International Telecommunication Union

ITU-R

Radiocommunication Sector of ITU

Report ITU-R S.2173-1

(07/2014)

Multi-carrier based transmission techniques for satellite systems

S Series

Fixed satellite service

Foreword

The role of the Radiocommunication Sector is to ensure the rational, equitable, efficient and economical use of the radio-frequency spectrum by all radiocommunication services, including satellite services, and carry out studies without limit of frequency range on the basis of which Recommendations are adopted.

The regulatory and policy functions of the Radiocommunication Sector are performed by World and Regional Radiocommunication Conferences and Radiocommunication Assemblies supported by Study Groups.

Policy on Intellectual Property Right (IPR)

ITU-R policy on IPR is described in the Common Patent Policy for ITU-T/ITU-R/ISO/IEC referenced in Annex 1 of Resolution ITU-R 1. Forms to be used for the submission of patent statements and licensing declarations by patent holders are available from http://www.itu.int/ITU-R/go/patents/en where the Guidelines for Implementation of the Common Patent Policy for ITU-T/ITU-R/ISO/IEC and the ITU-R patent information database can also be found.

Series of ITU-R Reports

(Also available online at http://www.itu.int/publ/R-REP/en)

Series	Title
BO	Satellite delivery
BR	Recording for production, archival and play-out; film for television
BS	Broadcasting service (sound)
BT	Broadcasting service (television)
F	Fixed service
M	Mobile, radiodetermination, amateur and related satellite services
P	Radiowave propagation
RA	Radio astronomy
RS	Remote sensing systems
S	**Fixed-satellite service**
SA	Space applications and meteorology
SF	Frequency sharing and coordination between fixed-satellite and fixed service systems
SM	Spectrum management

***Note**: This ITU-R Report was approved in English by the Study Group under the procedure detailed in Resolution ITU-R 1.*

Electronic Publication
Geneva, 2014

REPORT ITU-R S.2173-1

Multi-carrier based transmission techniques for satellite systems

(Questions ITU-R 46-3/4 and ITU-R 73-2/4)

(2010, 2014)

TABLE OF CONTENTS

Page

Abbreviations

3DTV	Three-dimensional television
ACK	Acknowledgment message
ACM	Adaptive coding and modulation
AM/AM	Amplitude-to-amplitude
AM/PM	Amplitude-to-phase
APSK	Amplitude and phase shift keying
ARQ	Automatic repeat request
ATC	Ancillary terrestrial component
AVC	Advanced video coding
AWGN	Additive white Gaussian noise
BCH	Bose-Chaudhuri-Hocquenghem
BER	Bit error ratio
B-GAN	Broadband global area network
BLER	Block error rate
BPA	Belief propagation algorithm
BPS	Bent-pipe satellite
BPSK	Binary phase shift keying
BSM	Broadband satellite multimedia
BSS	Broadcasting-satellite service
BTC	Block turbo code
CCDF	Complementary cumulative distribution function
CCM	Constant coding and modulation
CDM	Code-division multiplexing
CGC	Complementary ground component
CI-OFDM	Carrier interferometry orthogonal frequency-division multiplexing
CN	Core network
CNR	Carrier-to-noise ratio
COFDM	Coded orthogonal frequency-division multiplexing
CP	Cyclic prefix
CPA	Chase-Pyndiah algorithm
CRC	Cyclic redundancy check
CTC	Convolutional turbo codes
DAB	Digital audio broadcasting
DBS	Direct broadcasting satellite

D/C	Down-converter
DLP	Digital light processing
DSDT	Direct spectrum division transmission
DSL	Digital subscriber line
DTH	Direct-to-home
DVB–H	Digital video broadcasting-handheld
DVB–RCS	Digital video broadcasting-return channel via satellite
DVB–S	Digital video broadcasting-satellite
DVB–S2	Digital video broadcasting-satellite-second generation
DVB–SH	Digital video broadcasting-satellite services to handheld
DVB-T	Digital video broadcasting-terrestrial
DVB-T2	Digital video broadcasting-terrestrial-second generation
E_b/N_0	Bit energy to noise spectral density ratio
E_s/N_0	Symbol energy to noise spectral density ratio
E-S	Earth-to-space
ESPN	Entertainment and sports programming network
ETRI	Electronics and telecommunications research institute
ETSI	European telecommunications standards institute
FCC	Federal Communications Commission
FDD	Frequency-division duplex
FDM	Frequency-division multiplexing
FDMA	Frequency-division multiple access
FEC	Forward error correction
FES	Fixed earth station
FFT	Fast Fourier transform
FPGA	Field-programmable gate array
FSS	Fixed-satellite service
GEO	Geo-stationary orbit
GI	Guard interval
H-ARQ	Hybrid ARQ
HDTV	High definition television
HIHO	Hard-input hard-output
HPA	High-power amplifiers
HTS	High-throughput satellites
IBO	Input-backoff

ICI	Inter-channel interference
IFFT	Inverse fast Fourier transform
IPDC	Internet protocol datacast
IPoS	Internet protocol over satellite
ISI	Inter-symbol interference
LDPC	Low density parity check
LNA	Low noise amplifier
LoS	Line-of-sight
LTE	Long term evolution
L-TWTA	Linearized travelling wave tube amplifier
LUT	Look up table
MAP	Maximum a posteriori
MC-CDMA	Multi-carrier code-division multiple access
MCSS	Multi-carrier satellite system
MEO	Medium-earth orbit
MF-TDMA	Multi-frequency TDMA
MLSD	Maximum likelihood sequence decoding
MODCOD	Modulation and channel code combination
MPU	Multi-carrier processing unit
MSS	Mobile-satellite service
NACK	Negative acknowledgment message
O3B	Other 3 billion
OBO	Output-backoff
OBP	On-board processing
OECD	Organisation for economic co-operation and development
OFDM	Orthogonal frequency-division multiplexing
OFDMA	Orthogonal frequency-division multiplexing-frequency-division multiple access
PAPR	Peak to average power ratio
PER	Packet error rate
PTS	Partial transmit sequence
QAM	Quadrature amplitude modulation
QEF	Quasi-error-free
QoS	Quality of service
QPSK	Quadrature phase-shift keying
RCFEC	Rate-compatible FEC

RF	Radio-frequency
RSC	Recursive systematic convolutional
RSM	Regenerative satellite mesh
SCSS	Single-carrier satellite communications system
S-DARS	Satellite – digital audio radio service
SDR	Software defined radio
SDTV	Standard definition television
SFN	Single frequency network
SISO	Soft-input soft-output
SNR	Signal to noise ratio
SOVA	Soft output Viterbi algorithm
SR/ARQ	Selective-repeat ARQ
S-RAN	Satellite radio access network
SRS	Signal regeneration satellite
SSPA	Solid state power amplifier
S & W/ARQ	Stop and wait ARQ
TC	Turbo codes
TD	Total degradation
TDD	Time-division duplex
TDM	Time-division multiplexing
TIA	Telecommunications industry association
TWTA	Travelling wave tube amplifier
U/C	Up-converter
UW	Unique word
VSA	Vector signal analyser
VSG	Vector signal generator
WLAN	Wireless local area network

1 Introduction

This Report presents an overview of multi-carrier based transmission techniques over satellite links. Sections 2 and 3 of this Report give a general outline of satellite system design by providing examples that include applications, scenarios and satellite services where multi-carrier transmissions can be used. In § 4 of this Report, basic digital satellite transmission models are described, where more in-depth attention is given to each functional block and its operating principles in § 5-8. Some recent satellite system standards are described in § 9, and simulation and experimental results of multi-carrier transmissions over satellite links are evaluated in § 10 for various performance parameters. Emerging trends in technology that may be used for satellite systems are addressed in § 11.

2 Applications and scenarios

2.1 High definition television/Three-dimensional television

HDTV

High definition television (HDTV) is an example of a service that is traditionally delivered by broadcast satellite services, such as direct-to-home (DTH). HDTV content is widely delivered by satellite broadcast network using either the popular digital transmission standards: digital video broadcasting – satellite (DVB-S) or more recent digital video broadcasting – satellite – second generation (DVB-S2)[1]. Currently, many terrestrial broadcasting networks are delivering HDTV content using the digital video broadcasting – terrestrial (DVB-T) and digital video broadcasting – terrestrial – second generation (DVB-T2) standards, which are also digital formats. While in many parts of the world, analogue broadcasting formats are still employed, the forced transition from analogue to digital use of broadcast spectrum will eliminate its use in developed nations leading this transition.

Satellite DTH presents a viable alternative to terrestrial wired and terrestrial broadcast for delivery of HDTV, as well as, standard definition television (SDTV) content, as it requires very little infrastructure for the set-up of user terminals, while providing wide signal coverage.

The DVB-T and DVB-T2 formats make use of orthogonal frequency-division multiplexing (OFDM) as a means of combating the frequency selective fading that is prevalent in terrestrial networks; but this technology has other benefits as well. The use of OFDM allows for a single frequency network (SFN) – where repeaters are used as gap fillers to enhance signal coverage – without the need for complex equalization filters. In addition, the use of OFDM for broadcast systems allows for more efficient use of spectrum, since it reduces the number of guard bands required between sub-channels. In § 5.2 the benefits of OFDM for satellites are discussed, which includes the reduction in the number of guard bands required for satellite applications. However, due to the high peak-to-average power ratio (PAPR) problems of OFDM for satellite high power amplifiers and perhaps the generally slow time-to-deployment of satellites, OFDM for broadcast satellite has not caught on to date. Yet, it should be noted that new integrated MSS networks, which are described in § 3.2, will make use of OFDM.

Spectrum requirements for HDTV broadcasts are typically 3-5 times larger than those of SDTV, depending on what type of video and audio compression are used [1]. As is explained in § 8.2, the use of DVB-S2 can increase the number of HDTV channels that a satellite DTH service provider can distribute by roughly 33%. This combined with new MPEG-4 compression can enable the delivery of much more HDTV content in the programming of service providers.

3DTV

Three-dimensional television (3DTV) is a relatively new application that adds depth to a traditional two-dimensional HDTV image. 3DTV-ready televisions have already been released to market, the majority of which require the use of 3D glasses that produce a stereoscopic effect to give depth to the image. Although, there are some “auto-stereoscopic” televisions that can produce 3DTV images without the need for glasses. However, these televisions have a narrow viewing angle, lower resolution and can cause eye fatigue [2]. Television broadcasters have also begun to create 3D content. For example, in the United States of America, the Entertainment and Sports Programming Network (ESPN) offered limited 3D programs covering World Cup soccer games, while Sony and Sky Perfect JSAT Corporation have broadcasted 3D World Cup programs in Japan. Additionally,

---

1 For more information on DVB-S and DVB-S2 see § 8.

the Discovery Channel and ESPN are planning to broadcast dedicated 3DTV channels in the near future. It is also worth mentioning that the DVB group is starting work on their first phase to standardize the 3DTV format [3].

There are still many questions to be sorted out for the 3DTV format, for which the answer will have an impact on the bandwidth required to deliver its content. For example, the MPEG-2 and H.264/MPEG-4 advanced video coding (AVC) codecs are currently both being used to deliver SDTV and HDTV formats. Additionally, there are three standards for 3DTV signals, which affect its delivery: checkerboard pattern, panels or full resolution. Checkerboard and panels signals are more simple signals that offer lower resolution than full resolution signals; however, they do not require any additional bandwidth for signal delivery when compared with traditional HDTV. Checkerboard pattern signals enjoy being-first to market on digital light processing (DLP) devices; although they are more difficult to compress than panels signals. The full resolution signal format is created by adding a "depth signal" to the traditional HDTV signal. This additional signal causes an increase in the amount of bandwidth required to deliver the 3DTV signal; however, by using MPEG's latest multi-video coding compression standard for 3DTV, it is possible to compress the full resolution signal such that an additional 70% of bandwidth is required when compared with an HDTV signal [2]. This means that service providers could deliver roughly three 3DTV signals in the same bandwidth as five HDTV signals. This could be more achievable if the operator were to upgrade from DVB-S to DVB-S2, which makes use of MPEG-4 and is more bandwidth efficient. Otherwise, for a lower resolution, the service provider could implement simpler formats, making the adoption of 3DTV by service providers a very cost-effective way to deliver new services at no additional costs (bandwidth).

2.2 Mobile multimedia

The mobile wireless industry has been enjoying very healthy growth over the past decade. For example, US-based AT & T's network has seen a 5 000% increase in data traffic on its mobile network over the past 3 years, driven in part by Apple's iPhone. Other US networks, such as Verizon Wireless have also seen a substantial increase in data traffic on their networks [1]. The increase in data traffic can be attributed to the widespread adoption of new "smartphones" and other mobile internet devices (e.g. laptops with mobile internet cards), whose applications have ever increasing demands for bandwidth. These devices are capable of accessing email, the internet and using several types of applications that include access to several forms of media (e.g. news, photos and videos). As consumers expect their devices to be connected "wherever and whenever", the demand for ubiquitous mobile devices will become more prevalent.

Terrestrial connections outside of city centres can be limited and as a result, the use of satellites for mobile broadband and multimedia is an interesting complement or alternative [4]. The use of ubiquitous hybrid satellite/terrestrial handhelds as a solution for the delivery of mobile multimedia is becoming a prevalent idea. The digital video broadcast – satellite handheld (DVB-SH) standard specifies the transmission scheme for a hybrid satellite/terrestrial device that can receive digital broadcasts for mobile television from both terrestrial and satellite sources. Another way to leverage satellites for mobile multimedia is the use of integrated mobile-satellite service (MSS) systems, which are described in § 3.2. A mobile smartphone could be switched between the satellite component and terrestrial component of the network depending on coverage. In this way, some or all[2] applications used to access content on the terrestrial network could be accessed using the satellite component of the integrated network. An example of a smartphone developed for use on

[2] Some applications may be latency sensitive, which may be problematic for geosynchronous orbit satellite networks, where too much latency may render certain applications unusable.

an integrated MSS is the Genus™ smartphone released by Terrestar, that will make use of AT & T's network (for the terrestrial component) and Terrestar's satellite network (for the satellite component). It is also worth noting that the digital video broadcast – return channel via satellite (DVB-RCS) standard has a mobile option specified in it. This option would allow for mobile and nomadic interactive applications – which can include broadband multimedia access – over satellite.

2.3 Broadband Internet

The definition of broadband differs depending on which government; committee; agency; association; or body is discussing it. In general, broadband Internet is defined as high data rate/speed Internet access that is always connected[3] at speeds much faster than traditional dial-up Internet (56 kbit/s). Modern definitions for broadband can range from 5 to 2 000 times the rate of 56 kbit/s dial-up Internet [5]. However, the definition of broadband Internet will continue to evolve as data rates continue to grow. In [6], the Organisation for Economic Co-operation and Development (OECD), defines broadband as Internet connection at rates exceeding 256 kbit/s for downloads. In [7], the FCC defines the broadband in seven tiers of service, with the first tier ("basic broadband") having a rate in the range of 768-1 500 kbit/s for downloads. In [8], Industry Canada – the department of the government of Canada managing the country's broadband plan – defines broadband as having a connection of at least 1.5 Mbit/s download and 384 kbit/s upload speeds. In [5], the ITU defines broadband Internet as an Internet service providing data rates of 1.5-2 Mbit/s or more for download, which is the median download speed for a typical digital subscriber line (DSL) connection.

National broadband strategies, which seek nationwide connectivity through government subsidization with the goal of further economic growth, are leading to increased interest in the deployment of broadband Internet. Many of these strategies seek to connect citizens that are yet to be served by broadband, with each institution having their own particular definition of what constitutes a broadband connection. These citizens are typically in rural communities where broadband penetration is low due to the high cost of building infrastructure. As acknowledged in the federal communications commission's (FCC) National Broadband Plan [9], broadband satellite could be a viable way to achieve broadband penetration for those in communities that are underserved and where it would be prohibitively expensive to build terrestrial infrastructure [10]. Currently, most satellite Internet services are provided at speeds that are lower than the minimum data rates set forth in [5, 7, 8], however, there are many that would qualify as a broadband service under the OECD figures. To access government subsidies for Internet broadband, satellite companies will have to meet the minimum rates as defined in each national strategy.

High-throughput satellites (HTS) are being developed by several companies worldwide that could meet the minimum rates defined for broadband. For example, in the US, Viasat and Hughes Network Systems are developing HTS, where they are expected to provide download speeds of 2-10 Mbit/s and 5-25 Mbit/s, respectively [9]. However, these HTS will operate in geosynchronous orbits and, as a result, are more susceptible to latency than LEO and MEO satellite. Too much latency could hinder the use of interactive real-time applications. However, the latency issue is less relevant for applications that require best-effort network performance, such as e-mail and Internet browsing. To attempt to overcome the latency issue, the Other 3 Billion (O3B) [11] partnership is designing a satellite broadband Internet service that will use medium-earth orbit (MEO) satellites. This design will include the deployment of several MEO satellites – at 1/5 the distance of GEO satellites – that use flexible spot beams to connect developing nations to broadband Internet. O3B

[3] Always connected means that the service is available without having to first connect to it, such as is the case with dial-up.

claims that by reducing the distance of satellites from the earth, the latency of round-trip delay will be far less.

Although standards that define user interaction via satellite, such as DVB-RCS, do not include the use of OFDM, its use may be of some benefit. Although not a necessity, routing and capacity flexibility are desirable characteristics for broadband satellites; where efficient routing and capacity allocation greatly improves the performance of the system. More flexible routing and capacity can be achieved by using multiple carriers per beam in a frequency-division multiplexing (FDM) manner. That is, by implementing OFDM. This could be easily accomplished where on-board processing is available on the satellite by using efficient fast-Fourier transform (FFT) engines.

3 Satellite systems examples

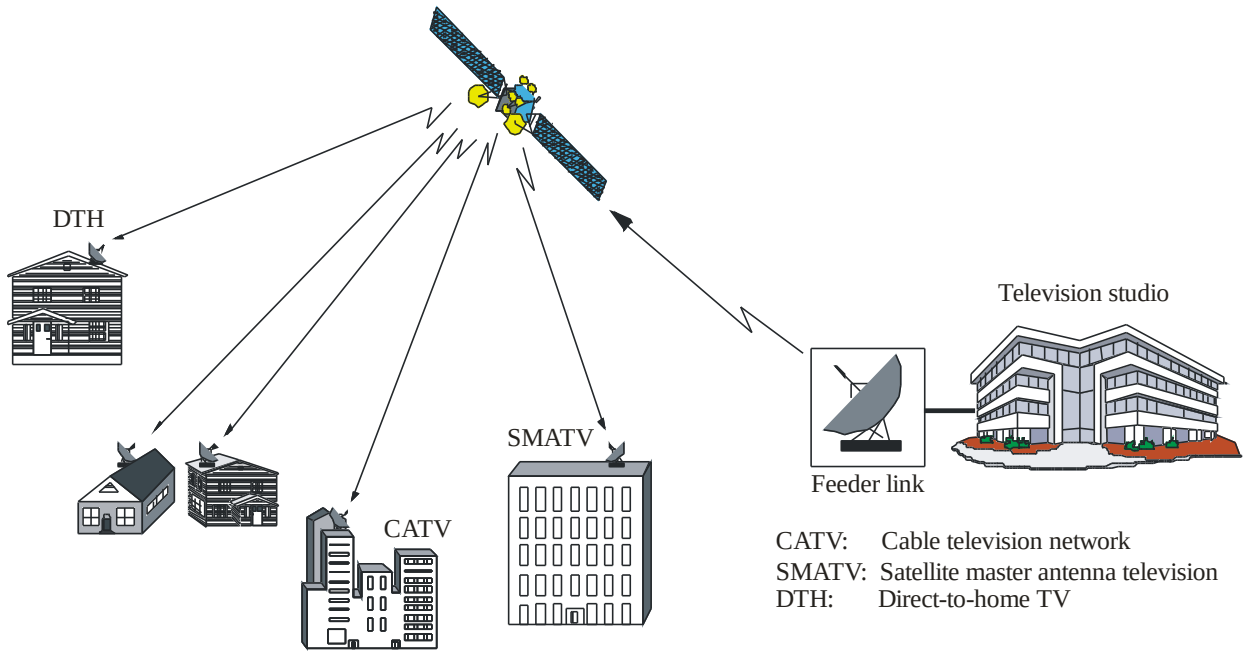
3.1 17/24 and 21/24 GHz BSS systems

A broadcasting-satellite service (BSS) is defined as a radiocommunication service in which signals transmitted or re-transmitted by space stations are intended for direct reception by the general public. The satellites implemented for BSS are often called direct broadcasting-satellites (DBS's). The antennas required for BSSs should be smaller than those used for fixed-satellite services (FSS's) [12]. Figure 1 demonstrates an example of a BSS with a star-network topology, where communication is one-way, from feeder link to user terminals via satellite.

In 2003, new frequency allocations were made for BSS, which vary depending upon the Region of operation.

FIGURE 1

Example of BSS



In Region 2, 17.3-17.8 GHz for BSS and 24.75-25.25 GHz for FSS feeder links for BSS are allocated for operation as a primary service. These bands are shared with the Earth-to-space (E-S) segment of FSS's. Where in the 17.3-17.8 GHz bands the use of E-S segments for FSS is strictly limited to feeder links for BSS. It should also be noted that the 24.75-25.25 GHz bands are shared with FSS's; however, feeder link for BSSs have priority over other types of FSS's.

In Regions 1 and 3, the 21.4-22.0 GHz band BSS can operate as a primary service. Region 3 also uses 24.75-25.25 GHz for FSS feederlinks for BSS, where FSS feederlinks for BSS have priority over other types of FSSs. In particular, the introduction of HDTV BSSs in the 21.4-22 GHz band is governed by Resolution 525 (Rev.WRC-07) of the Radio Regulations (RR). In its Annex (Section I – General Provisions) Resolution 525 (Rev.WRC-07) states that all services other than the BSS, in the band 21.4-22.0 GHz in Regions 1 and 3 operating in accordance with the Table of Frequency allocations may operate subject to not causing harmful interference to BSS (HDTV) systems nor claiming protection from such systems. It further elaborates that it shall be understood that the introduction of an operational BSS (HDTV) system in the band 21.4-22.0 GHz in Regions 1 and 3 should be regulated by an interim procedure in a flexible and equitable manner until the date to be decided by WRC-12.

3.2 Integrated MSS systems

An integrated MSS system is defined in [13, 14] as a system employing a satellite component and ground component where the ground component is complementary to the satellite component and operates as and is an integral part of the MSS system. In such systems the ground component is controlled by the satellite resource and network management system. Further, the ground component uses the same portions of MSS frequency bands as the associated operational MSS.

An integrated system provides a combined (integrated) single network that uses both a traditional MSS link and terrestrial transmission paths to serve mobile end-users. A typical integrated system comprises one or more multi-spot beam satellites and a nation-wide or regional ensemble of terrestrial cell sites, where both terrestrial and space components communicate with mobile terminals using a common set of MSS frequencies. Such systems are referred to as MSS-ancillary terrestrial component (MSS-ATC) in the United States of America and Canada, and MSS-complementary ground component (MSS-CGC) in Europe and are implemented in the 1-3 GHz bands.

Integrated systems will likely have various service components, including traditional MSS services. A rather large portion of this portfolio of services will be devoted to the provision of broadband services – including multimedia broadband services – to handheld or portable terminals. These handhelds and portable terminals are expected to have form and cost factors very similar to terrestrial cellular terminals of today [15]. There is a variety of approaches that system operators may choose to implement their baseline service provisioning, coverage goals and end-user demand.

Initial examples of integrated satellite systems are represented by satellite digital audio broadcasting (DAB) to mobile receivers mounted on vehicles in the early 2000s. Broadcasting of high-quality digital radio channels to mobile users was achieved by means of S-band frequencies around 2 GHz. This was called Satellite – Digital Audio Radio Service (S-DARS). Two companies in the USA, XM-Radio and Sirius, provided these services [16, 17]. The remarkable success achieved by XM-Radio and Sirius was based on traditional counter-measures such as a high link margin; time, frequency and satellite diversity; and the use of terrestrial repeaters in urban areas.

