modulation and coding for nonlinear satellite channels. *IEEE Transactions on Communications* 1984; **COM-32**(5):616–626.
5. Thomas CM, Weidner MY, Durrani SH. Digital amplitude-phase keying with *m*-ary alphabets. *IEEE Transactions on Communications* 1974; **COM-22**(2):168–180.
6. Foschini G, Gitlin R, Weinstein S. Optimization of two-dimensional signal constellations in the presence of Gaussian noise. *IEEE Transactions on Communications* 1974; **22**(1):28–38.
7. Einarsson G. Signal design for the amplitude-limited Gaussian channel by error bound optimization. *IEEE Transactions on Communications* 1979: **27**(1):152–158.
8. Shamai S, Bar-David I. The capacity of average and peak-power-limited quadrature Gaussian channels. *IEEE Transactions on Information Theory* 1995; **41**:1060–1071.
9. Müller R, Wachsmann U, Huber J. Multilevel coding for peak power limited complex Gaussian channels. *Proceedings of the IEEE International Symposium on Information Theory* (*ISIT'97*), June–July 1997; 103.
10. Berrou C, Glavieux A, Thitimajshima P. Near Shannon-limit error correcting coding and decoding: turbo codes. *Proceedings of the IEEE International Conference on Communications* (*ICC'93*), Geneva, Switzerland, May 1993.
11. Gallager R. Low-density parity-check codes. *Ph.D. Dissertation*, Massachussets Institute of Technology, 1963.
12. Ungerboeck G. Channel coding with multilevel phase signals. *IEEE Transactions on Information Theory* 1982; **IT-28**(1):55–67.
13. Robertson P, Worz T. Bandwidth-efficient turbo trellis-coded modulation using punctured component codes. *IEEE Journal on Selected Areas in Communications* 1998; **16**(2):206–218.
14. Caire G, Taricco G, Biglieri E. Bit-interleaved coded modulation. *IEEE Transactions on Information Theory* 1998; **44**:927–947.
15. ETSI Digital Video Broadcasting (DVB). Second generation framing structure, channel coding and modulation systems for broadcasting, interactive services, news gathering and other broadband satellite applications. European Telecommunications Standards Institute Std. EN 302 307, 2003.
16. Benedetto S, Biglieri E, Castellani V. *Digital Transmission Theory*. Prentice-Hall: Englewood Cliffs, NJ, 1987.
17. Dembo A, Zeitouni O. *Large Deviations Techniques and Applications* (2nd edn). Applications of Mathematics. Springer: Berlin, 1998; (38).
18. Benedetto S, Divsalar D, Montorsi G, Pollara F. Serial concatenated of interleaved codes: performance analysis, design and iterative decoding. *IEEE Transactions on Information Theory* 1998; **44**(3):909–926.
19. Tarable A, Montorsi G, Benedetto S. Mapping interleaver laws to parallel turbo decoder architectures. *Proceedings of the 3rd International Symposium on Turbo Codes and Applications*, Brest, France, September 2003.


Copyright © 2006 John Wiley & Sons, Ltd.

DOI: 10.1002/sat

20. Chang J-J, Hwang D-J, Lin M-C. Some extended results on the search for good convolutional codes. *IEEE Transactions on Information Theory* 1997; **43**(5):1682–1697.
21. Kahale N, Urbanke R. On the minimum distance of parallel and serially concatenated codes. *IEEE Transactions on Information Theory* 1997, in press.
22. Martinez A, Ponticelli B. Efficient coding schemes for high order modulations in satellite broadcasting applications. *Proceedings of the 7th International ESA Workshop on Digital Signal Processing Techniques for Space Communication*, Sesimbra, Portugal, October 2001.
23. Bahl L, Cocke J, Jelinek F, Raviv J. Optimal decoding of linear codes for minimizing the symbol error rate. *IEEE Transactions on Information Theory* 1974; **20**:284–287.
24. ten Brink S. Designing iterative decoding schemes with the extrinsic information transfer chart. *AEÜ—International Journal of Electronics and Communications* 2000; **54**(6):389–398.
25. Brannstrom F, Rasmussen LK, Grant A. Optimal scheduling for multiple serially concatenated systems. *Proceedings of the 3rd International Symposium on Turbo Codes and Applications*, Brest, France, September 2003.
26. Guillén i Fàbregas A, Martinez A, Caire G. Error probability of bit-interleaved coded modulation using the Gaussian approximation. *Proceedings of the 38th Annual Conference on Information Sciences and Systems* (*CISS*) *2004*, Princeton, U.S., March 2004.
27. Sason I, Shamai S. Improved upper bounds on the ML decoding error probability of parallel and serial concatenated turbo codes via their ensemble distance spectrum. *IEEE Transactions on Information Theory* 2000; **46**(1):24–47.
28. Divsalar D, Jin H, McEliece RJ. Coding theorems for 'turbo-like' codes. *Proceedings of the 36th Allerton Conference on Communication*, *Control and Computing*, Allerton, IL, U.S., September 1998; 210–219.

## AUTHORS' BIOGRAPHIES



**Riccardo De Gaudenzi** was born in Italy in 1960. He received his Doctor Engineer degree (cum Laude) in electronic engineering from the University of Pisa, Italy in 1985 and the PhD from the Technical University of Delft, The Netherlands in 1999. From 1986 to 1988 he was with the European Space Agency (ESA), Stations and Communications Engineering Department, Darmstadt (Germany) where he was involved in satellite telecommunication ground systems design and testing. In particular, he followed the development of two new ESA's satellite tracking systems. In 1988, he joined ESA's Research and Technology Centre (ESTEC), Noordwijk, The Netherlands where in 2000 he has been appointed head of the Communication Systems Section and since 2005 he is Head of the RF Payload and Systems Division. The division is responsible for the definition and development of advanced satellite system, subsystems and technologies for telecommunications, navigation and earth observation applications. In 1996 he spent 1 year with Qualcomm Inc., San Diego, U.S.A., in the Globalstar LEO project system group under an ESA fellowship. His current interest is mainly related with efficient digital modulation and access techniques for fixed and mobile satellite services, synchronization topics, adaptive interference mitigation techniques and communication systems simulation techniques. From 2001 to 2005 he has been serving as Associate Editor for CDMA and Synchronization for IEEE Transactions on Communications. He is co-recipient of the VTS Jack Neubauer Best System Paper Award from the IEEE Vehicular Technology Society.



**Albert Guillén i Fàbregas** was born in Barcelona, Catalunya, Spain, in 1974. He received the Telecommunications Engineering Degree and the Electronics Engineering Degree from Universitat Politècnica de Catalunya, Barcelona, Catalunya, Spain, and the Politecnico di Torino, Torino, Italy, respectively, both in 1999, under the Double Degree Socrates-Erasmus Scholarship Program of the European Community, and the PhD in Communication Systems from Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland, in 2004.

From August 1998 to March 1999, he conducted his Final Research Project at the Center for Communications and Signal Processing Research (CCSPR), at the New Jersey Institute of Technology (NJIT), Newark, NJ, supported with a Nokia-NJIT

Copyright © 2006 John Wiley & Sons, Ltd.

DOI: 10.1002/sat