More recently, in Korea and Japan, S-DMB service to hand-held user terminals was successfully deployed utilizing a geostationary (GEO) satellite [18]. The S-DMB system was based on code division multiplexing (CDM) technology, described in Recommendation ITU-R BO.1130-4 [19]. In

Europe, the "Unlimited Mobile TV" concept was introduced [20], which is based on the ETSI standard of DVB – Satellite services to Handheld (SH) devices published in March 2008 [21]. With this concept, a hand-held mobile terminal is supposed to receive broadcast signals from both the satellite and the CGCs. Video services from ICO G1 (a GEO satellite that covers the entire United States of America, Puerto Rico and the US Virgin Islands) are based on the DVB-SH standard as well. The DVB-SH system was designed for frequencies below 3 GHz (S-band). It complements the existing DVB-handheld (DVB-H) terrestrial coverage and analogously uses the DVB IP datacast (IPDC) set for content delivery, electronic service guide and service purchase, and protection standards.

3.3 Ka-band broadband systems

Increasing demand for broadband services can be effectively handled by a satellite system using high frequency bands such as the Ku and Ka bands. Especially, to provide high-speed Internet and television (TV) services to maritime and air vehicles, a satellite system may be the only possible option. In this case, an active array antenna that is mounted on a moving vehicle is used to track a satellite and provide seamless connections.

For Ka-band broadband systems, the volume of traffic on the forward link, which provides connections from the satellite gateway to the user terminals, is much greater than that on the return link, which provides connections from the user terminals to the satellite gateway. Recommendation ITU-R BO.1709-1 specifies three air interface standards, which can be used to implement broadband satellite networks [22], as shown in Table 1.

TABLE 1

Air interface standards for broadband satellite systems in Recommendation ITU-R S.1709-1

Standard name / Scheme	ETSI EN 301 790	TIA-1008-A	ETSI-RSM-A
Network topology	Star or mesh	Star	Star or mesh
Forward link scheme	DVB-S	DVB-S	High rate TDMA
Forward link data rate (Mbit/s)	1-45	1-45	100, 133.33, 400
Return link modulation	QPSK	O-QPSK	O-QPSK
Return link multiple access	Multi-frequency (MF)-TDMA	MF-TDMA	FDMA-TDMA
Return link data rate	No restriction	64, 128, 256, 512, 1 024, 2 048 ksymbol/s	2, 16, 128, 512 kbits/s

The DVB via satellite (DVB-S) standard, which is specified as the forward link scheme for the first and second air interfaces in Table 1, describes an air interface for the satellite multi-program TV service. As a result of the rapid evolution of digital satellite communication technology since the introduction of DVB-S[4] in 1994, a second generation standard for digital video broadcasting via satellite, DVB-S2, has been published. Remarkable improvements are achievable for DVB-S2, including a new channel coding scheme and higher-order modulation, which used as part of

[4] For more information on DVB-S see § 9.1.

an adaptive coding and modulation (ACM) technique. ACM is used mainly to compensate for rain attenuation, which is more prevalent in high frequency bands. The new channel code in DVB-S2 is a concatenated code with low density parity check (LDPC) codes and Bose–Chaudhuri–Hocquenghem (BCH) codes. The higher-order modulation scheme is based on amplitude and phase shift keying (APSK) constellations up to order 32, which has been shown to be more robust to non-linearities caused by high power amplifier distortion than quadrature amplitude modulation (QAM) [23][5].

The air interface in the first column of Table 1 is called a DVB- return channel by satellite (DVB-RCS), and in combination with DVB-S/S2 it provides two-way broadband satellite systems. Because the current version of DVB-RCS does not consider mobile service environments, an advanced version for comprehensive support of mobile and nomadic terminals is currently under study[6]. The second air interface in Table 1 is the Internet Protocol over Satellite (IPoS) standard that has been developed by Telecommunications Industry Association (TIA) in the USA. The third air interface in Table 1 is the Broadband Satellite Multimedia (BSM) standard developed by ETSI. Important characteristics of the BSM architecture are that its operational functions are separated into two types: satellite dependent functions and satellite independent functions. The purpose of this separation is to provide the capacity to incorporate future market developments, as well as the flexibility to include different market segment-based solutions in the higher layers of the protocol stack. Among the several types of BSM air interface families, Regenerative Satellite Mesh (RSM) – A is based on a satellite with onboard processing (OBP) such as SPACEWAY by Hughes Network Systems, which supports a fully meshed topology. With RSM-A, data can be transmitted between any pair of satellite terminals in a single hop.

The Electronics and Telecommunication Research Institute (ETRI) in Korea has developed a mobile broadband interactive satellite access technology (MoBISAT) system based on the DVB-S/DVB-RCS standard for satellite video multicasting and Internet via wireless local area networks (WLANs) [24]. The system can provide broadband services to passengers and the crews of land, maritime, and air vehicles, via installation of group user terminals with a two-way active antenna. To access the Internet inside vehicles, WLANs can be provided, enabling the use of laptop PCs or PDAs on board the vehicle.

4 System implementation methods

4.1 Single and multi-beam satellite systems

A geosynchronous orbit satellite communication system with large satellite antenna(s) can provide high data rate (or high speed) services to small user terminals thanks to a large satellite antenna gain. A geosynchronous orbit satellite system with multi-beam antennas has a larger capacity than a system with a single global beam over the same service area [25]. For a multi-beam geosynchronous orbit satellite system, we can easily synchronize all the downlink (satellite-to-user link) signals from a single satellite because the satellite is the only source of the signal. For a synchronous multi-beam satellite system, the beam signals received by the user are also synchronized, regardless of the user location. Therefore, inter-beam interference can be mitigated.

---

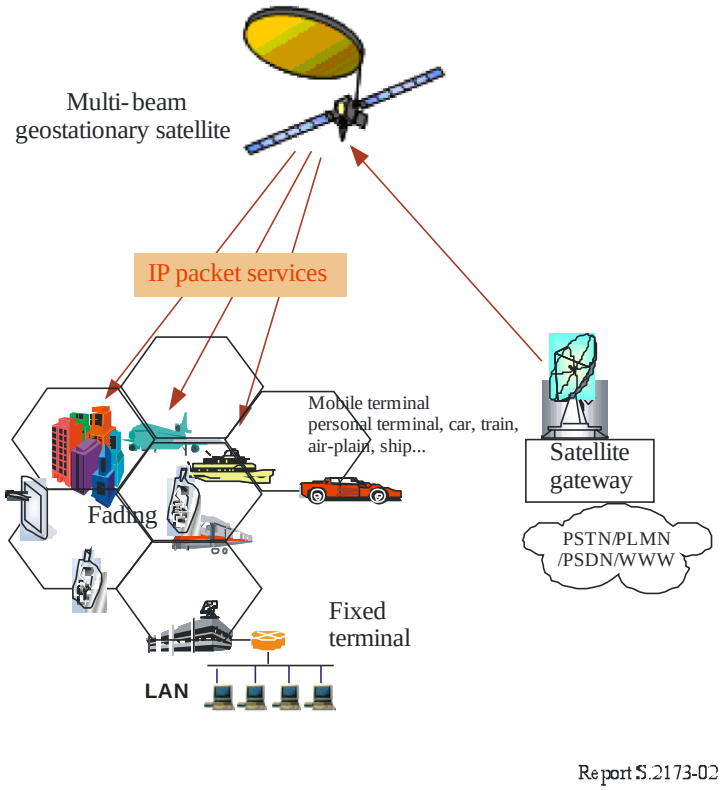
5 For more information on DVB-S2 see § 9.2.

6 For more information on DVB-RCS see § 8.3.

Figure 2 illustrates a multi-beam satellite system providing IP packet services. Services for mobile users are linked to a terrestrial IP core network through a fixed earth station (FES) and satellite. The geosynchronous orbit satellite has to be equipped with a large directional antenna, in order of about 20 meters, so as to be able to provide high-speed services for nomadic and portable terminals equipped with small antennas. The FES performs adaptive resource allocation on the satellite downlink and is a gateway that links user services to the terrestrial network. When the satellite has on-board processing capability, it can perform adaptive resource allocation [26].

FIGURE 2

Multi-beam satellite system providing IP packet services



4.2 Digital satellite transmission system

Figure 3 shows a simplified digital satellite transmission system viewed for a single link. Depending on how the system is implemented, a number of modulated and filtered data can be multiplexed and sent through this link. In this section, the basic principles of each functional block are described and a detailed description of each follows in § 5-7.

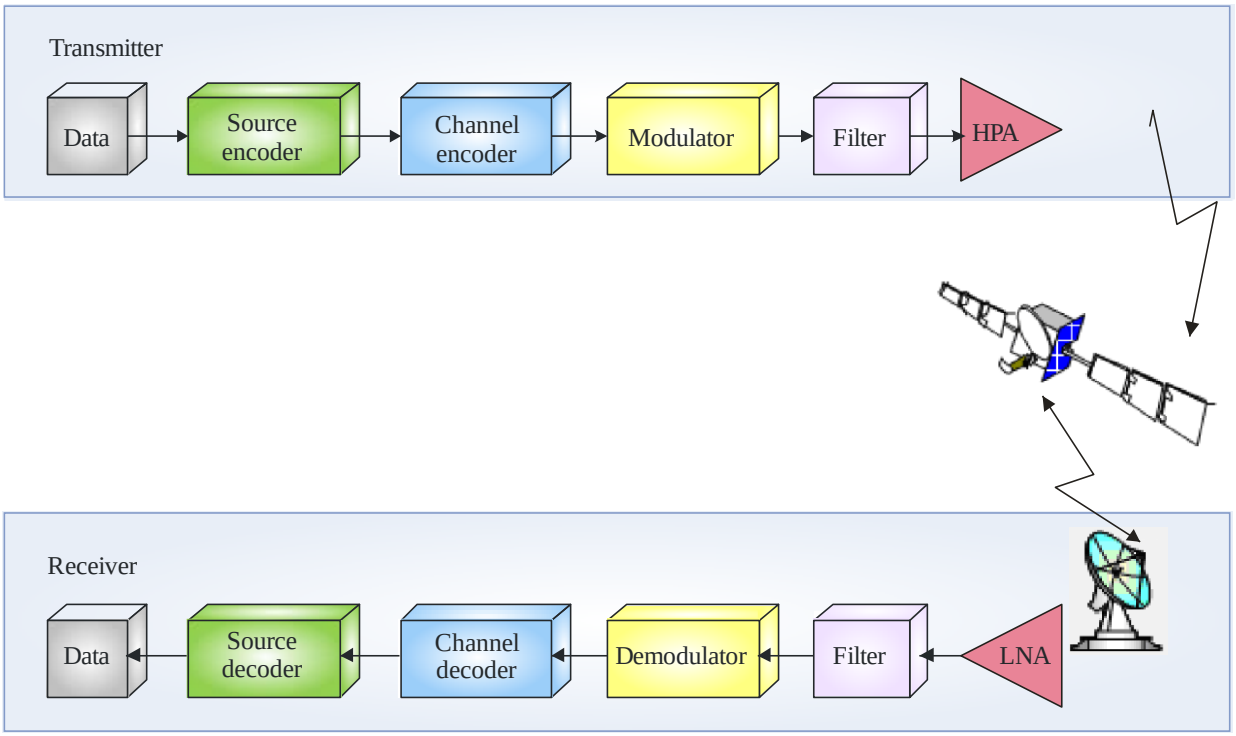
The source encoder in Fig. 3 transforms information into a given digital representation, and can possibly include a mechanism for the removal of any inherent redundancy within the data. Channel encoding is applied to the output of the source encoder. Channel coding, which is also called error control coding, is a digital processing technique aimed at averaging the effects of channel noise over several transmitted signals. A detailed discussion on channel coding is provided in § 7.

The modulator converts the digital information into a signal that can be easily transmitted through the channel. This is usually achieved by translating the frequency of the source information onto a carrier frequency, which is a frequency that is much higher than that of the un-modulated source signal. To modulate the digital information, the digital stream is first put into a baseband representation. The most commonly used digital modulation scheme for satellite systems is phase

shift keying (PSK), in which the data determines the phase of the transmitted signal. In bandwidth-limited conditions – as is common for satellite systems – utilization of multi-level modulation most often will increase the bit-rate of the satellite system. Discussion on multi-level modulation schemes specified in recent satellite system standards are described in § 9.

FIGURE 3

Digital satellite transmission system



This Report emphasizes the utilization of a multi-carrier modulation scheme and its utilization for multiple access. Section 5 describes details on multi-carrier modulations.

After filtering is applied to remove the sideband signals, the modulated signal is amplified by a high power amplifier (HPA). As the transmitted signal travels a long distance, and as a result, is subject to larger propagation loss than typical terrestrial systems, a very high power amplification is required. Generally two different types of the HPA's are used by satellite systems: travelling wave tube amplifier (TWTA) and solid state power amplifier (SSPA).

At the receiver, signal recovery operations are applied in the reverse order that they were carried out by the transmitter. The received signal is passed through a low noise amplifier (LNA) and then filtered to remove the sideband signals. The demodulator uses the filtered signal to create decisions on the transmitted digital information and passes those decisions on to channel decoder. Depending on the requirements of the channel decoder, the decisions produced by the demodulator can be hard (i.e. zero or one), or soft. Soft-decision decoding uses metrics based on the likelihood that a transmitted bit is zero or one to make a final decision. This soft-decision metric is required by iterative channel decoders, which are discussed in detail in § 7.

5 Multi-carrier and multiple access systems

5.1 Basics of multi-carrier transmission

Multi-carrier techniques have become very popular for high data rate applications, and many next-generation terrestrial standards include multi-carrier-based modulation and multiple access techniques. The most powerful advantage of multi-carrier techniques is their capability to eliminating problem deriving from intersymbol interference (ISI), alleviating the high burden of complex time-domain equalizers while maintaining a high rate of transmission. Particularly, in rapidly varying channels, such as wireless channels, the burden of maintaining accurate values for a many-tap equalizer can be a major burden.

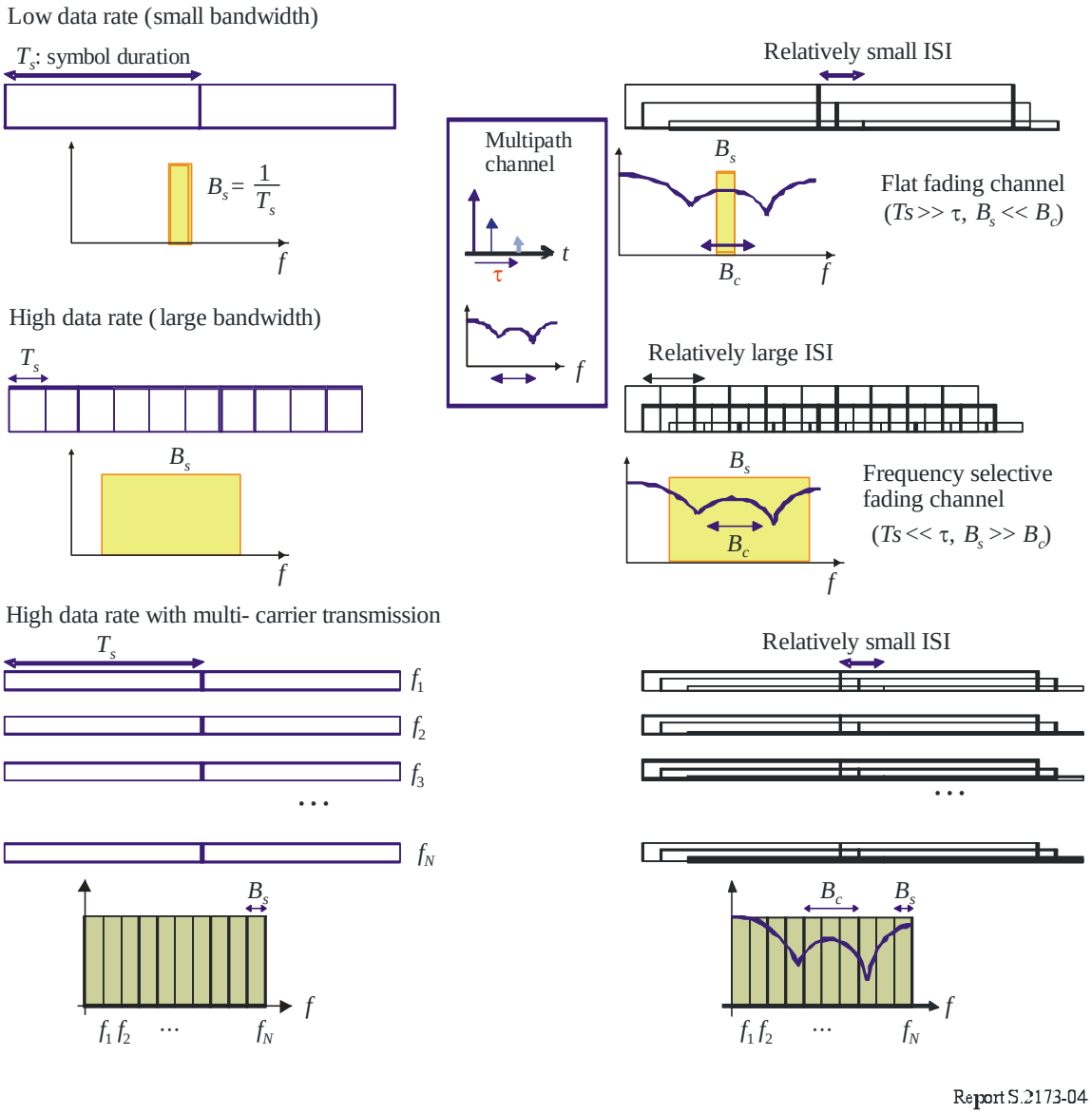
In order to have a channel that is ISI-free, the symbol time *Ts* has to be much longer than the delay spread of the channel, τ. Typically, it is assumed that *Ts* should be at least ten times larger than τ in order to satisfy an ISI-free condition. The bit error rate (BER) performance of a digital communication system seriously degrades if *Ts* approaches τ. In order to satisfy high data rate requirements, the desired Ts is usually much smaller than τ, which leads to severe ISI problems. In the frequency domain, the channel characteristics become selective. In this situation, channel gains are attenuated and enhanced over the system bandwidth, *Bs*, because the coherence bandwidth of the channel, *Bc* is larger than *Bs*.

Figure 4 illustrates the basic concept of multi-carrier transmissions. Multi-carrier modulation divides the high-rate transmit bit-stream into *N* lower-rate substreams, each of which has $NTs >> \tau$, and thus ISI problem can effectively be eliminated. These individual substreams can then be sent over *N* parallel subchannels, maintaining the total desired data rate. In the frequency domain, the subcarriers have *N* times less bandwidth, which equivalently ensures a much smaller subcarrier bandwidth than *Bc*; or in other words, a relatively flat fading condition for each substream.

Orthogonal frequency division multiplexing (OFDM) is a technology where each subchannel in a multi-carrier transmission is orthogonal to all other subchannels. By using OFDM, the bandwidth efficiency can be further increased when compared with a conventional frequency division multiplexing (FDM) system. In addition, no complex carrier de-multiplexing filters are needed to separate the carriers, as simple fast Fourier transform (FFT) engines can be used to do demodulation. Figure 5 shows a simple block diagram for an OFDM modulator. *N* baseband modulated symbols are first converted to *N* parallel streams with symbol periods that are *N* times longer, and are then passed through an inverse FFT (IFFT) engine. Although *N* times longer, the lengthened symbol duration can reduce the problems arising from ISI between the modulated symbols that compose a particular OFDM symbol. If a delay spread remains between OFDM symbols, a guard time or guard interval (GI) is inserted between OFDM symbols in order to eliminate the remaining delay spread. As plain insertion of a GI may cause inter channel interference (ICI) between sub-carriers, a cyclically extended GI must be inserted. This cyclic GI is called a cyclic prefix (CP), as each OFDM symbol is preceded by a periodic extension of the signal itself.

FIGURE 4

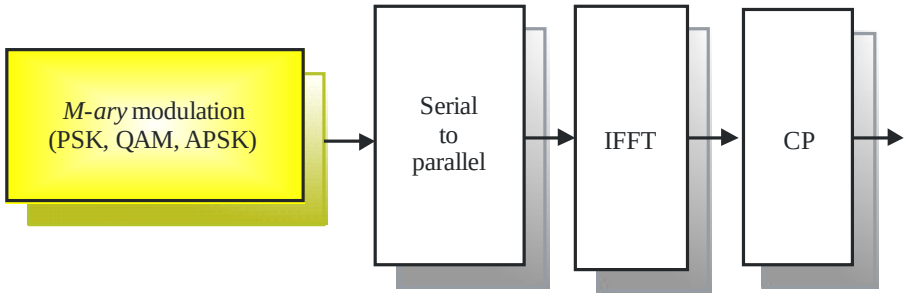
Basic concept of multi-carrier transmission



Report S.2173-04

FIGURE 5

Block diagram of an OFDM modulator



Report S.2173-05

5.2 Multi-carrier transmission over a satellite link

Many terrestrial system standards have adopted OFDM-based transmission schemes. Considering that OFDM-type systems are largely used in terrestrial networks as a means of providing good spectral and energy efficiency over frequency selective channels, the idea of utilizing them for satellite systems may seem to lack justification. This is due to the fact that conventional satellite systems do not experience frequency selective channel. Moreover, multi-carrier systems suffer from high peak to average power ratio (PAPR) problems, which impose a high burden on the high power amplifiers of satellite systems.

Even with the aforementioned obstacles, there are a number of factors that may make this technology attractive for satellite services [27]:

1. A satellite component of a next-generation network may be regarded as providing coverage extension for service continuity of the terrestrial component by allowing for vertical handover. For cost-effective vertical handover, future satellite radio interfaces may be compatible and have a high degree of commonality with terrestrial interfaces. This may enable the reuse of terrestrial component technology, to minimize the user terminal chipset and network equipment for low cost and fast development.
2. In broadband satellite communications where routing and capacity flexibility over the coverage area require the use of multiple carriers per beam with FDM.
3. In a scenario where an on-board switch is also used to route the time division multiplexing (TDM) data from a given gateway to different beams, OFDM may simplify the architecture of the on-board switch. No complex carrier de-multiplexing filters are needed to separate the carriers because simple FFT engines can be used.
4. Most importantly, OFDM techniques may represent a means to increase the spectral efficiency by compacting carriers, eliminating guard-bands for applications requiring high speed transmissions.

An existing example where OFDM transmission is adopted for satellite is the DVB-SH standard. DVB-SH uses the same signal format defined in DVB-H for terrestrial systems. The main reason for the adoption of OFDM in DVB-SH is that satellite and terrestrial transmitters in this network topology form a single frequency network (SFN). This kind of configuration uses the available frequency resources efficiently, as the same frequencies are used for both satellite and terrestrial transmitters because the same content is transmitted on both links. Terrestrial transmitters can be fed by receiving the satellite signal, if the transmitted and received signal is sufficiently isolated [28]. In addition to the high spectral efficiency of multi-carrier transmission, there is the added flexibility of allocating resources for adaptive use, i.e. not only in terms of time but also in terms of frequency domains [26]. In order to reduce the high PAPR incurred in the multi-carrier transmission technique, a number of approaches can be applied, which are discussed in § 5.

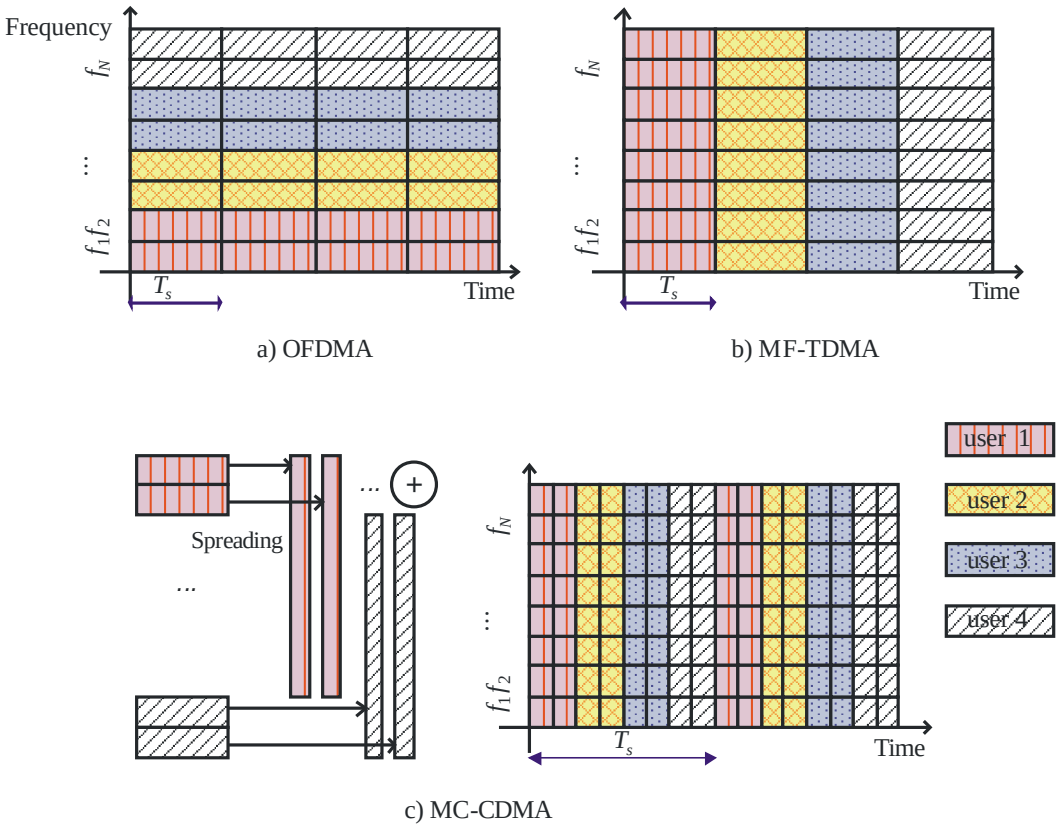
5.3 Multi-carrier based multiple access schemes

Due to increasing interests in multi-carrier transmissions schemes, the multiple access schemes associated with them are considered as one of the promising technologies for next-generation networks. Examples of these are multi-carrier code division multiple-access (MC-CDMA), multi-frequency time division multiple access (MF-TDMA) and orthogonal frequency division multiplexing – or frequency division multiple access (OFDM-FDMA; OFDMA). An example application where one of these schemes is employed is MF-TDMA, as defined in DVB-RCS.

The basic concept of multi-carrier based multiple access scheme is that all N subcarriers are shared by multiple users in the system. Figure 6 compares the basic concept of OFDMA, MF-TDMA, and MC-CDMA. In OFDMA, a number of users share subcarriers, of which a portion is allocated to each user. On the other hand, in MC-CDMA, each user's data is first spread by an allocated spreading code and are then transmitted using all the subcarriers simultaneously.

FIGURE 6

Comparison of OFDMA, MF-TDMA and MC-CDMA



6 Peak-to-average power ratio reduction technologies

6.1 Introduction

One of the major drawbacks of multi-carrier schemes is the high peak-to-average power ratio (PAPR) of the transmitted signal. In satellite communication systems, the earth station transmitter and the satellite transponder employ high-power amplifiers (HPA)s. It is desirable to operate the HPAs as close as possible to their saturating point to maximize their radio-frequency (RF) power efficiency. However, operation of these HPAs near their saturation point seriously degrades the performance of modulation schemes which do not have a constant envelope. Therefore, the selection of a suitable modulation scheme combined with a PAPR reduction method is required in order to reap the advantages of multi-carrier satellite systems. It should be noted that non-discrete individual beam signals described in the MC-CDMA technique are often referred to as OFDM signals. The following section reviews a number of techniques that may be used to reduce the PAPR of multi-carrier signals.

6.2 Peak-to-average power ratio reduction techniques

Block scaling

In the case of block scaling, critical OFDM symbols are identified by comparing their amplitude with one or more thresholds. If any of the samples in a symbol exceed the thresholds, then the amplitude of the whole symbol is reduced by multiplying each of its samples by a factor less than unity. This method requires the transmission of side information to tell the receiver how to restore the signal to its original amplitude. Although this causes noise enhancement, it is expected that coding can more easily correct errors introduced by Gaussian noise than bursts of errors resulting from the clipping effect of a saturated amplifier.

Channel coding

A number of techniques suggest employing a proper form of channel coding to avoid transmission of the symbols that exhibit a high PAPR. The basic idea of these techniques is to exploit the redundancy introduced by a properly chosen code, which allows the system to refrain from transmitting those sequences that give rise to a high PAPR. It is also desirable to exploit the properties of the code to perform some sort of error correction.

By considering the instantaneous power of an OFDM signal in its discrete time form, it can be shown that the resulting PAPR depends on the auto-correlation function of the sequence of data to be transmitted once modulation parameters are fixed. Studies have demonstrated that it would be possible to transmit only the symbols belonging to a subset which would ensure a lower PAPR. However, the computational requirements to identify a set of suitable sequences are prohibitively large for practical numbers of sub-carriers. An automatic coding technique is preferred for an efficient implementation.

Golay sequences are known to have the interesting property of producing a PAPR which is always limited, when transmitted with phase-modulated carriers. The PAPR for a Golay sequence is upper bound at 3 dB. However, the codes have a rate ranging from 0.3 to 0.4. It has been demonstrated that as the number of sub-carriers increases, the redundancy required for proper PAPR reduction is high, for only a modest error correction capability. This substantially reduces the spectral efficiency of the system.

Partial transmit sequence scheme

Partial transmit sequence (PTS) scheme is one of the most efficient approaches in terms of PAPR reduction performance. With this scheme, a phase shift is applied to disjoint sub-blocks of the data sequence, hence the name partial transmit sequence. The combination of such blocks with different phase shifts gives the wanted different alternatives for transmission. PTS is flexible in that it can be used with an arbitrary number of subcarriers and any type of modulation scheme. However, it is difficult to implement because its implementation has exponentially increasing complexity; although, a reduced-complexity technique can be used at the sacrifice of some performance. The PTS scheme also requires the transmission of a substantial amount of side information. Reference [29] investigated a reduced complexity PTS scheme for a satellite system using MC-CDMA.

Sub-carrier scrambling

For subcarrier scrambling, each user's data symbol is spread by the orthogonal user code in the frequency domain and thus, there is a correlated pattern among subcarriers. Since the correlated pattern among subcarriers influences the PAPR, the PAPR is highly dependent on the patterns of the orthogonal user codes used. This implies that the PAPR of the multi-carrier signal will be reduced if disturbances are introduced into the correlation or coherence among the subcarriers [30].

One simple way to accomplish this is to scramble the polarities of the subcarriers. Randomly scrambling the polarities of the subcarriers removes the correlation, and limits the coherent combination among the subcarriers irrespective of the user codes used. Therefore, the PAPR of each user code (or code combinations) becomes statistically equivalent to the others and is much smaller than that of the worst-case PAPR of no scrambling.

A further PAPR reduction can be achieved by using the fixed scrambling patterns that are separately optimized for each user code (or user code combination) and holding the scrambling pattern while the same code combination is maintained. If the optimum fixed scrambling patterns for each user code combination are searched prior to system operation, they can be allocated dynamically according to the current user code combination during the system operation. Besides its great PAPR reduction capability, this method has two advantages. First, it does not require any additional real-time computation for PAPR reduction. Secondly, it requires negligible side information. In this scheme, the system has to broadcast to the users which scramble pattern is currently in use. The amount of information for this is relatively small and needs to be sent only at the beginning of each connection/disconnection procedure.

Dynamic allocation of the scrambling pattern may add a processing load in busy traffic conditions. In order to avoid this, the system can employ a sub-optimum solution. From an extensive search of scrambling patterns, a single scrambling pattern that maintains a PAPR close to the minimum over all code combinations can be found. This sub-optimum solution can be found before system operation via extensive computer search, meaning that the system can reduce the PAPR without any additional processing burden or any side information.

Tone reservation

In this scheme, the system abstains from transmitting data on a limited number of sub-carriers, which are then modulated in a way that opposes the high peaks that result from data modulated sub-carriers. With a proper combination of the reserved tones, a peak-reduction signal can be superposed onto the information bearing OFDM symbol, such that the peak-reduction signal is in phase opposition to the original symbol at the time instants where there are peak amplitudes. Consequently, the modified symbol has a lower PAPR.

The generation of the peak-reduction signal needs to take place for each OFDM symbol through an optimization algorithm. The number of sub-carriers allocated to peak-reduction must be sufficient to be effective, but sufficiently small so as not to excessively sacrifice the information rate.

Pulse superposition

The pulse superposition technique acts directly on the highest power peaks of the OFDM signal. Once the signal has been generated, an anti-peak signal is added to it, resulting in a signal having a lower PAPR. The peak reduction is thus completely executed in the time domain, and can use all available information on the characteristics of the signal to be modified because the signal has already been generated when the PAPR reduction is applied. Pulse superposition obtains a reduced PAPR by construction. The generation of the anti-peak signal is made directly in the time domain and thus does not necessitate additional inverse fast Fourier transform (IFFT) processing.

The anti-peak signal is built from the replica of a carefully selected elementary pulse. The shape of the elementary pulse is very important, because it is expected to reduce a high peak without leading to the creation of secondary peaks, and it must have a spectrum without components that are out of the band of the original signal. The effectiveness of the pulse superposition technique is similar to that of the block scaling approach, but it does not suffer from noise enhancement because weak signals are unaffected.

Non-linear companding transforms

Non-linear companding transforms compand the original OFDM signals using strict monotone increasing functions. The companded signals at the transmitter can be recovered correctly through the corresponding inversion of the non-linear transform function at the receiver. Non-linear companding transforms enlarge the small signals and compress the large signals to increase the immunity of small signals to noise.

The non-linear companding transform is in essence a type of clipping scheme, but it does not suffer from the in-band distortion, out-of-band radiation and peak regrowth after digital to analogue conversion that affects simple clipping schemes. Since the distribution of the original OFDM signals are known to have Rayleigh distributions, the non-linear companding transform function can be obtained theoretically on the basis of the desired distribution of the companded OFDM signals. Two types of non-linear companding transforms based on the error function and the exponential function have been proposed.

Non-linear companding transforms are a type of non-linear process that may lead to significant distortion and performance loss caused by companding noise. However, unlike AWGN, companding noise is generated by a known process that can be recreated at the receiver, and subsequently removed. An iterative type receiver has been proposed to eliminate companding noise for a companded and filtered OFDM system.

6.3 CI-OFDM

6.3.1 Introduction

The implementation of a multi-carrier satellite system is impeded by its HPA's inability to deal with the high PAPR of OFDM signals. This makes PAPR suppression techniques highly attractive for the design of a multi-carrier satellite system. Carrier Interferometry OFDM (CI-OFDM) [31] is a low-complexity PAPR suppression scheme that has been shown to mitigate the PAPR of a communications system, while providing good packet-error rate (PER) performance – as is shown in § 10.3.2 of this Report.

6.3.2 CI-spreading technology

CI-OFDM is a type of sub-carrier scrambling technology that can be implemented in an OFDM system at the cost of an additional FFT module at the transmitter and receive end. Therefore, the number of OFDM carriers has little impact in increasing the scale of hardware [32].

Figure 7 shows the basic configuration of a CI-OFDM transmitter, as described in [31]. Here, modulated symbol samples, s_i, are multiplied by a spreading code $\mathbf{CI}_i$, where $CI_i^k = \exp(j2\pi ki/N)$, and multiplied by each subcarrier. This effectively creates a CI-OFDM signal where each symbol is transmitted on all subcarriers. The CI-OFDM modulated samples, x_i:

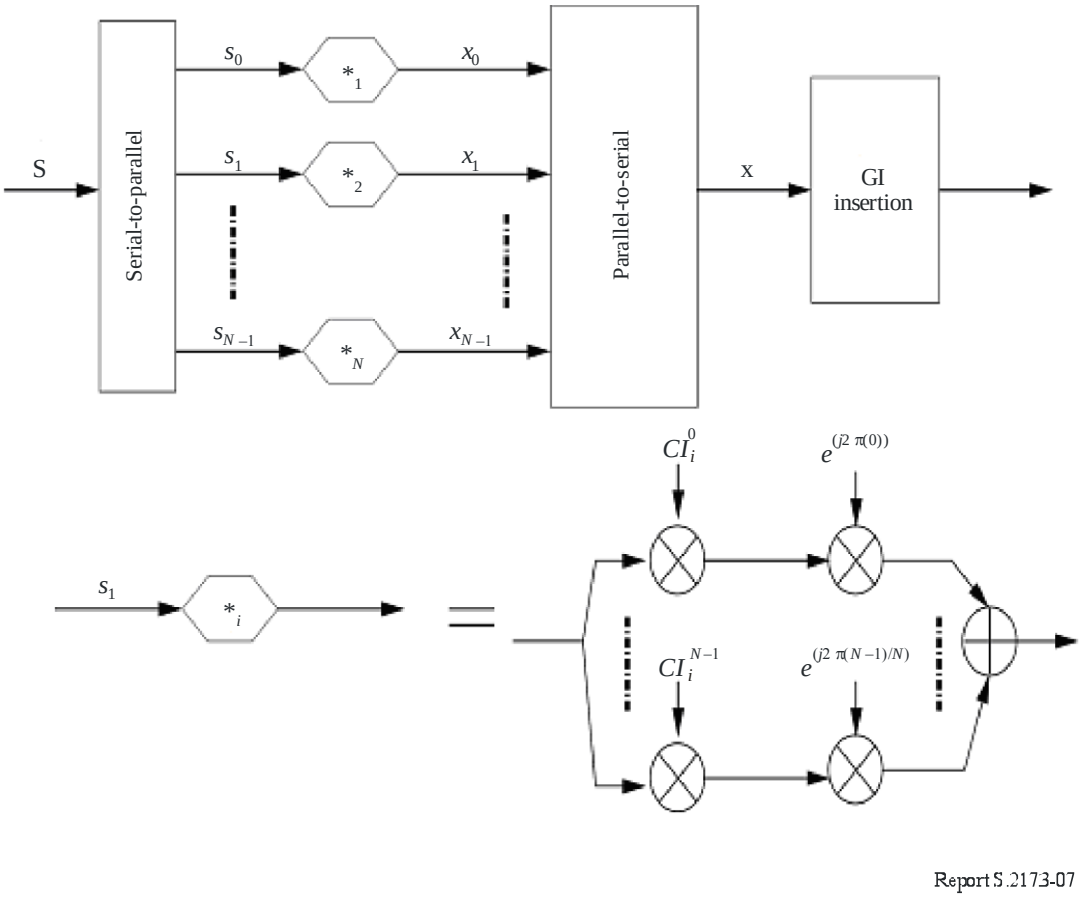
$$x_i = \sum_{l=0}^{N-1} \sum_{k=0}^{N-1} e^{j2\pi ki/N} e^{j2\pi l/N} s_i \tag{1}$$

are then combined to create a multi-carrier vector, **x**, representing the time samples of the CI-OFDM modulated signal. A GI, which is chosen based on the characteristics of the channel, is appended to **x** to avoid interference between subsequent multi-carrier symbols due to the delay spread of the channel. Each individually spread OFDM symbol has its peak when all other symbols are at their minimum, avoiding the constructive addition of subcarrier peaks that leads to high

a PAPR. This creates a nearly flat CI-OFDM signal, as demonstrated in Fig. 8. It should be noted that the summations in equation (1) are two inverse discrete Fourier transforms (IDFT)s of s_i. The two IDFT operations performed in equation (1) can be replaced by two inverse FFT engines [32], as depicted in Fig. 9.

FIGURE 7

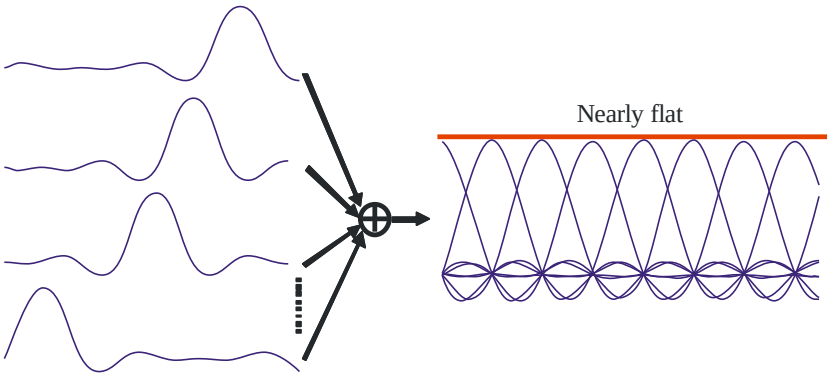
Basic configuration of CI-OFDM transmitter



Report S.2173-07

FIGURE 8

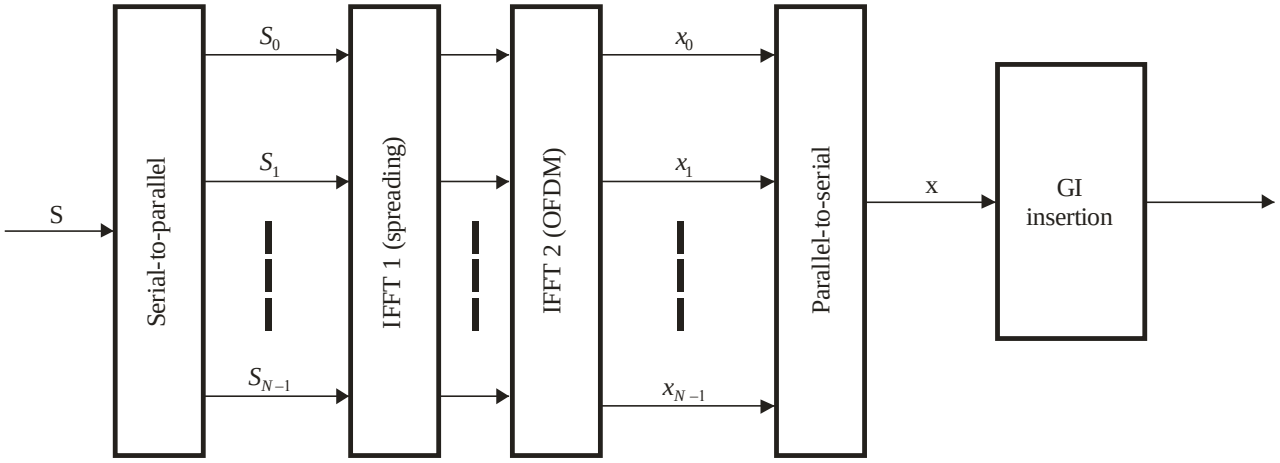
Peak reduction effect of CI-OFDM



Report S.2173-08

FIGURE 9

Configuration of CI-OFDM transmitter using FFT engines



Report S.2173-09

6.4 Power amplifier linearization: a technique to reduce the effect of PAPR

An HPA's non-linear behaviour is one of the most significant factors in determining the PAPR performance, and the linearization of the power amplifier is the most effective method to be used in overcoming this difficulty. Linearization of the power amplifier is more important to multi-carrier systems, because they are more sensitive to non-linear distortions due to their comparatively high PAPR. Over the years, a number of linearization technologies have been developed and predistortion has been the most common approach utilised by new systems today. Although predistortion techniques are not a direct way of reducing PAPR, they can be an effective solution to reduce the performance degradation due to PAPR.

The essence of predistortion is to precede an HPA with a non-linearity that is the inverse of the HPA's non-linear behaviour. This non-linear device can be designed digitally using a mapping predistorter (a look-up table: LUT) or a polynomial function based on Cartesian or polar representation. All implementations of predistortion should be adaptive to changes in operating conditions (supply voltage, temperature and antenna loading). Therefore, one of the most important factors in an adaptive predictor is fast convergence rate. The mapping predistorter has the advantage of performing any order of non-linearity and any modulation technique. On the other hand, a major drawback with the mapping predistorter is the large size of the look-up table to obtain an acceptable accuracy, which results in long adaptation time. A polynomial-based predistortion with adaptive capability can be a solution. Another limitation to the predistortion approach is that the HPA must still be operated with a substantial back-off in order for predistortion to be effective.

6.5 Direct spectrum division transmission: a technique to reduce the effect of PAPR

In general, multi-carrier transmission is well-known to increase the effects of PAPR and thus demands that the amplifiers are backed-off compared to single carrier transmission. This is considered to be a big drawback of multi-carrier transmission. Direct spectrum division transmission (DSDT) is one of the promising options that allow a reduction in the effects of PAPR to overcome this drawback [33].

FIGURE 10

Direct spectrum division transmission (DSDT) via satellite

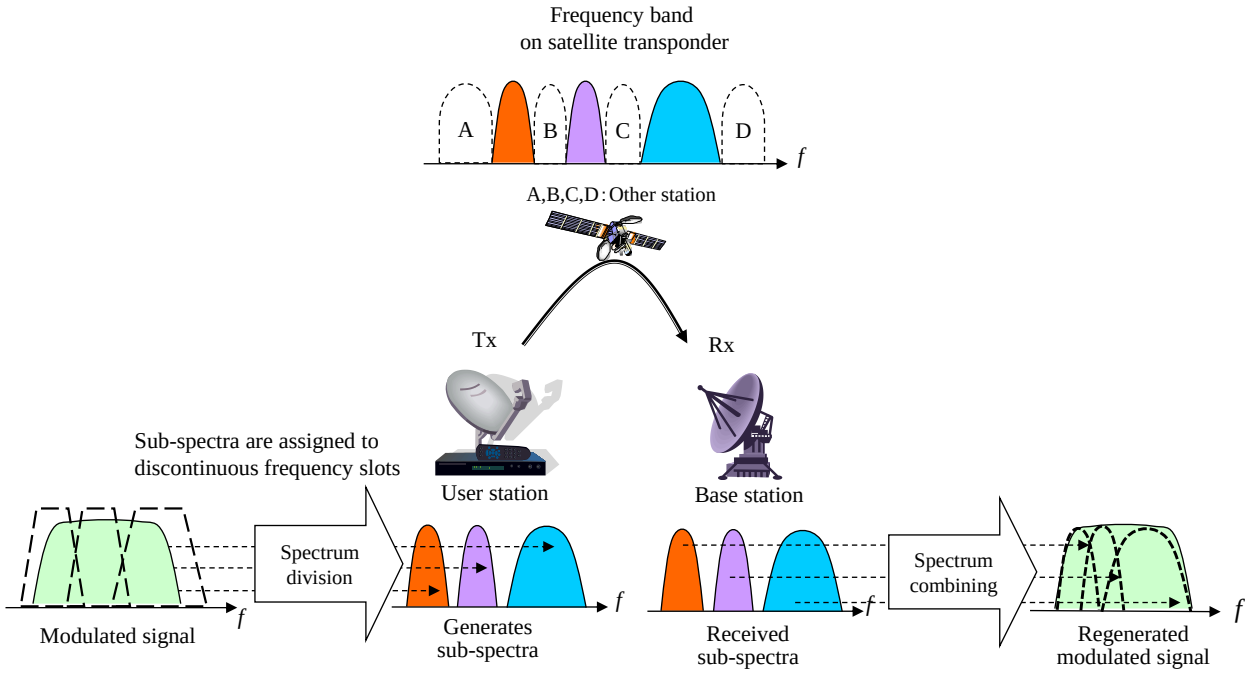


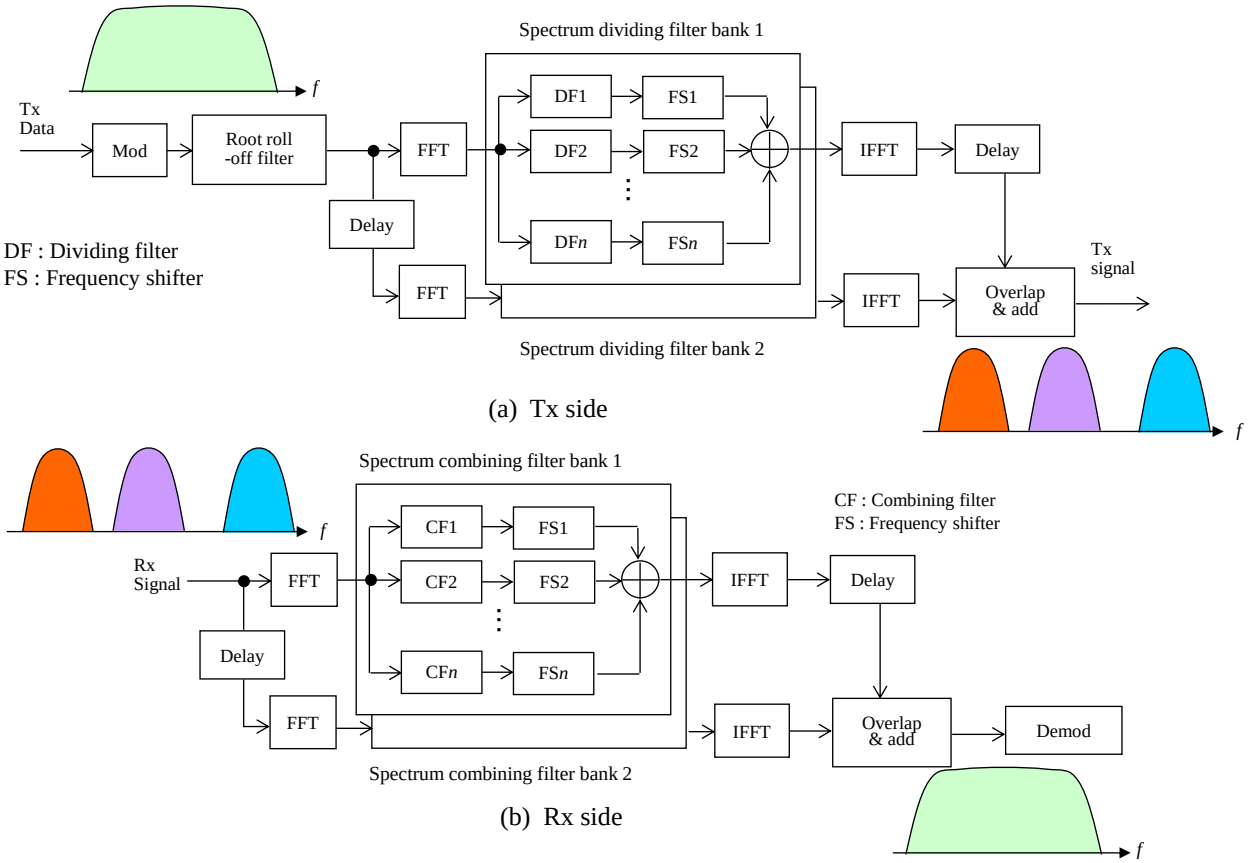
Figure 10 shows the concept of communications by the DSDT via satellite. In the transmitter (Tx) side, the single-carrier modulated signal is converted into the frequency domain. Its spectrum is divided into multiple sub-spectra and assigned to unused discontinuous frequency slots. The multiple sub-spectra are converted into the time domain signal and transmitted. In the receiver (Rx) side, the received signal is converted into a frequency domain signal, and the sub-spectra are shifted and combined. This retrieves the Tx signal.

Since the proposed technique divides the single carrier signal into multiple sub-spectra after modulation, it is different from other multi-carrier based transmission techniques that modulate signals carrier by carrier. This enables the PAPR to be suppressed. In addition, since the spectrum can be flexibly divided to fit the unused frequency slots, availability of satellite transponder would be improved.

Figure 11(a) shows the spectrum division circuit in the Tx side. The modulated signal is converted into the frequency domain by Fast Fourier transformation (FFT), and its spectrum is divided into multiple sub-spectra and assigned to discontinuous frequency slots by a spectrum division filter bank. To utilize the transition bands of the sub-spectra as pass bands after bandwidth combination in the Rx side, all dividing filters have the same transition bandwidth and the –3 dB frequency of each sub-spectrum is adjusted to match that of the adjacent one. The multiple sub-spectra are converted into a time domain signal by inverse FFT (IFFT). Figure 11(b) shows the spectrum combining circuit in the Rx side. The received sub-spectra are combined by using a spectrum combining filter bank. Root roll-off filters are used as the combining filters. Since both sub-spectra and combining filters have root roll-off characteristics, each sub-spectrum after the spectrum combining filter offers full roll-off characteristics.

FIGURE 11

Circuit configuration and spectrum image of the DSDT



7 Channel coding techniques

7.1 Channel coding

Digital data transmission channels are subject to various impairments including noise, distortion, and interference. Therefore, the output of a channel differs from its input and decoding errors can result from impaired transmission. Although we can achieve the required performance by simply by using sufficient power, in many cases error-control techniques can provide the required accuracy with less energy.

Error control techniques can be classified into two main categories: forward error correction (FEC) and Automatic Repeat reQuest (ARQ). In FEC, redundant bits are added to the information data at the transmitting end and, by utilizing that redundancy, an attempt is made to correct any errors that may have been contributed by the channel at the receiving end. On the other hand, in ARQ, the receiving terminal does not attempt to correct the errors, but attempts to detect them. In the event that an error is detected, the receiver simply requests the re-transmission of the data.

Contrasted with source coding – where source data is compressed by removing redundant information – FEC coding, which is often called channel coding adds redundant information to aid

in error correction. The purpose of channel coding is to average the effects of channel impairments over several transmitted bits. In order to achieve this purpose, the channel encoder transforms the information sequence into a longer binary-coded sequence by adding redundant or parity check symbols.

Channel coding can be classified into two types: block coding and convolutional coding. The channel encoder adds redundant bits according to the coding method used. The channel decoder transforms the received sequence into an estimated information sequence. Since channel impairments may cause some decoding errors, the channel decoder must be implemented in a way that minimizes the probability of decoding error. This is done by using statistical information about the aforementioned channel impairments.

With block codes, algebraic properties are very important for constructing good classes of codes and developing decoding algorithms. Therefore, the theory on decoding techniques for block codes is quite well developed to the point where decoding techniques based on algebraic calculation are usually processed quickly, and thus are computationally efficient. However, most conventional (i.e. non-iterative decoding) techniques for block codes require the incorporation of hard decisions, and it is generally difficult to incorporate soft decisions. However, LDPC codes, being a special class of block codes, are an efficient means of providing performance approximating the Shannon limit, by using iterative soft-decision decoding.

Channel codes can also be classified as either systematic or non-systematic codes. In a systematic code, the information bits – called systematic bits – compose part of the codeword, with the rest of the code bits being called parity bits. Any linear block code can be put into systematic form.

Compared with block codes, convolutional codes have a more probabilistic flavor and are difficult to analyze mathematically. Convolutional codes are decoded by using probabilistic methods such as the Viterbi algorithm [34]. The main advantage of such decoders is that they can incorporate soft decisions with little increase in decoder complexity, and as a result, can achieve better performance than when hard decisions are incorporated.

Multi-stage coding techniques may be applied by concatenating multiple coding schemes, when a large amount of coding gain is required. The original intention was to create a longer code that could attain better performance, but whose decoding can be performed by relatively simple component codes. The details on this multi-stage concatenated coding scheme are described in § 6.2. Recently iterative decoding techniques along with the associated coding schemes have been developed to achieve performance approximating the Shannon limit. More detailed discussions on iterative decoding will follow in §§ 6.3 and 6.4.

7.2 Concatenated codes

Concatenated codes are the combination of two or more encodings of a message-word, either in serial or in parallel.

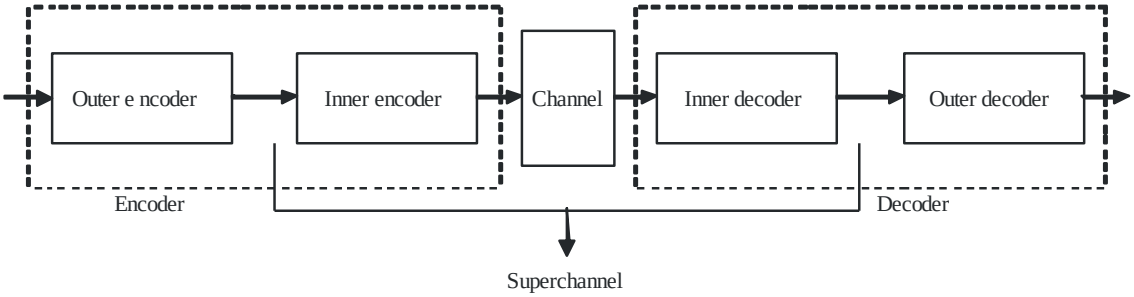
7.2.1 Single-level concatenated codes

Concatenated codes were originally proposed by Forney in [35] as the serial concatenation of two channel codes, with one code being the *outer code* and the other, the *inner code*. Concatenated codes were proposed as a way of easily obtaining long codes (from smaller codes), which are required by the Shannon channel coding theorem for capacity approaching performance. Figure 12 depicts the proposed concatenated code of Forney, which is referred to as a single-level concatenation, due to the fact that there is only one concatenation. The decoding of the proposed concatenated code is to done in stages using a separate inner and outer decoder. That is, the result of the inner decoder is passed on to the outer decoder to do a second decoding. This, in effect, creates

a virtual "superchannel" between the outer encoder and decoder. In other words, the inner decoder is meant to improve the channel conditions, presenting the outer encoder with a channel that is less susceptible to errors. The inner code is typically a convolutional code and the outer code is typically a Reed-Solomon code [36], as is the case for the digital video broadcast – satellite (DVB-S) standard [37]. However, a concatenated code composed of a LDPC inner code and a BCH outer code is employed by the digital video broadcast – satellite – second generation (DVB-S2) standard [38].

FIGURE 12

Single-level concatenated code proposed by Forney



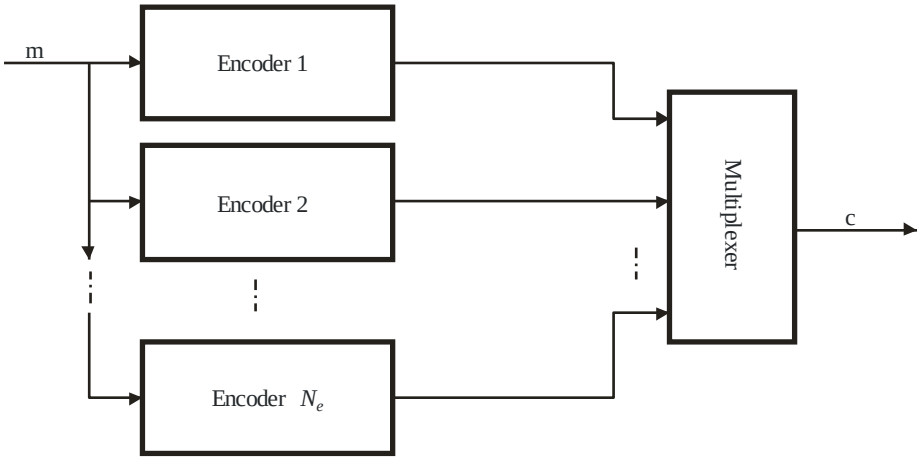
Report S.2173-12

7.2.2 Multi-level concatenated codes

Single-level concatenation is a special case of multi-level concatenation. In a multi-level concatenation there are more than two encoding and decoding stages used. Multiple stages of encoding are combined to form a multi-level concatenated code. Multi-level concatenations can provide a more flexible, more powerful concatenated code than single-level concatenations [39]. As with single-level concatenation, separating the encoding and decoding stages reduces the decoding complexity. This is because the decoding can take place on each codeword generated by each encoder independently, as opposed to decoding all codewords composing the entire concatenated code jointly. This property of concatenated codes allows for the use of iterative decoding techniques such as turbo decoding, as will be discussed in § 7.2.

FIGURE 13

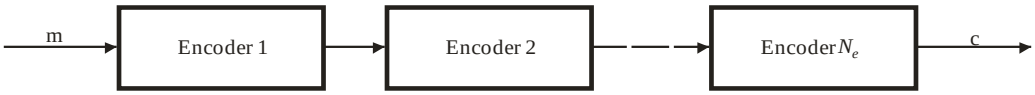
Parallel multi-level concatenated code encoder



Report S2173-13

FIGURE 14

Serial multi-level concatenated code encoder



Report S2173-14

In general, there are two types of multi-level concatenated codes: serial and parallel. In both cases N_e encoders are used to encode a message-word, **m**, to create a new codeword, **c**. In the case of parallel concatenation – as depicted in Fig. 13 – **m** is passed to N_e separate encoders and their results are multiplexed (concatenated) together to create the codeword, **c**. For both block and convolutional codes are used for more than one stage of parallel encoding, interleaving (for encoders 2 to N_e) must be done to randomize the bits of **m** in order maintain quasi-independence between the outputs of the encoders. Though, there are no interleavers included in Fig. 8, it is assumed to be done at each encoder.

In the case of serial concatenation – as depicted in Fig. 14 – **m** is passed to encoders, one at a time, with the output of each encoder being passed to the next one in the chain. Serial concatenated codes are often referred to as **product codes** in the case where only block codes are used. For serial concatenation, **m** = **M** is a matrix with N_e dimensions. For serial concatenation an interleaver must be put between each encoding stage to transpose the output of each encoder, so as to ensure that independence is maintained between each encoder output. In Fig. 14, an interleaver is assumed to be part of encoders 2 to N_e. Figure 15 demonstrates the structure of a serially concatenated block codes (i.e. product code with (n_1, k_1) and (n_2, k_2) systematic block codes). The output codeword, **c** = **C**, is a codeword matrix which is composed of the input message-matrix, **M**; a parity-check portion, $\mathbf{P}_1$, representing the encoding of **M** by Encoder 1; a parity-check portion, $\mathbf{P}_2$, representing the encoding of **M** by Encoder 2; and a parity-check portion, **P'**, representing the encoding of the bits from $\mathbf{P}_1$ and $\mathbf{P}_2$. The output of the concatenated block code encoder can also be expressed mathematically as:

$$\mathrm{C} = \mathrm{G}_1^T \, \mathrm{M} \mathrm{G}_2 \tag{2}$$

where:

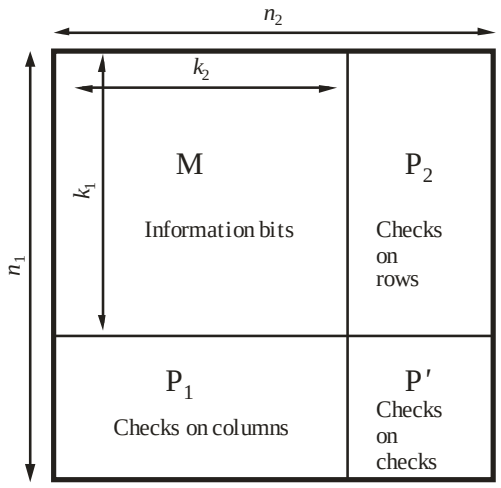
$(\cdot)^T$: is the transpose of a matrix

$\mathbf{G}_1$: is the generator matrix for Encoder 1

$\mathbf{G}_2$: is the generator matrix for Encoder 2.

FIGURE 15

Structure of serially concatenated block code



Report S2173-15

7.3 Turbo codes

7.3.1 Introduction

In 1993, a new iterative decoding technique for concatenated channel codes was introduced by Berrou *et al.*, as described in [40]. Although this new development was initially met with much scepticism, it has opened a new avenue to greatly improving the energy efficiency of communication systems through the use of channel codes. The original implementation of this technique was shown using the parallel concatenation of two distinct convolutional codes and called a convolutional turbo code (CTC). However, it is important to note that turbo coding does not refer to how the channel codes are designed, but rather to the method by which these codes are decoded. From there on, turbo decoding principles were extended to all sorts of channel codes, creating both serial and parallel versions of convolutional and linear block codes.

A turbo code (TC) is a concatenated code that is decoded using an iterative decoding process. It was determined that excellent performance could be achieved with just two serial or parallel concatenations, which is the number of concatenations used in all 3G and 4G telecommunications standards, as well as the DVB standards for satellite. At present CTCs have been implemented in many 3G telecommunications standards (e.g. UMTS, HPSA, HPSA+) and are included as part of the 4G telecommunications standard (long term evolution (LTE) and WiMAX). In some of these standards there is an option to use block turbo codes (BTC)s (e.g. WiMAX), but CTCs are still listed as the preferred code.

7.3.2 Convolutional turbo codes

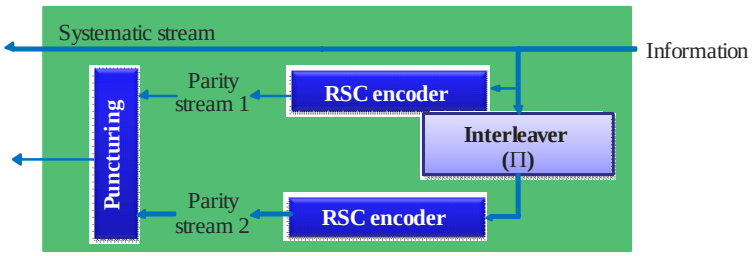
The encoder of the first CTC introduced by Berrou et al. was built using a parallel concatenation of two recursive systematic convolutional (RSC) codes and a random interleaver, inserted between them [41], as shown in Fig. 16. Two component RSC codes are used to encode the same information bits, and the outputs from the two RSC codes are then punctured and multiplexed. Generally, both RSC codes have a code rate of 1/2 with one parity and systematic output bit for every input information bit. Figure 17 shows an example of an encoder for a RSC code, having a rate of 1/2 and a constraint length, K, of 3. With this configuration, the first RSC encoder generates systematic and parity stream 1, while the second RSC encoder generates parity stream 2 by an interleaved version of the same input information. This results in overall code rate of 1/3, having one parity bit from each stream and one systematic bit. Alternating puncturing of the parity bits from each stream produces a rate 1/2 code [42].

The constraint length of convolutional codes is usually proportional to complexity of the encoder and decoder. For turbo codes, RSC component codes with relatively short K are used. The excellent performance of turbo codes – even with these simple component codes – results from the use of a long-length-interleaver between the component codes and the recursive structure of the component codes.

The purpose of the interleaver is to decorrelate the inputs to the component RSC encoders, so that the parity streams from each encoder are independent. This improves the likelihood that an error pattern will eventually be correctable by one of the two decoders. The interleaving is usually performed in a pseudorandom manner over a fixed length block.

FIGURE 16

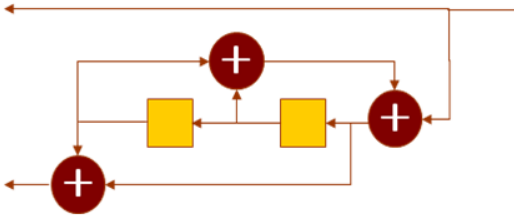
Encoder structure of a CTC



Report S.2173-16

FIGURE 17

RSC encoder in the CTC encoder

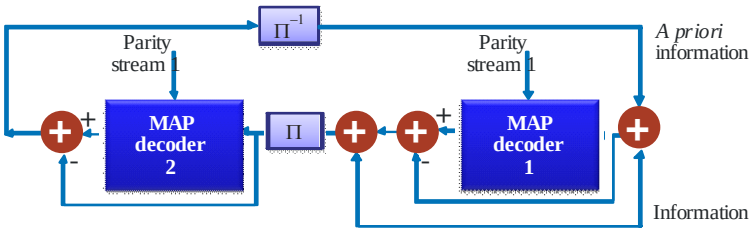


Report S2173-17

The structure of an iterative decoder for CTCs is shown in Fig. 18. There are two maximum *a posteriori* (MAP) decoders which correspond to two component RSC encoders. The MAP decoder is a soft input soft output (SISO) decoder, and it produces extrinsic information indicating reliabilities of decoded outputs. Initially the *a priori* information is zero. The difference between the output of the first decoder and the original inputted information is calculated to produce extrinsic information derived from the decoding process. The initial (channel) information is then added to the extrinsic information to produce the input information for decoder 2. Interleaving (Π) is then used to create the correct ordering of the input information for decoder 2. At the output of the second decoder, the extrinsic information is derived and de-interleaved (Π^{-1}) before being passed back to decoder 1 as *a priori* information to be used in the next iteration. This process is iterated several times.

FIGURE 18

Iterative decoder for the CTC encoder



Compared to classical CTCs, a slightly different type of codes, named duo-binary turbo codes are specified in DVB-RCS [43] and the IEEE WiMAX standard. For the turbo code with information block size of n symbols, the encoder generates $3n$ symbols with a code rate of 1/3. Referring to Fig. 19, for an information symbol with two bits, [A B], the first RSC code generates a parity symbol with two bits [W1 Y1] and the second RSC generates the other parity symbol for interleaved version of [A B] with two bits [W2 Y2].

7.3.3 Block turbo codes

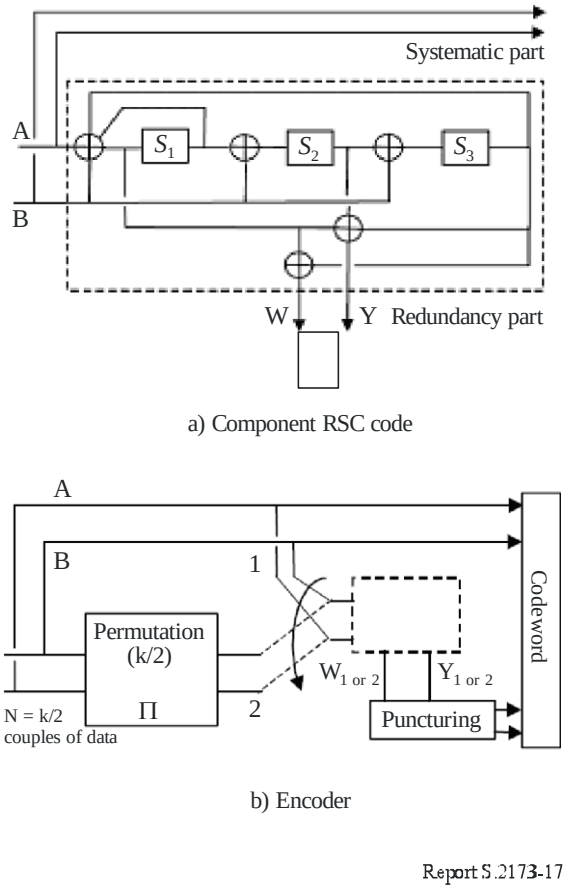
Pyndiah *et al.*, first introduced BTCs, which are product codes combined with iterative decoding algorithms [44]. After their introduction and detailed investigation [45], these codes have proven to be applicable to many areas, including WiMAX systems. For a BTC, serial concatenation with block interleaving is normally used. Although parallel concatenation is also possible for block turbo codes, serial concatenation is usually preferred, leading to the so-called product codes.

Figure 15 shows a two dimensional product code using a (n_1, k_1) block code and a (n_2, k_2) block code as component codes. For the product code of Fig. 15, the check digits on check digits are the same whether the checks on rows or on columns are computed first. It is theoretically possible to construct m-dimensional product codes for m larger than 2. Therefore, in the m dimensional product codes we have ($n_1 \times n_2 \times \ldots nm$) bits (symbols) of encoded block [46].

Block codes modified so as to increase their minimum distance are often used as component codes; this results in a large increase in the minimum distance of the product code at the expense of a small decrease in its rate, and hence a small increase in the required bandwidth [44]. Such modified codes include expurgated and extended codes [47].

FIGURE 19

Schematic diagram of duo-binary turbo codes [43]



Increasing the dimension of product codes, which is an extension of making 2-dimensional product codes from linear block codes, results in an increase in the minimum distance of the product code. High-dimensional schemes often produce better performance than low dimensional schemes with a similar code rate or data frame length. Much more complex component codes are needed for a 2-dimensional code in order to produce both the same code rate and the same frame length as those of a 3-dimensional code. Assuming that a high-dimensional BTC can produce a performance similar to that of a 2-dimensional code, a less complex component code can be employed in designing the high-dimensional BTC at the expense of additional decoding iterations. This can be an important advantage for high-dimensional schemes because for BTCs, the parallel decoding of codeword sequences is possible in each dimension. Therefore, the overall decoding speed of the high-dimensional code can be faster than that of the 2-dimensional code, if properly designed.

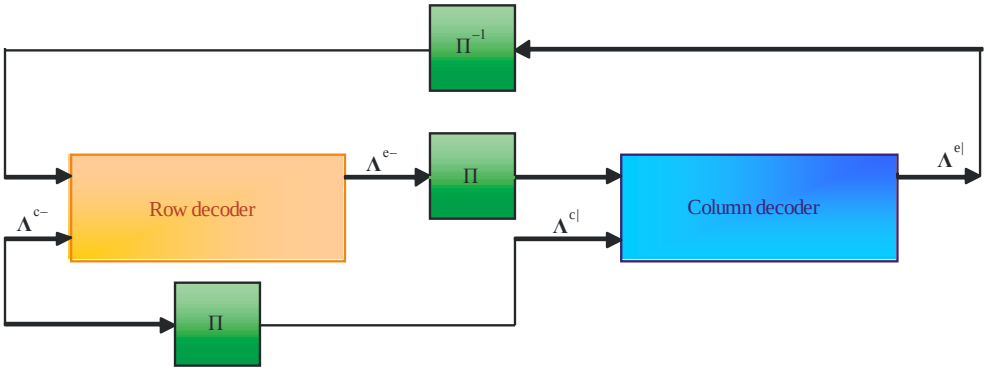
Most soft-output decoding algorithms for BTCs are based on algebraic decoding methods, which inherently are not capable of being modified for soft-decision decoding and usually have to be resorted to by the decoder several times in order to estimate soft outputs. For example, the Chase-Pyndiah Algorithm (CPA) [44, 45] modifies the Chase algorithm [48] to compute soft outputs. Trellis-based soft output decoding can also be applied to BTCs by using the fact that all linear block codes can be represented using trellis structures [46], thus allowing the use of iterative decoding using the soft output Viterbi algorithm (SOVA) [49] and MAP algorithm [50].

The decoding structure for a 2-dimensional BTC – shown in Fig. 20 – is very similar to the one in described by the CTC decoder in Fig. 18. Since for BTCs the codeword is represented by a codeword matrix, **C**, the channel information is also a matrix, $\mathbf{\Lambda}^c$, representing the reliabilities or log-likelihood ratios determined based on the received samples of the transmitted symbols, **R**. Note

that depending upon whether decoding is done on the rows or the columns of **C** the notation for $\mathbf{\Lambda}^{c}$ differs. That is, when decoding takes place on the rows of **C**, the channel input, $\mathbf{\Lambda}^{c-}$, is orientated row-by-row; and when decoding takes place on the columns of **C**, the channel input, $\mathbf{\Lambda}^{c|}$, is orientation column-by-column. The decoder outputs the extrinsic information, $\mathbf{\Lambda}^{e}$, derived from the decoding operation, which is arranged row-by-row ($\mathbf{\Lambda}^{e-}$) at the output of the row decoder and column-by-column ($\mathbf{\Lambda}^{e|}$) at the output of the column decoder. A soft-input, $\mathbf{\Lambda}$, is then created by the decoder from a combination of the channel and extrinsic information and used by the decoder to make a decision, **D**. Interleavers (Π) are used to perform a matrix transpose, which re-orients the reliability information between each decoding stage. Note that de-interleaving is expressed: Π^{-1}.

FIGURE 20

Decoder for CPA of 2-dimensional BTC



Report S.2173-20

Owing to the structure of a product code (as expressed in Fig. 14), each codeword in **C** can be decoded independently. That is, all codewords encoded using the rows of the input message-word, **M** (row codewords of **C**), can be decoded independently of all codewords encoded using the columns of **M** (column codewords of **C**). Note that when a codeword is decoded, new information about each bit in that given codeword is produced. This information is used advantageously to decode another codeword which shares bits with that given codeword[7]. That is, for a BTC when codewords are decoded row-wise, new information about each bit in that codeword can be used as *a priori* information in decoding of each column codeword, and vice versa, using extrinsic information derived from column-wise decoding to decode row codewords. This is the principle used for iterative decoding of turbo codes: new information obtained from a previous decoding can be used as *a priori* information in a future decoding. In this way, the results of a previous decoding could be revisited using new results obtained from the most recent decoding of **C**. Much like completing a cross-word puzzle by filling in the "across" answers first, those choices will impact the choices made when filling in the "down" answers. As each section is revisited, entries are modified so that they make sense with the clues given. The channel samples, used to create the

[7] For clarification, review Fig. 15. In this figure it can be seen that each row codeword shares one bit with n_2 different column codewords, while each column codeword shares one bit with n_1 different row codewords.

channel information, $\mathbf{\Lambda}^c$, are the clues. The reader should note that to complete one iteration, two decodings must take place (one row-wise and one column-wise)[8].

In summary, soft output decoding is first performed for n_2 (n_1, k_1) component codes, and then interleaving is performed on the soft outputs so that they can be used as extrinsic information to soft output decoding process of n_1 (n_2, k_2) codes. As for the decoding of CTCs, these processes are re-iterated.

7.3.4 Methods for decoding turbo codes

MAP (BCJR) algorithm

Iterative decoding of turbo codes essentially requires a decoding algorithm that can produce soft outputs, as the soft outputs of one component decoder become the inputs to another component decoder. The soft output decoding algorithm used for the first CTC in [40] was MAP algorithm, which is also known as BCJR algorithm [50]. The principle of BCJR decoding is that the metric associated with a particular branch in the trellis is given by the product of probabilities associated with choosing the most appropriate start-node, and following the most appropriate set of branches that arrive at the most appropriate end-node at the end of the trellis [42].

The soft output for decoded information, u_k, which is the information bit associated with the transition from time k–1 to k in the trellis, can be estimated by calculating the ratio of the probabilities that a transmitted bit, given the observation of the received vector, **y**, is 1 or 0, that is:

$$L(\hat{u}_k) = \frac{\Pr\{u_k = 1 \mid \text{observation}\}}{\Pr\{u_k = 0 \mid \text{observation}\}} \tag{3}$$

In order to reduce the complexity of computing the products of several probabilities, we often use logarithmic values of (2), which is often called log-MAP, and represented by:

$$L(\hat{u}_k) = \log \frac{\Pr\{u_k = 1 \mid \mathbf{y}\}}{\Pr\{u_k = 0 \mid \mathbf{y}\}} = \log \frac{\sum_s \sum_{s'} \gamma_{+1}\left(y_k, s_{k-1}^1, s_{k-1}^1\right) \alpha_{k-1}(s_{k-1}^1) \beta_k(s_k^1)}{\sum_s \sum_{s'} \gamma_{-1}\left(y_k, s_{k-1}^0, s_k^0\right) \alpha_{k-1}(s_{k-1}^0) \beta_k(s_k^0)} \tag{4}$$

where y_k is the received signal at time k, and s, α, β, and γ are the state index, forward path metric, backward path metric and branch path metric, respectively. Their relationships are described by Fig. 21. Note that $\text{sgn}(L(\hat{u}_k))$ is the sign of is the decoded information and $|L(\hat{u}_k)|$ is the soft output, i.e. the reliability value. The forward and backword path metrics are calculated in a recursive manner as represented by:

$$\alpha_k(s) = \frac{\sum_{s'} \sum_{i=\pm 1} \gamma_i(y_k, s', s) \alpha_{k-1}(s')}{\sum_s \sum_{s'} \sum_{j=\pm 1} \gamma_j(y_k, s', s) \alpha_{k-1}(s')} \tag{5}$$

---

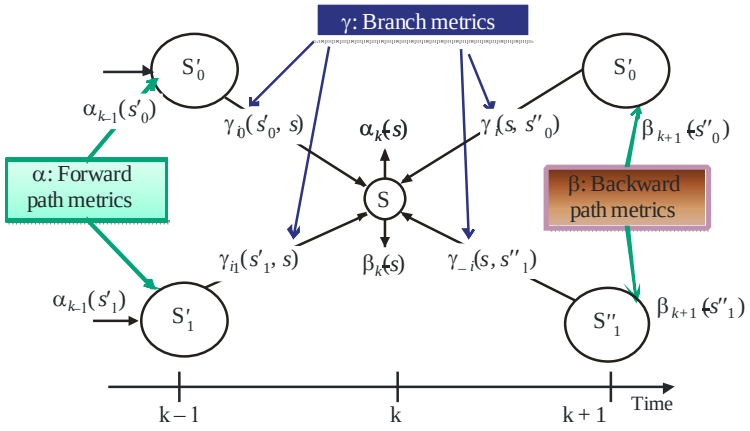
[8] Note that the number of decodings comprising one iteration is relative to the number of block codes composing that particular Turbo code.

and

$$\beta_k(s) = \frac{\sum_{s'}\sum_{i=\pm1}\gamma_i(y_{k+1}, s, s')\,\beta_{k+1}(s')}{\sum_{s}\sum_{s'}\sum_{j=\pm1}\gamma_j(y_{k+1}, s', s)\,\alpha_k(s')} \tag{6}$$

FIGURE 21

Metric values calculated for MAP algorithm

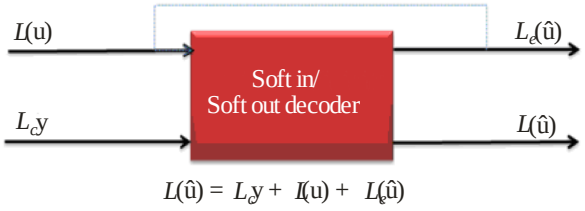


Report S.2173-21

Once the soft outputs are obtained, iterative decoding can be performed by estimating the associated extrinsic information and transferring it to the next decoding stage. Figure 22 shows the relationship between soft inputs and outputs of a SISO decoder.

FIGURE 22

SISO decoder



Report S.2173-22

Initially, the *a priori* information $L(\mathbf{u})$ is set to zero, where L_c is the channel gain. Assuming that soft outputs are available from the first decoding process of decoder 1 in Fig. 18, the *a priori* information going into the decoder 2 is the difference between the output and the input to the first decoding process. The information passed to the next component decoder, called the extrinsic information, is used as *a priori* information by the next decoding stage in conjunction with the received values, **y**, that is:

$$L_e^1(\hat{\mathbf{u}}) = L^1(\hat{\mathbf{u}}) - \left[L_c y + L^1(\mathbf{u})\right] \tag{7}$$

and

$$L_e^2(\hat{\mathbf{u}}) = L^2(\hat{\mathbf{u}}) - \left[L_c y + L^2(\mathbf{u})\right] \tag{8}$$

Max-log-MAP and SOVA

There are iterative SISO algorithms other than the MAP algorithm that can be used for the decoding of CTCs. First, by using the fact that $\log(\sum_i \exp(a_i)) \approx \max(a_i)$, at high SNR, a large portion of the summation in equation (3) can be eliminated, reducing the computational complexity of log-MAP algorithm; however this comes at the expense of some performance degradation. This is called Max-log-MAP.

Another iterative decoding algorithm is the soft output Viterbi algorithm (SOVA) which has lower complexity, but inferior performance when compared with MAP algorithm [49]. As the name indicates, this algorithm is the soft output version of the Viterbi algorithm. The basic principle of the SOVA is that it estimating the reliability, $|L(\hat{u}_k)|$, as the difference of the path metrics between an associated surviving path and contending path. This is very natural if we consider the situation when the difference of path metrics is very small or even zero. For such a situation, this means that the reliability of choosing a surviving path is very low.

SOVA-based decoding is generally inferior to the MAP-based decoding in terms of BER performance [51, 52]. In [51], the authors show that the original SOVA introduces two types of distortions in its output. They show that the performance of the SOVA can be improved by normalizing its output and eliminating the correlation between intrinsic and the extrinsic information derived in each decoding iteration. A modified updating rule can also be applied to improve performance. It was discovered that the original SOVA proposed in [49] omits some updates for ease of implementation, leading to overestimation of the reliability [52]. It was proved that the modified SOVA of [49] is equivalent to Max-log-MAP algorithm [53]. Application of SOVA-based SISO decoding for BTCs are well addressed in [46].

Chase-Pyndiah algorithm

Pyndiah *et. al.* introduced the idea of BTCs in [44, 45]. This paper included an algorithm called CPA, adapts the Chase decoding algorithm for block codes [48] by creating soft-decisions to decode BTCs using iterative decoding.

The CPA of a 2-dimensional BTC operates on the exchange of reliability messages between a row and column decoder. Each decoder uses the Chase algorithm [48], which is a reduced-complexity version of the maximum likelihood sequence decoding (MLSD) algorithm. However, the CPA creates soft-output reliabilities on the decisions made by the Chase algorithm to perform iterative decoding. The purpose of this section is not to describe the Chase algorithm, but rather how the CPA works using the Chase algorithm. For more information on the Chase algorithm, the interested reader is referred to [48].

Since the CPA is predicated on a hard-decision decoding algorithm, modifications had to be made in order to estimate the reliability of the hard-decisions made by each decoder. The reliability of a hard-decision on the entry from row *i* and column *j* of the codeword matrix, $d_{i,j}$, is found by examining the closest competing codewords $\mathbf{c}^+$ and $\mathbf{c}^-$ having different entries (1 or 0) at bit *j* (for row decoding) or bit *i* (for column decoding). Based on this, the extrinsic information is derived as:

$$\lambda^e_{i,j} = \begin{cases} \sum_{l=1, l\neq j}^{n} 2 d_{i,l} p_l \lambda_{i,l} & \text{if both } \mathbf{c}^+ \text{ and } \mathbf{c}^- \text{ exist} \\ \beta_m d_{i,j} & \text{otherwise} \end{cases} \quad (9)$$

where pl = $c^+_{i,l} - c^-_{i,l}$ and $d_{i,j}$ are antipodal (i.e. either +1 or –1; where +1 represents a bit 1 and –1 represents a bit 0); and it is possible that one or both of $\mathbf{c}^+$ and $\mathbf{c}^-$ may not exist due to the fact that the Chase algorithm operates using a subset of the 2^n possible codewords. In the event that there are not two competing codewords, the Pyndiah algorithm uses a scaling factor, β_m, which approximates the value of the extrinsic information for the m^{th} decoding. These values, which are a function of *m*, are specified in [48]. The values for β_m in [45] are approximated using computer simulation of a $(64,57)^2$ BTC (the superscript "*x*" indicates *x* dimensional BTC). A simple, yet effective alternative for approximating β_m, which does not require computer simulation, is presented in [54]. A soft-input value, $\lambda_{i,j}$, is computed by the decoder for each entry in **C**. These soft-values are used to select the codewords in each subset, the competing codewords, as well as, in computing the extrinsic information. The soft-input is computed using:

$$\lambda_{i,j} = \lambda^c_{i,j} + \alpha_m \lambda^e_{i,j} \quad (10)$$

where α_m is a scaling factor, which is either pre-defined for increasing values of *m* [45], or computed after every decoding by assuming that the extrinsic information is governed by Gaussian statistics [54]; a method that is very similar to the one used in computing the scaling factor in [51]. The above equations demonstrate how to compute the soft-input and extrinsic likelihood values for a row decoder. For the column decoder, the decoding operations are the same, except the summation in equation (9) is summed over *i* instead of *j*.

7.4 Low density parity check codes

7.4.1 Introduction

Low-density parity-check (LDPC) codes are a special class of linear block codes (LBC's) which are characterized by a "sparse parity-check matrix". Originally proposed by Gallager in his 1960 doctoral dissertation (published by the MIT Press in 1963 [55]), except for brief attention from Tanner in 1981 [56], LDPC codes were long forgotten until their rediscovery in the mid-1990s [57]. Much like TCs, greater interest in LDPC codes arose from new leaps in signal processing, resulting from enhanced computer processing and memory. When originally proposed by Gallager, digital signal processing was still very limited, and the use of his proposed soft-decision decoding algorithm was seemingly impossible, given the current state of technology.

LDPC codes are seen as the chief rival of TCs. The best performing LDPC codes have been shown to have nearly the same performance as the best TCs; however their major strength lies in the way that they are decoded. The message passing algorithm, which has come to be called the belief propagation algorithm (BPA) is very computationally efficient. The operations that compute and

update soft-information are parallelizable, and as a result, the BPA has much better computational simplicity than any decoding algorithms used for TCs. However, it should be noted that, in implementation, LDPC codes have significantly greater complexity in memory and routing than TCs [58].

Due to their sparse parity check matrices, LDPC codes were originally difficult to encode. This is because their generator matrices are non-sparse and thus required several computations to generate codewords. However, new LDPC encoding techniques allow for efficient encoding of LDPC codes, while not sacrificing their error-performance. One of the main drawbacks to using LDPC codes is that they require large block lengths (> 10 000) to attain error-performance near that or better than TCs. This may limit the type of real-world communication systems that LDPC codes are used for.

While TCs have been implemented in several standards, LDPC codes have taken time to gain real traction. One of the main reasons could be that it is generally more difficult to design rate-compatible LDPC codes than rate-compatible TCs. Rate-compatible codes produce multiple code rates from a single encoder and decoder. The importance of rate compatible codes is becoming great since most of the next generation wireless systems identify ACM as a mandatory requirement. However, rate-compatible LDPC have been included in the 3GPP2 standard [59] and have been shown to have excellent performance, while enabling flexible rate-selection [60]. Additionally,

LDPC codes are beginning to be included in many modern standards as an alternative to TCs. As a result, their deployment will more likely depend on the preference of the manufacturer. LDPC codes are the channel code of choice for the DVB-S2 standard [38] and are standardized in the 802.11n and 802.16e standards.

7.4.2 Description

A LDPC code is a LBC characterized by a sparse parity-check matrix. The parity-check matrix, **H**, of a LBC contains all vectors defining the null space of its generator matrix, **G**. The null space of **G** is the set of vectors that lie in the space perpendicular to the vector-space created by **G**. So **H** is a $(n-k) \times n$ binary matrix with the following property:

$$\mathrm{cH}^T = \mathrm{m}(\mathrm{GH}^T) = \mathrm{m}(0) = 0 \tag{11}$$

where **c** is a codeword vector generated by **G**; and **0** is a zero matrix (i.e. a matrix with all zero entries). This set of vectors can be used to verify whether a received codeword is valid. If the result of $\mathbf{cH}^{\mathrm{T}} \neq \mathbf{0}$, then c is not a valid codeword of **H**. Note that being a LBC, the rate of a LDPC code is $R_c = k/n$. A parity-check matrix is called low-density if there are far fewer '1' entries in the matrix than '0' entries. This type of matrix is referred to as sparse. This means that the number of codeword bits involved in one equation generated by the multiplication of **c** and a vector of **H** is very limited. There are two types of LDPC codes: *regular* and *irregular*. The distinction between a regular and irregular code has to do with the parity-check matrix of an LDPC code. A regular LDPC code is a LBC whose parity-check matrix contains exactly ω_c 1's in each column and exactly $\omega_r = \omega_c\ (n/(n-k))$ 1's in each row, where $\omega_r << (n-k)$. Accordingly, an irregular LDPC code is a LBC that does not have an equal number of 1's in either the rows or columns of it parity-check matrix.

7.4.3 Graphical representation of LDPC matrices

In [56], Tanner considered LDPC codes and demonstrated how they may be represented using a bi-partite graph. These graphs, which can be used in describing the LDPC decoding algorithm, have come to be known as *Tanner graphs*. A Tanner graph gives a diagrammatic representation of a LDPC code by showing which message-bits combine to create parity bits. This graph is composed

of nodes (memory cells) and vertices (connections). The nodes of the graph accept and transmit messages through vertices that connect them to neighbouring nodes. These neighbour nodes have the same functionality. Much like a neural network, nodes exchange information with other nodes via the links that connect them to form a decision on the information that they process as a whole. In a Tanner graph there are two types of nodes: *variable* and *check*. Variable nodes represent the codebits of each codeword. These nodes represent the entries in the columns of **H**. Check nodes represent the parity-check equations formed by the vectors of **G**. These nodes are represented by the rows of **H**. An entry in **H**, $h_{i,j}$ represents a connection or vertex between check node i and variable node j; where $h_{i,j} = 1$ represents a connection and $h_{i,j} = 0$ represents that there is no connection. All column entries in a row of **H** correspond to the bits involved in a parity-check equation, of which there are $(n-k)$; one for each parity bit.

Consider for example, a simple (10,5) LDPC code with a parity-check matrix:

$$\mathrm{H} = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix} \tag{12}$$

The parity-check matrix has characteristics such that $\omega_c = 2$ and $\omega_r = (n/(n-k))\ \omega_c = 4$ for all columns and rows respectively. This means that this is a regular LDPC code. Notice that 40% of the matrix is filled with non-zero entries, making it relatively sparse. Better LDPC codes are larger and tend to have sparser **H** matrices. Figure 23 depicts the Tanner graph of the (10,5) LDPC code. Notice how each variable node, ν_i, is connected to $\omega_c = 2$ check nodes, and that each check node, ρ_j, is connected to $\omega_r = 4$ variable nodes. If the first row of the **H** matrix is examined, it can be seen that there are 4 entries, {1, 2, 3, 4}. Observe that ρ_1 is also connected to ν_1, ν_2, ν_3 and ν_4. Note that for the first row of **H**, the operation:

$$\mathrm{ch}_1^T = \sum_{j=1}^{N} \oplus c_j h_{1,j} = \upsilon_1 \oplus \upsilon_2 \oplus \upsilon_3 \oplus \upsilon_4 = 0 \tag{13}$$

That is, the sum of all column entries modulo-2 in a row should be equal to zero.

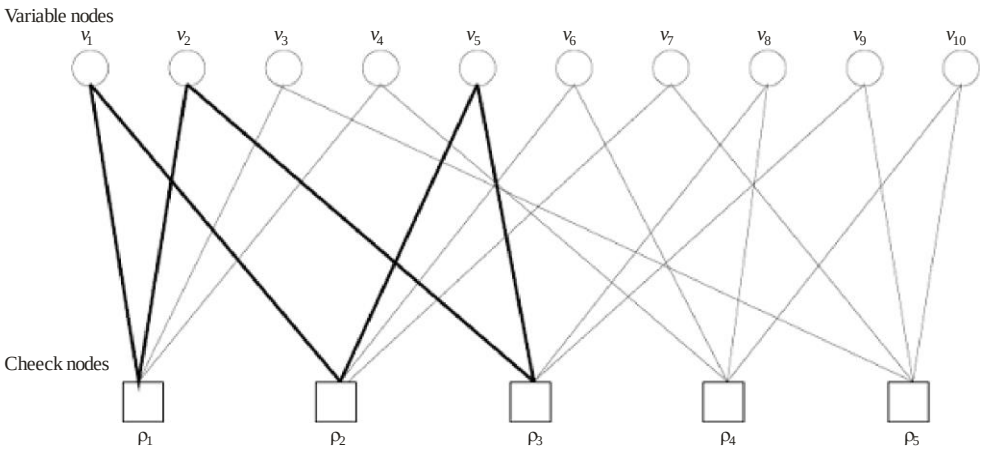
Cycles. A cycle (or loop) of length δ is a path that returns to the vertex at which it began by using δ unique vertices for each link in that path. This means that the shortest possible path taken to create a cycle is 4, for example, as depicted in Fig. 24. In Figure 23 a cycle with the minimum cycle length of 6 is depicted by bold lines[9]. The girth, γ, of a Tanner graph is the minimum cycle length of the graph. Therefore, for the LDPC code depicted in Fig. 23, $\gamma = 6$.

The girth and cycles of a LDPC code are relevant to the iterative decoding algorithm used to decode LDPC codes. This is because short cycles degrade the performance of the iterative decoding algorithm used to decode LDPC codes. Thus, codes with very large girth are desired.

[9] The cycle starting at ν_1 is [$\nu_1 \rightarrow \rho_2$, $\rho_2 \rightarrow \nu_5$, $\nu_5 \rightarrow \rho_3$, $\rho_3 \rightarrow \nu_2$, $\nu_2 \rightarrow \rho_1$, $\rho_1 \rightarrow \nu_1$].

FIGURE 23

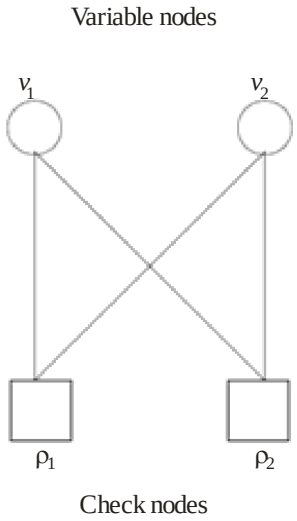
Tanner Graph for (10,5) LDPC code



Report S.2173-23

FIGURE 24

A cycle of length 4 (minimum possible cycle length)



Report S.2173-24

7.4.4 Decoding LDPC codes: belief propagation

In his doctoral dissertation, Gallager also included a near-optimal, iterative, soft-decision decoding algorithm. Though at the time, the algorithm was impossible to implement due to the requirement of advanced signal processing techniques. As a result, his contribution was forgotten, only to be rediscovered independently decades later by several different researchers in a variety of different fields of research [61]. In fact, in [61] McEliece *et al.*, discuss how Gallager's iterative message passing algorithm and the turbo decoding algorithm developed by Berrou *et al.*, are both instances of the BPA developed by Pearl [62]. In this Report we refer to the LDPC iterative decoding algorithm as it is most commonly referred to in literature: BPA.

The BPA is based on the exchange of "messages" between the variable nodes and check nodes in the Tanner graph. The nodes of the Tanner graph take the messages that they receive from all other nodes to which they are connected and use them to update information that they store on the

likelihood of a particular decision made by a variable or check node. For each variable node, the likelihood of a particular bit – represented by that particular variable node – is computed:

$$\lambda_j = \ln\left(\frac{\Pr\left[c_j = 1|\mathrm{r}\right]}{\Pr\left[c_j = 0|\mathrm{r}\right]}\right) \tag{14}$$

where λ_j is defined as the log-likelihood ratio (LLR) that the j^{th} bit in a codeword is 1 or 0, given that the received samples representing the entire codeword.

This equation can be broken down into two components: one representing the original channel information and one representing the check node equations defined by **H**[10]:

$$\lambda_j = \lambda_j^c + \sum_{l \in \xi_j} \mu_l \tag{15}$$

where λ_j^c represents the LLR, conditioned on the channel sample, r_j, which is the channel sample in which code bit c_j is transmitted; $\{\mu_l\}$ are the LLRs of the parity-check equation, $\{\rho_l\}$, which are from the set, ε_j, of check nodes connected to ν_j.

Figure 25 shows the operation of the BPA for one iteration. The algorithm has an initialization phase and an iteration phase, where one iteration is composed of two half iterations. In the initialization phase the variable nodes are provided with channel LLRs, λ_j^c, from the channel. No information is provided by the check nodes as there has not yet been a chance to evaluate the parity-check LLRs, therefore $\mu_i = 0$ for all i. In the first half iteration the variable nodes send their messages, λ_j, to the check nodes to which they are connected, as depicted in Fig. 25(a). The check nodes then compute the LLRs for each parity-check equations, μ_i for all i. The computation of each μ_i for the $(\mathrm{m}+1)^{\text{th}}$ iteration is done using:

$$\mu_i^{(m+1)} = -\left(\prod_{l \in \xi_i} \operatorname{sgn}\left(\lambda_l^{(m)} - \mu_i^{(m)}\right)\right) f\left(\sum_{l \in \xi_i} f\left(\left|\lambda_l^{(m)} - \mu_i^{(m)}\right|\right)\right) \tag{16}$$

where m is the current number of iteration used in decoding; ζi is the set of variable nodes that ρ_i is connected to; and where:

$$f(x) = -\ln\left(\tanh\left(-\frac{x}{2}\right)\right) \tag{17}$$

and

$$\operatorname{sgn}(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ -1 & \text{otherwise} \end{cases} \tag{18}$$

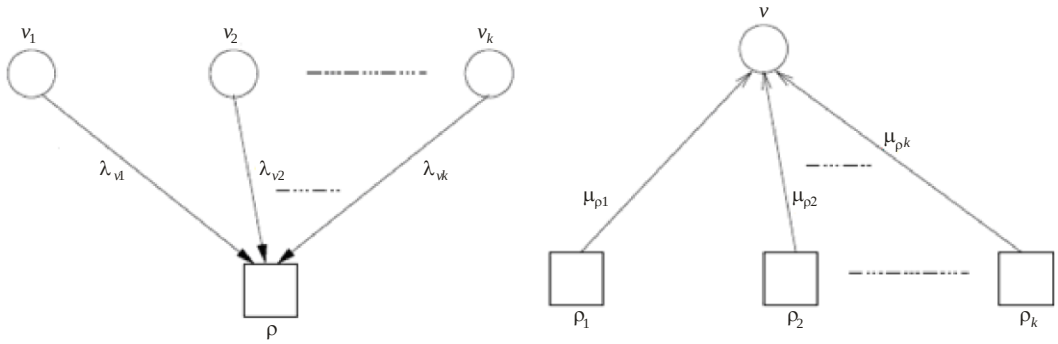
Notice that the μ_i from iteration m is subtracted from each λ_l from the same iteration. This is done to remove the contribution of μ_i from λ_l, which was added in the computation of λ_l in the previous

[10] For more information on log likelihood ratios of LDPC codes the reader is referred to the tutorial by J. R. Barry [63].

iteration. The performance of the iterative BPA is related to how well information is exchanged between independent sources. When old information is re-used it biases the decoder to repeatedly make the same decision. This is because re-using old information effectively creates a cycle in the Tanner graph. Cycles reduce the ability of the BPA to converge on the decision giving the optimal a posteriori probability (APP). In fact, in the case where the Tanner graph is cycle-free, the BPA always converges to the optimal APP decision after a finite number of iterations. It has been shown that the BPA gives approximately the APP solution for large enough LDPC codes with large girth [63] and as a result, can obtain near-optimal performance. That is, it takes a long time to bias the decoder decision. In re-using the information from a previous decision, a length-2 cycle is created, which is even worse than the shortest cycle that can exist for a LDPC code. As a result, in practise, all new parity-check LLRs are computed by subtracting μ_i from the previous iteration.

FIGURE 25

Operation of the BPA for one iteration



a) First ½ iteration of BPA b) Second ½ iteration of BPA

Report S2173-25

In the second half iteration the check nodes send their messages, μ_i, to all variable nodes that they are connected to, as depicted in Fig. 25(b). Having received the messages from each check node, the variable nodes compute:

$$\lambda_j^{(m+1)} = \lambda_j^c + \sum_{l \in \xi_j} \mu_l^{(m+1)} \tag{19}$$

to find the new LLR of each codebit, for all j. This process continues until either all parity-check equations are satisfied (i.e. $\mathbf{cH}^T = \mathbf{0}$), or the maximum number of iterations (as defined in implementation) is exceeded.

8 Link rate adaptation

Link rate adaptation is a technique that can be used by a communications system to transmit at data rates that closely match the instantaneous capacity of the channel. This method is usually accomplished by selecting the instantaneous modulation and channel coding combination that matches the error-free capacity of the channel based on the carrier-to-noise ratio (CNR) measured at the receiver. This technique is particularly well-suited for communication networks that are subject

to slow-fading. That is, where the channel is quasi-static for several packet transmissions. This requires the communications system to adapt less often and results in less overhead. This technique has become widely used, being implemented in several 3rd and 4th generation terrestrial standards, access standards, as well as, satellite standards: UMTS, HSPA, HSPA+; 802.16e Mobile WiMAX 3GPP LTE, 802.11n (Wi-Fi), DVB-S2, DVB-RCS.

8.1 Constant coding and modulation

Although, constant coding and modulation (CCM) is not a link rate adaption technique, it is still worth discussing. CCM is a special case of link rate adaptation techniques, where the rate is set once and beforehand to meet the worst-case scenario implementation requirements. That is, the rate is chosen such that it at least meets the bit-rate requirements of the services it will be providing for the given environment in which it is implemented. This method does not make efficient use of spectrum, as it continues to transmit at the minimum rate even when the channel could support higher data-rates. An analogy for CCM is driving an automobile at 20 km/h at all times, in case one drives into a school zone with the 20 km/h speed limit.

8.2 Adaptive coding and modulation

ACM is a technique whereby a combination of channel coding schemes and modulations are selected from a pre-defined list based on instantaneous CNR estimates to optimize the communications link's utilization. In the case where the channel is time-division duplex (TDD), the link CNR on both the forward and return link are roughly the same and as a result the CNR estimation could be done by the transmitter when it is in a receiving mode. In the case where the channel is frequency-division duplex (FDD), a feedback mechanism is required to inform the transmitter of the current state of the channel. If the propagation delay over the link is large, then it is imperative that the channel have slow-fading characteristics. Otherwise, the information fed-back to the transmitter will no longer reflect the actual channel conditions. It is also imperative that the feedback mechanism be low-overhead, so that the communications system does not unnecessarily waste bandwidth.

Implementations of ACM can differ depending on how the channel codes are selected and implemented. Some schemes use a set of pre-defined coding and modulation scheme such as is the case for the DVB-S2 standard. That is, the channel coding scheme is selected from a discrete collection of channel codes. Other schemes use a single code that can be punctured to obtain a specific rate; however, care must be taken to ensure that the code chosen is punctured in such a way that a very low error-floor is maintained. Schemes with pre-defined channel codes lack flexibility, in that, if a block is received in error, it must be re-transmitted in its entirety, even though it may be possible to decode for a few extra parity bits. In order to operate the ACM mode effectively, an intelligent control mechanism is required. The control mechanism usually includes a signal quality estimator, signal quality predictor and resource (or coding and modulation scheme) allocator. Recommendation ITU-R S.1061 discusses this issue, especially for compensating rain attenuation in satellite links. Even more noteworthy is that any received information may be entirely disregarded. That is, the information received about the failed packet may be dropped. In the case where a packet fails to decode it would be better to retain the information already received about that packet and combine it with the information that will be re-transmitted. Such schemes are usually called hybrid automatic repeat request (H-ARQ) protocols [64], and these are described in more detail in the next section.

8.3 Hybrid ARQ

ARQ protocols are usually operated by the data link layer. When a packet is detected in error by an error detection scheme – such as a cyclic redundancy check (CRC) – the information about that packet is retained; and a request for re-transmission is made by the receiver. ARQ can be combined with physical layer FEC schemes and the combined scheme, where this combination is called an H-ARQ scheme. H-ARQ schemes are classified into three categories namely type-I, type-II and type-III schemes [65, 66]. A type-I scheme uses the same channel code each time a re-transmission is requested, in the event that an erroneous packet is detected. The performance of type-I H-ARQ relies heavily on the error correction capability of the channel code. This scheme is not adequate time varying channel conditions because the fixed channel code cannot manage the limited channel resource efficiently. In a type-II scheme, the original packet could be composed of only information bits or a few parity bits for error correction. The initial coding rate is best selected to match the capacity of the channel [60]. In the event that packet errors are detected, a request for re-transmission is made, but a block of incremental redundancy bits are transmitted instead of the entire packet. New incremental redundancy bits are transmitted until the packet is received error-free. As type-II H-ARQ re-transmits only the added parity blocks, it is a good solution for prevailed line of sight (LoS) conditions, such as open areas. A type-III scheme uses a complementary code as an error correction code. Its self-decodablility is advantageous for heavy shadowing conditions, such as urban areas [67].

Rate-compatible FECs (RCFEC)s can be used for all H-ARQ protocols, although it is more likely to be used for type-II and -III H-ARQ. A RCFEC is usually achieved by puncturing a specified number of parity bits from a "mother code" with low-rate to obtain a higher rate code; however, it can also be achieved through the use of codeword extension algorithms to create a lower rate code from a high-rate one.

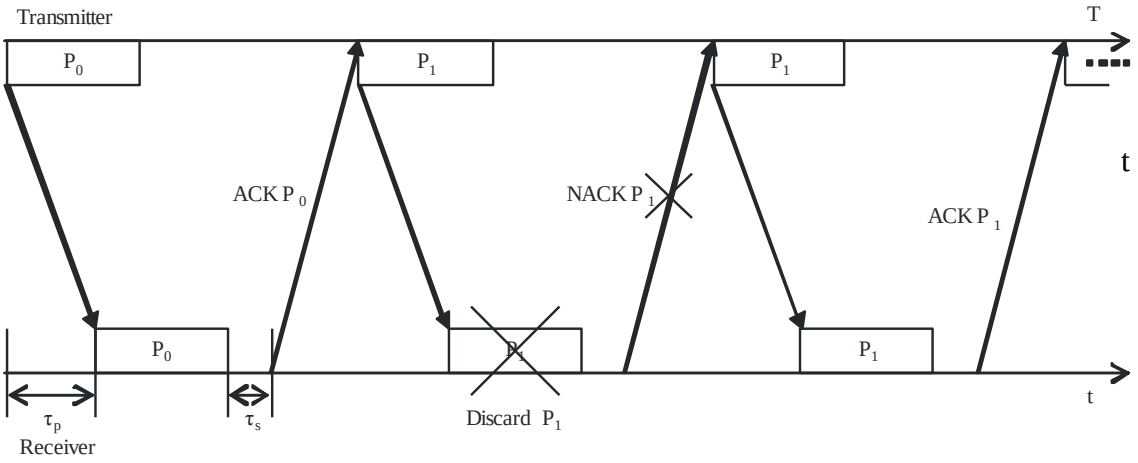
Figures 26 and 27 compare operational principles of original ARQ and type-II H-ARQ protocols. In general, there are three different ARQ protocols: stop-and-wait (S & W/ARQ), go-back-N and selective-repeat (SR/ARQ). Since the purpose of this section is not to explain how different ARQ schemes operate, the interested reader is referred to [68] for more information.

In Figures 26 and 27, the simplest ARQ protocol (S & W) is used[11]. This protocol works by having the transmitter send a packet to the receiver and waiting for the receiver to reply with an acknowledgement of successful receipt of that packet. Successful receipt is usually determined using a CRC. If the transmitter receives an acknowledgment message (ACK), then it transmits the next packet. However, if it receives a negative acknowledgement message (NACK), then it re-transmits the entire packet. The receiver will drop any information it obtained about that particular packet, and do decoding on the new packet as if it has received it for the first time. Figure 26 illustrates the operation of a S & W ARQ protocol. Note that τ_p and τ_s denote the propagation and processing delays respectively. In this example, packet P_0 is transmitted and received successfully, and the receiver acknowledges successful receipt to the transmitter. Following receipt of the ACK from the receiver, the transmitter begins sending the next packet (P_1), which is not received correctly. The receiver sends the transmitter a NACK message and discards P_1. Upon receipt of the NACK, the receiver re-transmits P_1. The receiver has successfully received P_1 after re-transmission and sends an ACK to the transmitter.

[11] It is acknowledged that the S & W/ARQ is highly inefficient and that a SR/ARQ protocol is most often used in practise. However, for the purposes of illustration it is easier to explain the concept of IR using S & W/ARQ.

FIGURE 26

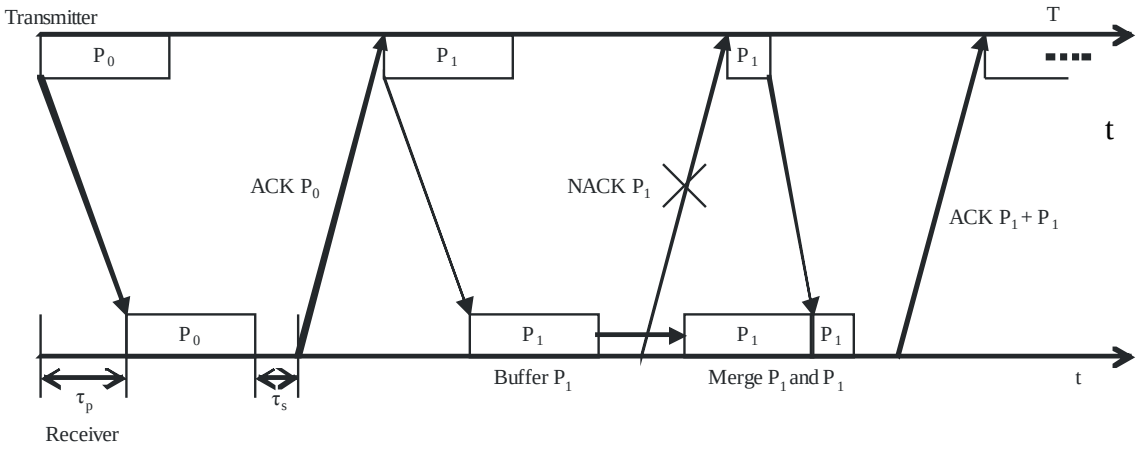
Operation for regular S & W/ARQ



Report S.2173-26

FIGURE 27

Operation for IR with S & W/ARQ



Report S.2173-27

In the regular S & W ARQ scheme, two things lead to the inefficient use of the channel[12]: the discarding of packet information already received and the re-transmission of the entire packet, should the original transmission fail. In discarding the entire packet, the communications system is, in effect, ignoring valuable information about the transmitted packet. Also, sometimes it may not be necessary to re-transmit the entire packet to successfully decode it. Sometimes successful decoding can take place at the cost of transmitting only a few extra parity bits, which represent that particular packet. In an type II S & W H-ARQ protocol, if the receiver cannot decode the packet, the transmitter will send a specific number of additional parity bits[13] and the receiver will buffer the

12 That is, aside from the transmitter having to wait for an ACK from the receiver before transmitting the next packet.

13 The number of additional parity bits can either be pre-determined or determined "on-the-fly" by estimating the number of bits necessary based on the instantaneous channel capacity. In the latter case, the NACK message may contain a demand for a specific number of additional parity bits.

unsuccessfully decoded packet. Upon receipt of these additional parity bits, the buffered packet will be merged with the parity portion and decoding will take place on this combination.

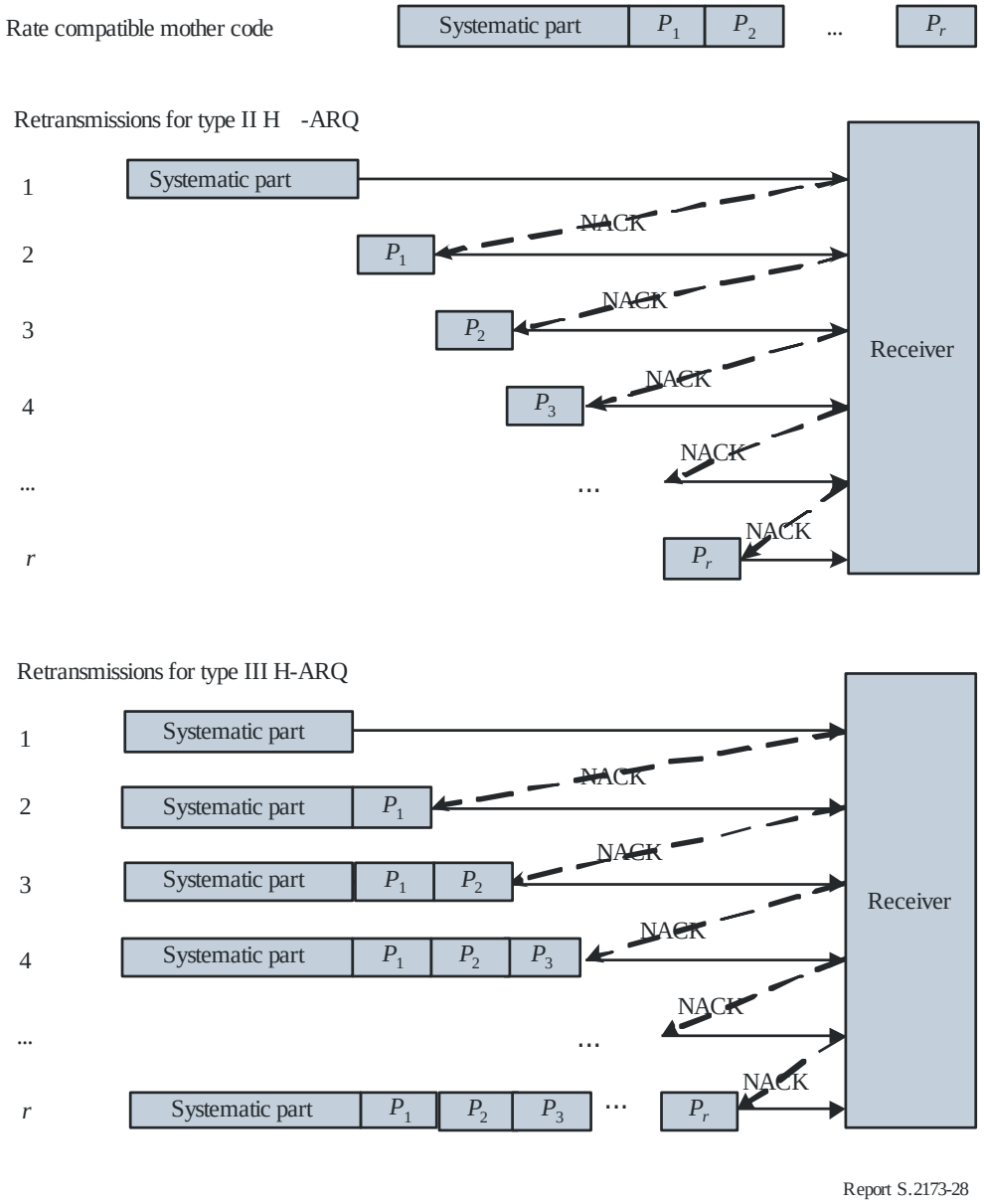
Figure 28 demonstrates the operation of a type-II S & W H-ARQ protocol. As was the case with Fig. 26, when P_0 is acknowledged, the transmitter sends P_1. The main difference with the regular S & W ARQ is that when P_1 is received incorrectly, the transmitter determines an appropriate number of parity bits, p_1, to transmit and this information is combined with the buffer copy of P_1 to do decoding. It should be noted that transmitting a small parity portion instead of an entire packet makes more efficient use of the channel. In addition, the buffering of unsuccessfully received packets enables the type-II H-ARQ scheme. In fact, if the entire packet were re-transmitted, simple packet combining would bring at least a 3 dB gain in CNR for each re-transmitted packet.

The original type-I hybrid ARQ does not combine re-transmitted packets with previously-received erroneous packet; however, the Chase combing algorithm could be applied. The basic idea of the Chase combining scheme is to combine multiple received copies of the coded packet weighted by the instantaneous SNR of each received packet prior to decoding [69]. The Chase code combining method provides time-diversity gain via a satellite channel. The difference between type-II and type-III lies in whether they re-transmit the data which have already been transmitted before. Figure 28 compares type-II and type-III H-ARQ schemes using RCFECs, which can be decomposed into a systematic portion with several parity parts, P_1 to P_r. In a type-II H-ARQ scheme, only an incremental parity part is transmitted at every re-transmission so that a longer codeword can be formed by combining it with what has already been received by the receiver. On the other hand, the type-III scheme transmits not only the incremental parity part but also the previously transmitted parts, so that the Chase combing algorithm can be employed by the receiver.

In 3G and 4G mobile terrestrial standards H-ARQ has been achieved using either the Chase code combining algorithm [69] or incremental redundancy (IR) [70]. While the Chase algorithm simply combines packets, IR allows for quick decoding of failed packets by transmitting an extra parity block – typically much smaller than the entire packet – to aid in decoding the failed packet. IR allows for easy adaptation to channel fluctuations, and as an H-ARQ protocol, makes more efficient use of the channel.

FIGURE 28

Comparison of type-II and -III H-ARQ schemes



9 Standards and transmission methods

9.1 DVB-S

The DVB-S standard was first released in 1994 and later modified in 1997 [37]. The document, as currently drafted describes the modulation and channel coding to be used for the delivery of digital video over satellite in the fixed-satellite and broadcasting-satellite services bands using the MPEG-2 codec. Over time, this standard has become the most widely used for the delivery of digital television over satellite [71]. Although this standard is to be phased out by the DVB-S2 standard, released in 2005, it continues to be used. In fact, in [71], it is stated that DVB-S will continue to be used in the short- and medium-term.

Modulation

DVB-S is designed for DTH services via satellite, which require robustness against noise and interference. As a result, the developers of the DVB-S standard decided to focus their main design objective on power rather than spectral efficiency. This means the use of a QPSK constellation as the DVB-S modulation. DVB-S does not change the modulation; however, the DVB group released a supplementary document [72] to deal with the case where the use of BPSK is desired.

Channel coding

The channel code used by the DVB-S standard is a single-level concatenated code, composed of an outer (255,239, $T = 8$) Reed-Solomon code and an inner convolutional code with rate 1/2 and a constraint length of 7. The convolutional code can be punctured to obtain the most appropriate rate and error correction for a given service. The available rates are 1/2, 2/3, 3/4, 5/6 and 7/8, giving a maximum spectral efficiency of 1.75 bit/s/Hz.

9.2 DVB-S2

Originally published in 2005, the DVB-S2 standard was created as the next evolutionary step for digital video broadcasting over satellite. The DVB group boasts that "it makes use of the latest modulation and coding techniques to deliver performance that approaches the theoretical limit for such systems" and "…it does so with a performance level that ensures we won't see a DVB-S3 for a very long time – if ever!" [71]. As a result, DVB-S2 has been adopted by several satellite broadcasters, including early adoption by BskyB in the UK and Ireland; Premiere in Germany; Sky in Italy; and DirectTV in the United States of America. It is also worth noting that Recommendation ITU-R BO.1784 specifies DVB-S2 as a suitable system for the development of a system for satellite broadcasting with flexible configuration [73].

The DVB-S2 standard uses link adaptation techniques using ACM, which combines high-order modulation with advanced coding techniques. This feature, when used with advanced video compression (e.g. MPEG-4), could enable widespread commercial HDTV services [38]. DVB-S2 was created to be backward compatible with DVB-S so to allow operators to continue using DVB-S as they update their networks; but not sacrificing the additional capacity that could be provided to newer, more advanced receivers, through the link adaptability of ACM. Table 2 from [71] shows the capacity improvement achieved by upgrading from DVB-S to DVB-S2. The example system in Table 2 suggests that the useful bit-rate gain can be as high as 36%, and the use of DVB-S2 could possibly provide of 3 extra SDTV channels or one extra HDTV channel in 27.5 MHz of bandwidth.

TABLE 2

Example comparison between DVB-S and DVB-S2 for TV broadcasting [71]

Satellite e.i.r.p. (dBW)	**51**		**53.7**	
System	DVB-S	DVB-S2	DVB-S	DVB-S2
Modulation & Code rates	QPSK 2/3	QPSK 3/4	QPSK 7/8	8-PSK 2/3
Symbol rate (Mbaud)	27.5 ($\alpha = 0.35$)	30.9 ($\alpha = 0.2$)	27.5 ($\alpha = 0.35$)	29.7 ($\alpha = 0.25$)
C/N (27.5 MHz) (dB)	5.1	5.1	7.8	7.8
Useful bitrate (Mbit/s)	**33.8**	**46 (gain = 36%)**	**44.4**	**58.8 (gain = 32%)**
Number of SDTV Programmes	7 MPEG-2 15 AVC	10 MPEG-2 21 AVC	10 MPEG-2 20 AVC	13 MPEG-2 26 AVC
Number of HDTV Programmes	1-2 MPEG-2 3-4 AVC	2 MPEG-2 5 AVC	2 MPEG-2 5 AVC	3 MPEG-2 6 AVC

Modulation

DVB-S2 provides several choices for modulation, which include QPSK, 8-PSK, 16-APSK and 32-APSK. The APSK constellations are chosen for satellite applications due to their robustness against the HPA distortion inherent in satellite transponders. This collection of constellations is meant to be used adaptively with channel coding in order to approach the capacity of the channel. The use of a 32-APSK constellation allows for a potential gain of 257% in spectral efficiency for the DVB-S2 standard (32-APSK 9/10) over the DVB-S standard (QPSK 7/8). The modulation used in DVB-S2 is shown in Fig. 3.

Channel coding

DVB-S2 uses a single-level concatenated code, composed of a BCH outer code and a LDPC inner code. The LDPC code – whose length can be either 16 200 or 64 800 bits, depending on the application chosen – performs extremely well, having performance near the Shannon Limit[14]. The BCH outer code acts as a check against the introduction of an error-floor, which is often seen in coding schemes with iterative decoders. The rate of the concatenated code is adaptive, having rates of 1/5, 1/4, 1/3, 2/5, 1/2, 3/5, 2/3, 3/4, 4/5, 5/6, 8/9 and 9/10.

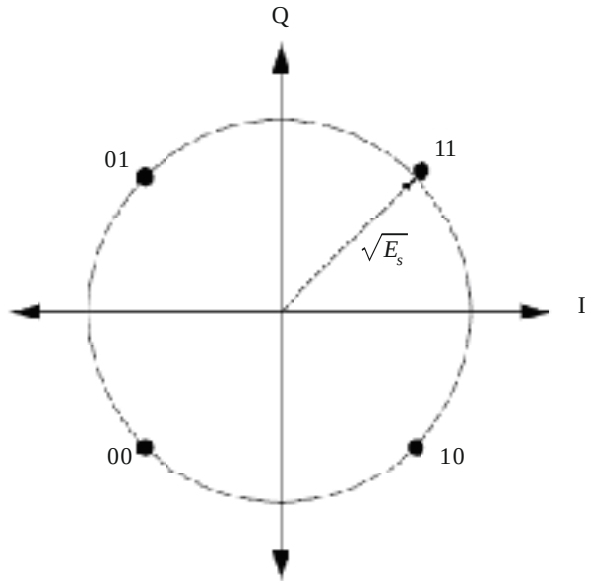
Link adaptation

The evolutionary component of DVB-S2 is the link adaptation that uses ACM. The receiving terminal is asked to give the transmitter updates on the carrier-to-noise ratio (CNR) packet-by-packet; which happens to be the length of one codeword. The receiver then uses a mechanism to select the appropriate modulation and coding combination (MODCOD) that achieves the highest quasi-error free throughput, for the given CNR. This information is relayed back using an 8-bit message. The DVB-S2 standard has left the mechanism for determining how to switch between different MODCODs up to design and implementation to encourage proprietary solutions.

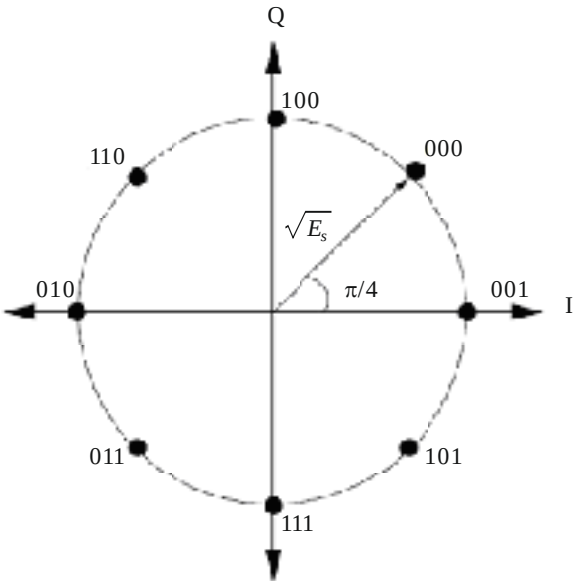
[14] Consult [74] for more information on the Shannon Limit.

FIGURE 29

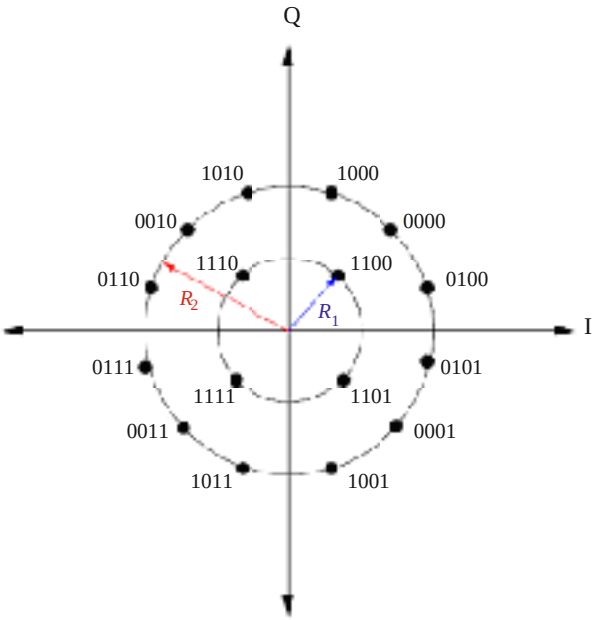
DVB-S2 modulation



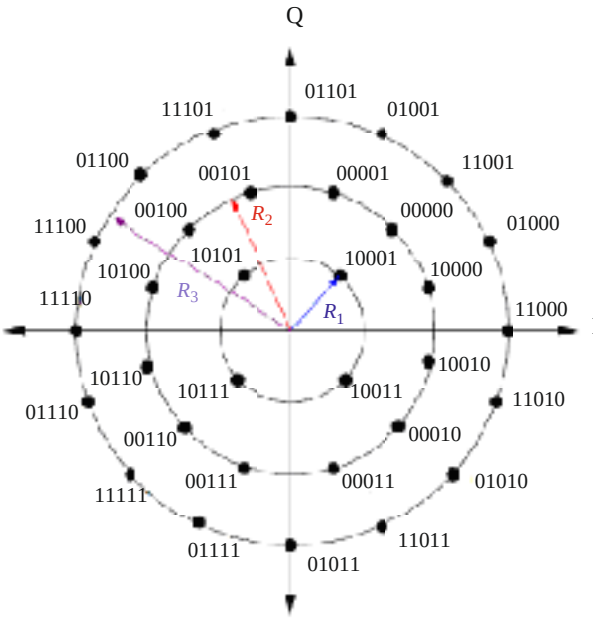
a) QPSK



b) 8-PSK



c) 16-APSK



d) 32-APSK

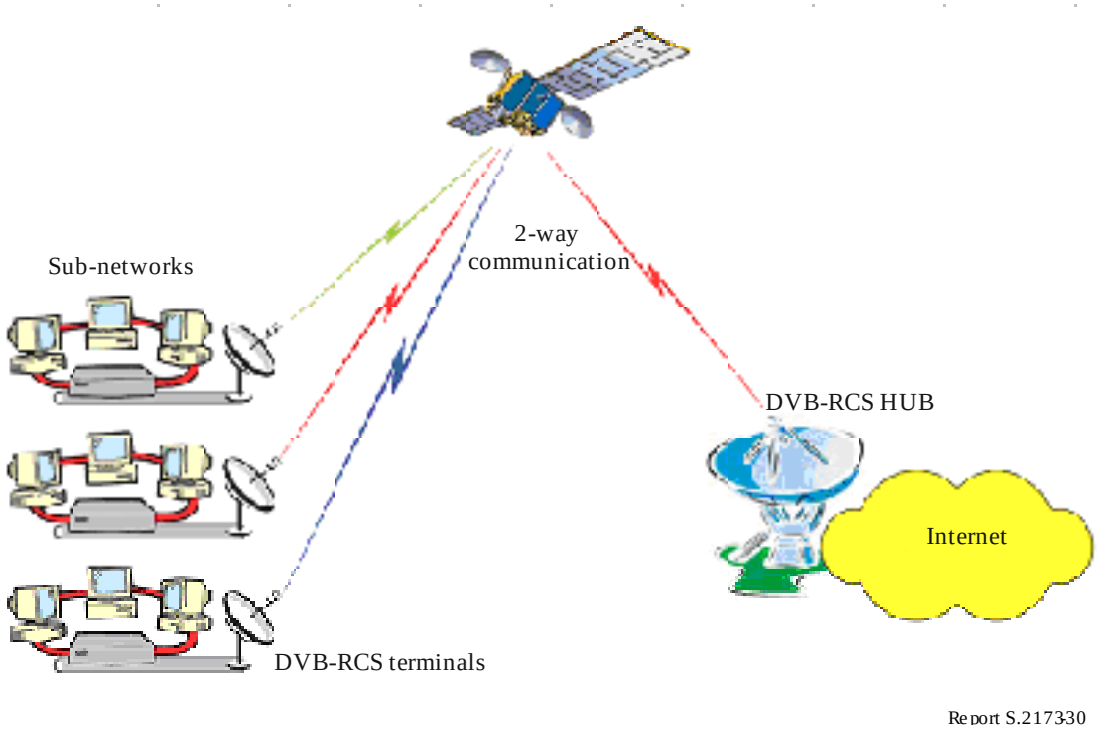
Report S.2173-29

9.3 DVB-RCS

The digital video broadcast – return channel via satellite (DVB-RCS) is an open standard that defines the air interface specifications for the two-way communication of broadband VSAT terminals. It delivers demand-assigned connections that are equivalent to ADSL or cable internet – with a maximum of 20 Mbit/s to each terminal in the forward link and 5 Mbit/s to each terminal in the return link. Figure 30 demonstrates the network architecture of DVB-RCS [75]. The most recent update to DVB-RCS (DVB-RCS+M) includes support for mobility and nomadic terminals, as well as, enhanced support for direct terminal-to-terminal mesh connectivity. As of mid-2007 there have been more than 150 DVB-RCS systems deployed worldwide, serving around 100 000 terminals. DVB-RCS user terminals offer broadband services which can include interactive DVB services, IPTV and full IP connectivity.

FIGURE 30

Network architecture for DVB-RCS [75]



Forward link

The forward link of DVB-RCS is designed to be shared amongst user and is implemented either using DVB-S or DVB-S2 depending on operator preference. In the case where DVB-S2 is used there are two transmission options: CCM and ACM.

Return link

The return link of DVB-RCS is a multi-frequency TDMA (MF-TDMA) link that is accessed using a demand-based assignment scheme, which attempts to maximize application functionality using quality of service classes (e.g. voice, video streaming, file transfers and web browsing). There are two channel codes that can be used for the return link. These are the concatenated codes of the DVB-S standard and a double binary circular recursive systematic CTC, which can be punctured to achieve multiple rates. The possible rates for the punctured CTC are 1/3, 2/5, 1/2, 2/3, 3/4, 4/5 and 6/7. While the code rate is adaptive using either the concatenated code or CTC, the modulation is a fixed QPSK constellation.

Continuous carrier operation (mobile option)

A mobile option is defined for DVB-RCS to deal with mobile or nomadic terminals. This is referred to as continuous carrier mode in the DVB-RCS standard with specifications for both the forward and return link. Both the forward and return links can use spectrum spreading.

The forward link of the mobile option uses the same specifications detailed in § 7.3.2 and can be operated in CCM mode for the case where DVB-S2 carriers are used.

The baseline transmission scheme for the return link is DVB-S2; where the return link can be operated in both CCM and ACM modes. A $\pi/2$-BPSK constellation can be used for continuous return link transmission in addition to the QPSK, 8-PSK, 16-APSK, 32-APSK constellations defined in DVB-S2. The use of concatenated codes with smaller frame size (4 096 bits) instead of the DVB-S2 channel codes is optional. These short-frame codes are composed of a BCH outer code and a LDPC inner code. In the event that the short frame size codes are used, the provisions of the DVB-S2 standard must still be followed.

9.4 DVB-SH

DVB-SH [21] is a standard designed for mobile TV, which leverages satellites to achieve coverage of large region, which may include entire countries. Terrestrial gap-fillers are used to complement the satellite component as in integrated MSSs, such that coverage is maintained in areas where direct reception of the satellite signal is not possible. This standard is designed for frequencies below 3 GHz and is envisioned for use in the S-Band (2.2 GHz) as a MSS.

The deployment of DVB-SH is largely in the development phase. It was confirmed in 2006 by the European Commission that a slice of S-Band could be used for MSSs and that these services could include a complementary ground component of an integrated satellite/terrestrial system. In addition, it is reported in [76] that authorities in France are considering the use of DVB-SH in ultra high frequency (UHF) bands. In 2008, US GEO satellite was launched for the deployment of a nationwide DVB-SH network that is capable of providing video services, navigation information, and emergency messaging [76]. The satellite is still holding trials on the efficacy of this service.

Architectures

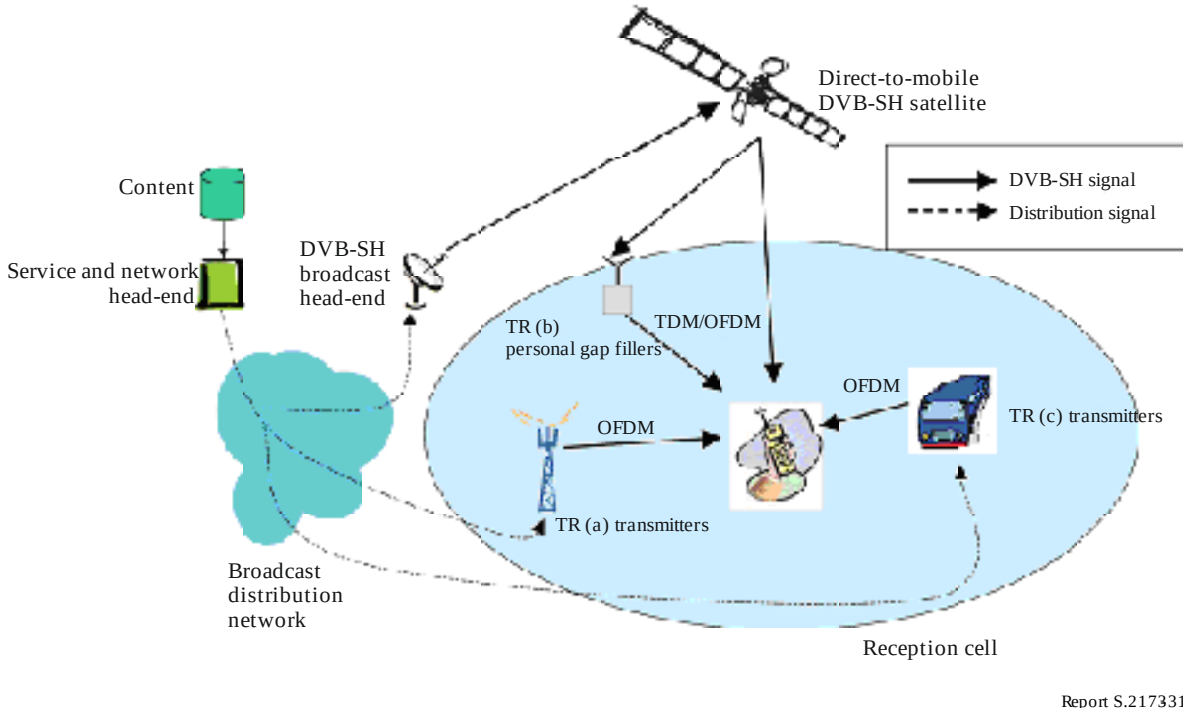
Figure 31 shows the network architecture for a DVB-SH network. The network operates as a broadcast system, which could be operated in one of two implementations: SH-A and SH-B. In SH-A, OFDM is used by both the satellite and terrestrial components. This would simplify the user equipment as only one chipset would be required to receive a DVB-SH transmission. The SH-A implementation also lends itself to be implemented as a SFN. In SH-B the satellite link transmits using a single-carrier TDM frame structure, while the terrestrial link transmits using OFDM. The use of single-carrier TDM on the satellite link is most likely done to avoid issues relating to the HPA's difficulties in dealing with the high PAPRs of OFDM.

Channel coding and modulation

The channel coding is the same in both single-carrier TDM and OFDM modes. The Turbo code of the 3GPP2 standard [59] is used, as is done for the DVB-H standard [77]. The 3GPP2 Turbo code is a CTC that employs two RSC encoders. This code achieves rates of 1/5, 2/9, 1/4, 2/7, 1/3, 2/5, 1/2 and 2/3 by puncturing the encoder output. It is also worth mentioning that a time interleaver is used in both modes to obtain time-diversity.

FIGURE 31

Network architecture for DVB-SH [74]



In single-carrier TDM mode, QPSK, 8-PSK and 16-APSK as defined in DVB-S2 [38] are used.

In OFDM mode, QPSK, 16-QAM and non-uniform 16-QAM constellations are used, with two types of non-uniform 16-QAM being available. The non-uniform 16-QAM constellations are defined as: $\{-4, -2, 2, 4\}\sqrt{\varepsilon_s}$ and $\{-6, -4, 4, 6\}\sqrt{\varepsilon_s}$; where the part in curly brackets being the possible amplitudes of the quadrature and in-phase components of the modulation ε_s is a energy-scaling factor for a unit-energy constellation[15]. Note that regular 16-QAM is defined as $\{-3, -1, 1, 3\}\sqrt{\varepsilon_s}$. These irregular 16-QAM constellation effectively bring the constellation points closer together (when compared with regular 16-QAM); however, the constellation points at the centre of the constellation have more protection against noise.

10 Performance parameters and models

This section demonstrates the performance of different implementations of a multi-carrier satellite communications system. Section 10.1 gives the performance of a multi-carrier satellite system (MCSS) with an ideal linear high power amplifier (HPA) and OFDM. Since the HPA is ideal there is no non-linear distortion and the performance of the MCSS is the same as it is for a single-carrier satellite system (SCSS) in an additive white Gaussian noise (AWGN) channel. Section 10.3 contrasts § 10.1 by giving the performance of a MCSS with a travelling wave tube amplifier (TWTA). The performance of the MCSS with TWTA is evaluated in an AWGN channel.

[15] That is, a constellation where the average energy of the constellation is equal to one.

10.1 Performance and spectral efficiency of a multi-carrier satellite system in linear channels

As mentioned earlier in this section, the MCSS with ideal linear HPA is equivalent to a SCSS. Therefore in this section, the results for a SCSS in an AWGN channel are presented. Table 3 presents a possible ACM scheme composed of different combinations of QAM constellations and BTCs. The superscripts in the BTC scheme indicate 2-dimensional or 3-dimensional BTCs. As component codes, expurgated versions of the BCH codes are used. The performance results are obtained using a FPGA implementation which makes use of a modified version of SOVA [46]. A particular combination of a modulation and channel code, with specified coding rate is referred to as a MODCOD. For sake of brevity, each MODCOD is listed by the modulation and the code rate of the specific channel code employed. For example, in Table 3 MODCODs are listed using the type of modulation and the code rate of the BCH BTC employed.

Table 3 also gives the AWGN performance of each particular MODCOD for the SCSS. Note that symbol-repetition (SR) is used to improve the robustness of the ACM scheme. For completeness, the signal-to-noise ratio (SNR or E_b/N_0) and CNR (or E_s/N_0) required to obtain a BER of 10^{-6} are presented side-by-side. Examining both is important, especially when observing results that include the use of SR. With a SR scheme, the same symbol is re-transmitted Nr times. The power of the received symbols is $10 \log_{10}(N_r E_s)$ dB for a particular symbol. Notice that for results presented in Table 3, where SR is used, that when the number of symbols transmitted doubles, the performance improves by 3 dB. The use of SR does not give a coding gain, and as a result is not more energy efficient. However, what a SR scheme does provide is a way to maintain availability in the event that the channel's capacity is lower than the rate of the lowest supported code. Presenting results in terms of SNR is a good way to compare the energy efficiency of different ACM schemes, whereas presenting results in terms of CNR allows for the creation of a look-up table detailing the CNRs at which to utilize each MODCOD.

Table 4 demonstrates the SNR needed for quasi-error-free (QEF) packet errors rates (PER)s using the DVB-S2 APSK constellations and channel codes as MODCOD. The received samples are decoded using 50 BPA decoding iterations. Table 4 is a modified version of the one presented on page 33 of the DVB-S2 standard [38]. As mentioned above, the energy efficiency of an ACM scheme is best observed in terms of SNR, whereas the operational usage of each MODCOD is best decided using CNR. For example, the QPSK 1/3 MODCOD can achieve a PER at lower SNR than the QPSK 1/4 MODCOD, while delivering a higher bit-rate. This means that compared to the rate 1/3 DVB-S2 channel code, the rate 1/4 code is inefficient. However, if the CNR were less than –1.24 dB, it would still be better to use the rate 1/4 code, given the QEF PER requirement. Additionally, it is clear that there is some overlap in the spectral efficiencies of some MODCODs. For example, the QPSK 9/10 MODCOD, which has a spectral efficiency of 1.79 bit/s/Hz, requires a higher CNR than the 8-PSK 3/5 MODCOD, which has a similar spectral efficiency. Clearly, it would be best to use the 8-PSK 3/5 MODCOD in this CNR range. Therefore, an efficient implementation of ACM only includes those MODCODs that have a CNR advantage over other MODCODs. If they do not have CNR where they have a greater spectral efficiency than all other MODCODs, they should not be used.

TABLE 3

Various BTC MODCODs for the SCSS

MODCOD	Spectral efficiency (Bit/s/Hz)[16]	E_s/N_0 (dB) @BER = 10^{-6}	E_b/N_0 (dB) @BER = 10^{-6}	BTC scheme
8SR-BPSK 0.037	0.037037	–12.5	1.81	$(15,10)^3$
4SR-BPSK 0.074	0.074074	–9.5	1.80	$(15,10)^3$
2SR-BPSK 0.15	0.148148	–6.5	1.79	$(15,10)^3$
BPSK 0.30	0.296296	–3.5	1.78	$(15,10)^3$
QPSK 0.30	0.592592	–0.5	1.77	$(15,10)^3$
QPSK 0.79	1.580246	5.5	3.51	$(63,56)^2$
8-PSK 0.30	0.888888	2.5	3.01	$(15,10)^3$
8-PSK 0.65	1.951093	8.0	5.10	$(31,25)^2$
16-QAM 0.79	3.160494	12.5	7.50	$(63,56)^2$
64-QAM 0.65	3.902185	15.5	9.59	$(31,25)^2$
64-QAM 0.79	4.740741	18.5	11.74	$(63,56)^2$

TABLE 4

Various DVB-S2 MODCODs for the SCSS

MODCOD	Spectral efficiency (Bit/s/Hz)	E_s/N_0 (dB) @ PER = 10^{-7}	E_b/N_0 (dB) @ PER = 10^{-7}
QPSK 1/4	0.490243	–2.35	0.75
QPSK 1/3	0.656448	–1.24	0.59
QPSK 2/5	0.789412	–0.30	0.73
QPSK 1/2	0.988858	1.00	1.05
QPSK 3/5	1.188304	2.23	1.48
QPSK 2/3	1.322253	3.10	1.89
QPSK 3/4	1.487473	4.03	2.31
QPSK 4/5	1.587196	4.68	2.67
QPSK 5/6	1.654663	5.18	2.99
QPSK 8/9	1.766451	6.20	3.73
QPSK 9/10	1.788612	6.42	3.89
8-PSK 3/5	1.779991	5.50	3.00

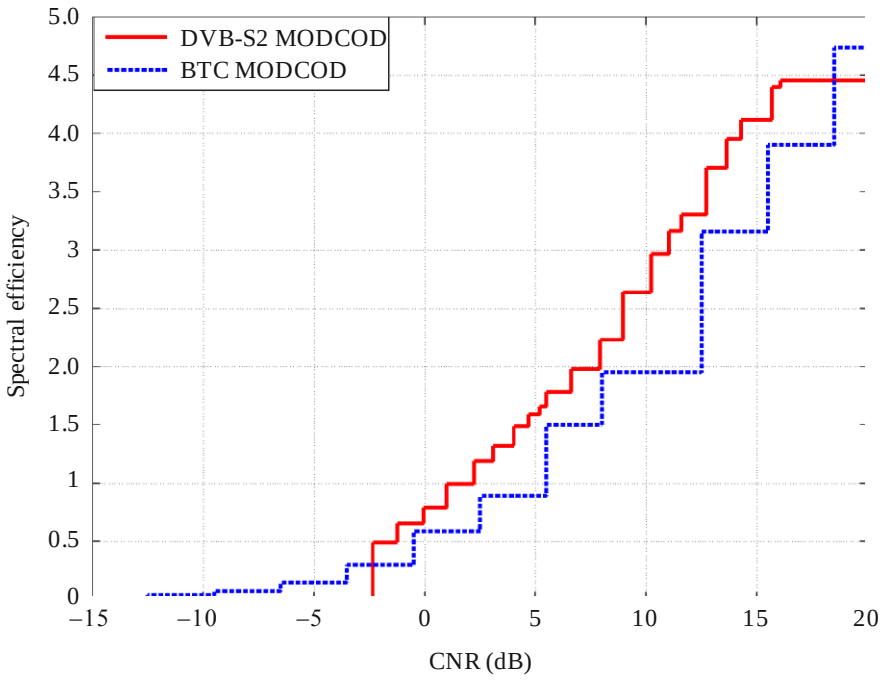
[16] Note that all results presented in this section assume that filtering is done at the minimum Nyquist rate. In practice, the spectral efficiency is also influenced by the amount of usable bandwidth determine by the filtering coefficient (α for a raised cosine filter). Thus, these results represent the theoretical maximum spectral efficiency.

TABLE 4 (*end*)

MODCOD	Spectral efficiency (Bit/s/Hz)	E_s/N_0 (dB) @ PER = 10^{-7}	E_b/N_0 (dB) @ PER = 10^{-7}
8-PSK 2/3	1.980636	6.62	3.65
8-PSK 3/4	2.228124	7.91	4.43
8-PSK 5/6	2.478562	9.35	5.41
8-PSK 8/9	2.646012	10.69	6.46
8-PSK 9/10	2.679207	10.98	6.70
16-APSK 2/3	2.637201	8.97	4.76
16-APSK 3/4	2.966728	10.21	5.49
16-APSK 4/5	3.165623	11.03	6.03
16-APSK 5/6	3.300184	11.61	6.42
16-APSK 8/9	3.523143	12.89	7.42
16-APSK 9/10	3.567342	13.13	7.61
32-APSK 3/4	3.703295	12.73	7.04
32-APSK 4/5	3.951571	13.64	7.67
32-APSK 5/6	4.119540	14.28	8.13
32-APSK 8/9	4.397854	15.69	9.26
32-APSK 9/10	4.453027	16.05	9.56

FIGURE 32

Comparison of BTC and DVB-S2 ACM schemes



Report S.2173-32

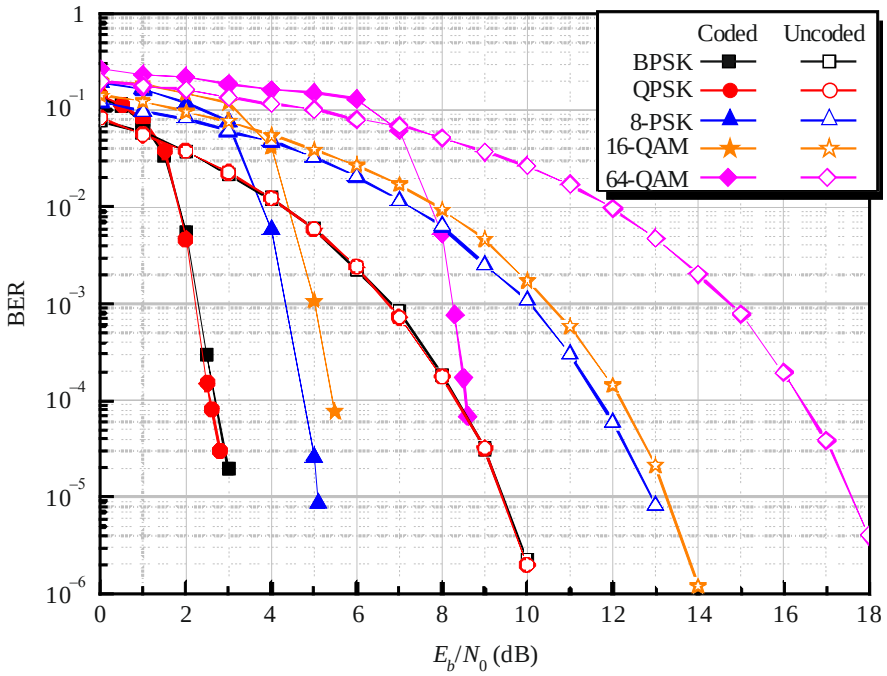
Figure 32 demonstrates the spectral efficiency of both the BTC and DVB-S2 ACM schemes by plotting the results from Tables 3 and 4. Although these results are listed for a BER of 10^{-6} and a PER of 10^{-7}, the comparison is still acceptable. This is due to the fact that modern channel codes (BTCs and LDPC) have very sharp slope after the "knee" of their error-rate curves, as demonstrated in Figs 33 to 39. This means that there is negligible difference (~0.25 dB) between the CNR required to obtain a PER of 10^{-2} and 10^{-7}, for example. In fact, it can be observed that when the PERs of the DVB-S2 codes are roughly 10^{-3}, their respective BERs are somewhere between 10^{-5} and 10^{-6}. Note that the curves are plotted using the maximum spectral efficiency generated by all MODCODs at each CNR. That is, if MODCOD *x* has higher spectral efficiency than MODCOD *y*, and MODCOD *x* has lower CNR than MODCOD *y*, then MODCOD *y* is omitted from Fig. 32. MODCODs not included in Fig. 32 are underlined in Table 4. Figure 32 suggests that the DVB-S2 ACM scheme has better energy efficiency than the BTC ACM scheme. However, the DVB-S2 ACM uses a longer block length than the BTC ACM. Note that the DVB-S2 packet length is designed to be the same as the block length by the DVB-S2 standard; whereas, there is no defined packet length for the BTC ACM. Long packet lengths are more difficult to handle when using packet windowing, which for a SR/ARQ re-transmission protocol requires sufficient buffer size in the transmitter and receiver.

Larger packet sizes require larger buffers and buffer size may be limited in some cases. An example is the user equipment of a mobile satellite service. In this case, it may be more advantageous to make use of the shorter length BTCs or the mobile option channel codes described in the DVB-RCS standard [43]. Additionally, larger packets can have latency issues, which may limit the use of longer length channel codes for latency sensitive applications such as voice. However, in the case where sufficient buffering is available for packet windowing and latency is not an issue, the DVB-S2 channel codes provide excellent performance.

Figures 33 to 35 show the BER performance of BTC MODCODs in Table 3, while Figs 36 to 39 show those of selected DVB-S2 MODCODs.

FIGURE 33

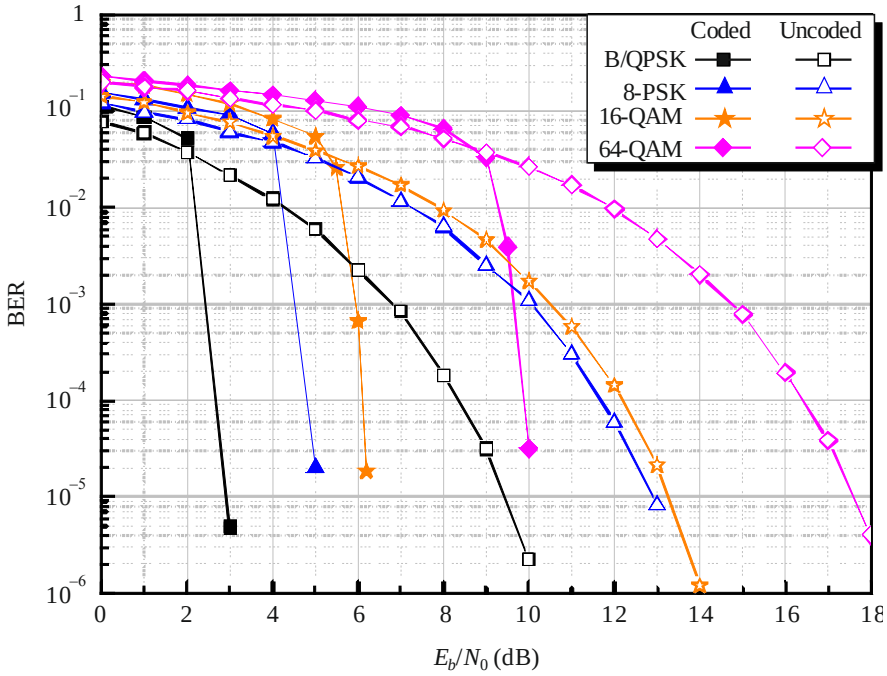
BER performance of BTC using the modified SOVA for $(31,25)^2$ in AWGN channel



Report S.217333

FIGURE 34

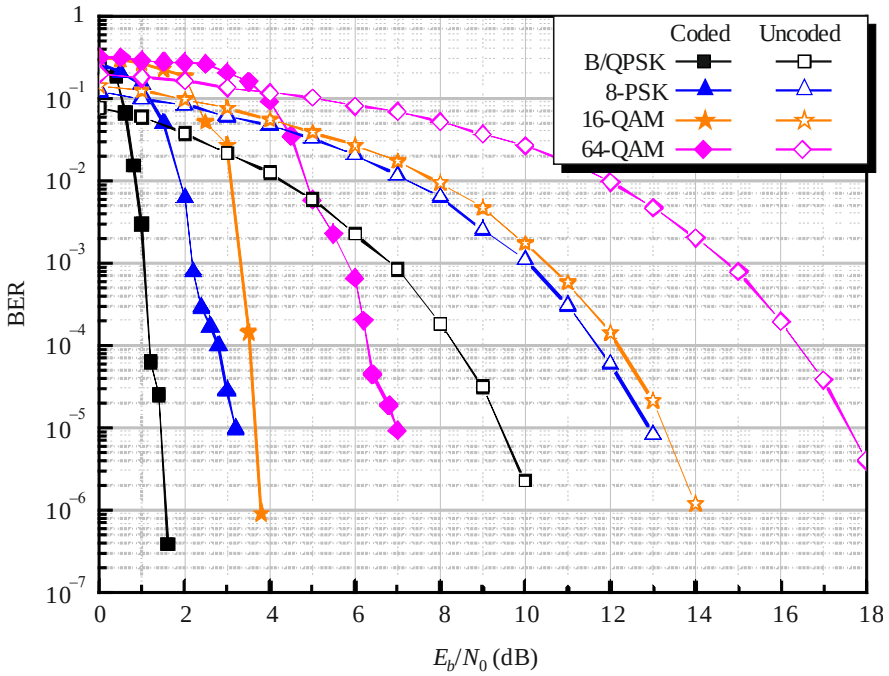
BER performance of BTC using the modified SOVA for $(63,56)^2$ in AWGN channel



Report S.217334

FIGURE 35

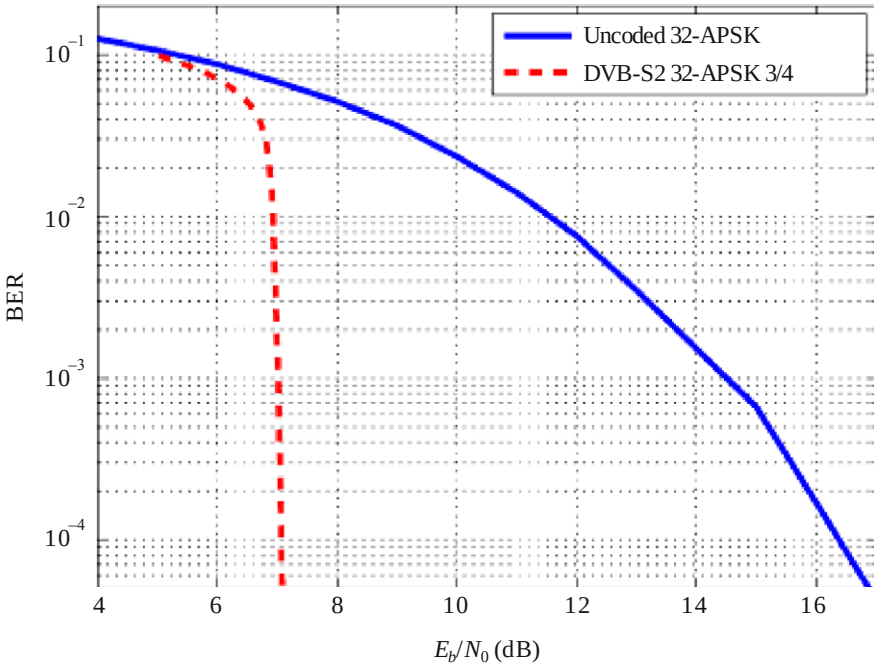
BER performance of BTC using the modified SOVA for $(15,10)^3$ in AWGN channel



Report S.217335

FIGURE 36

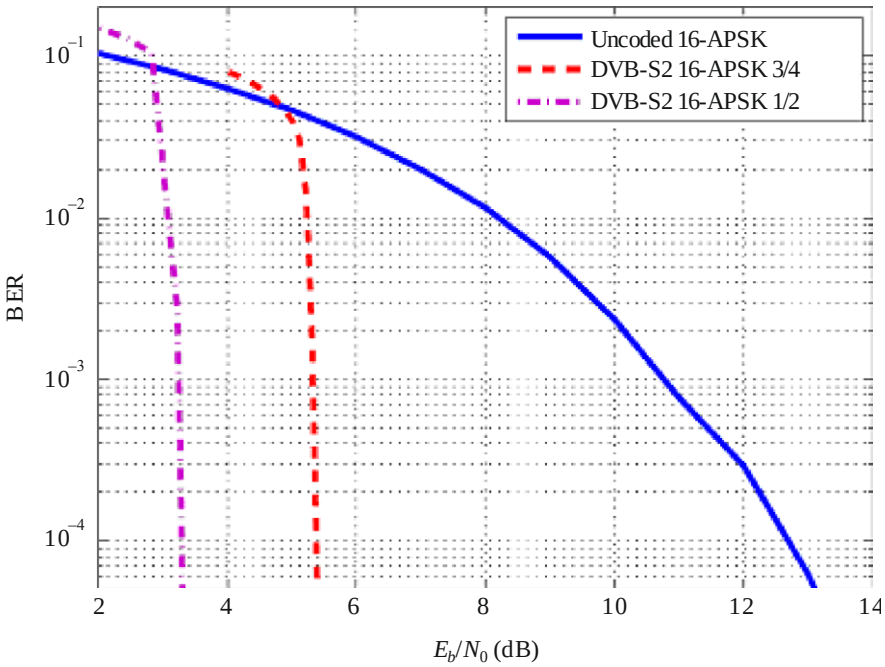
BER performance of DVB-S2 32-APSK in AWGN channel



Report S.217336

FIGURE 37

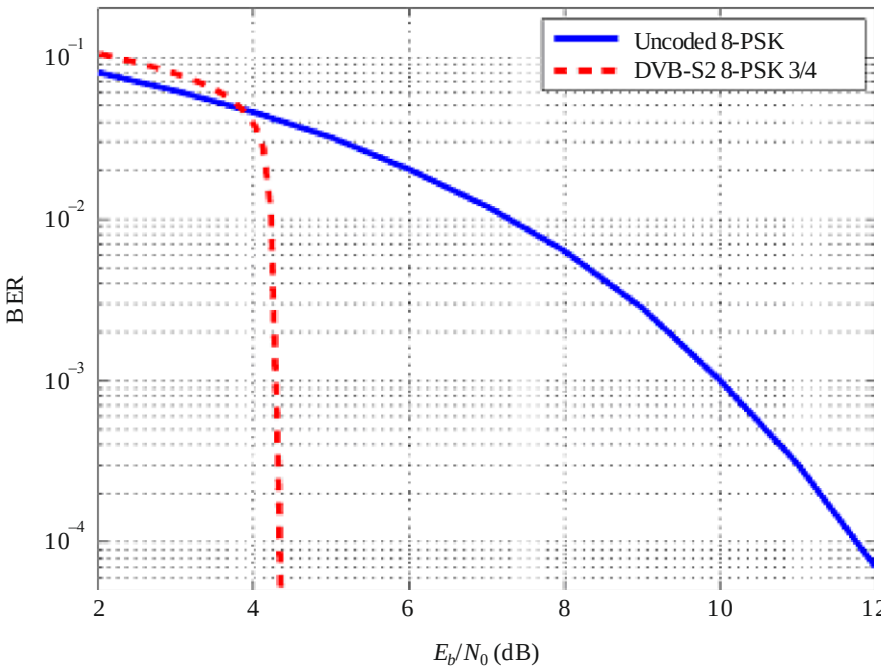
BER performance of DVB-S2 16-APSK in AWGN channel



Report S.217337

FIGURE 38

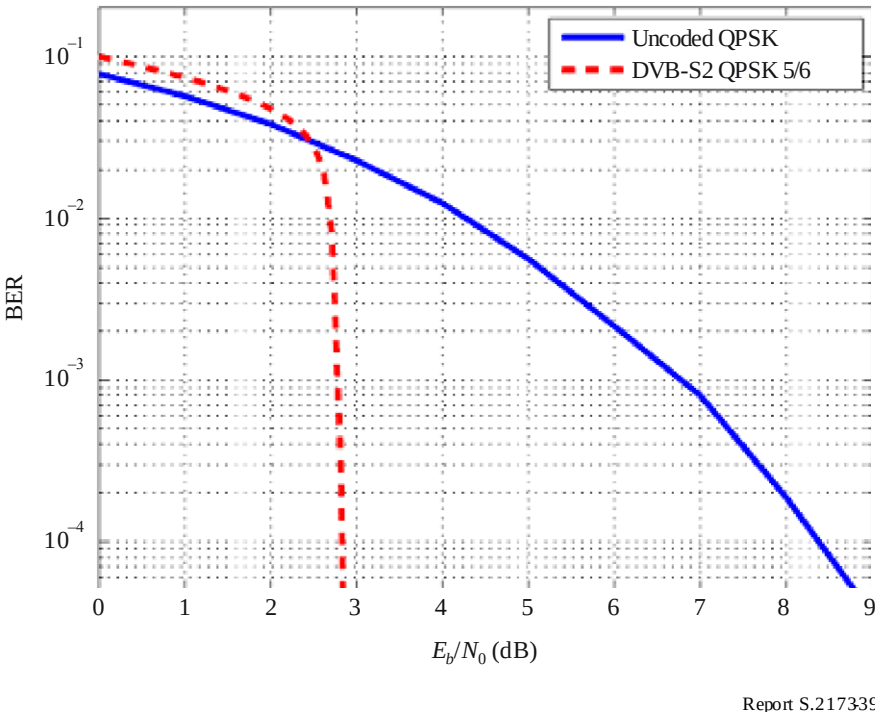
BER performance of DVB-S2 8-PSK in AWGN channel



Report S.217338

FIGURE 39

BER performance of DVB-S2 QPSK in AWGN channel



10.2 Evaluation of direct spectrum division transmission (DSDT)

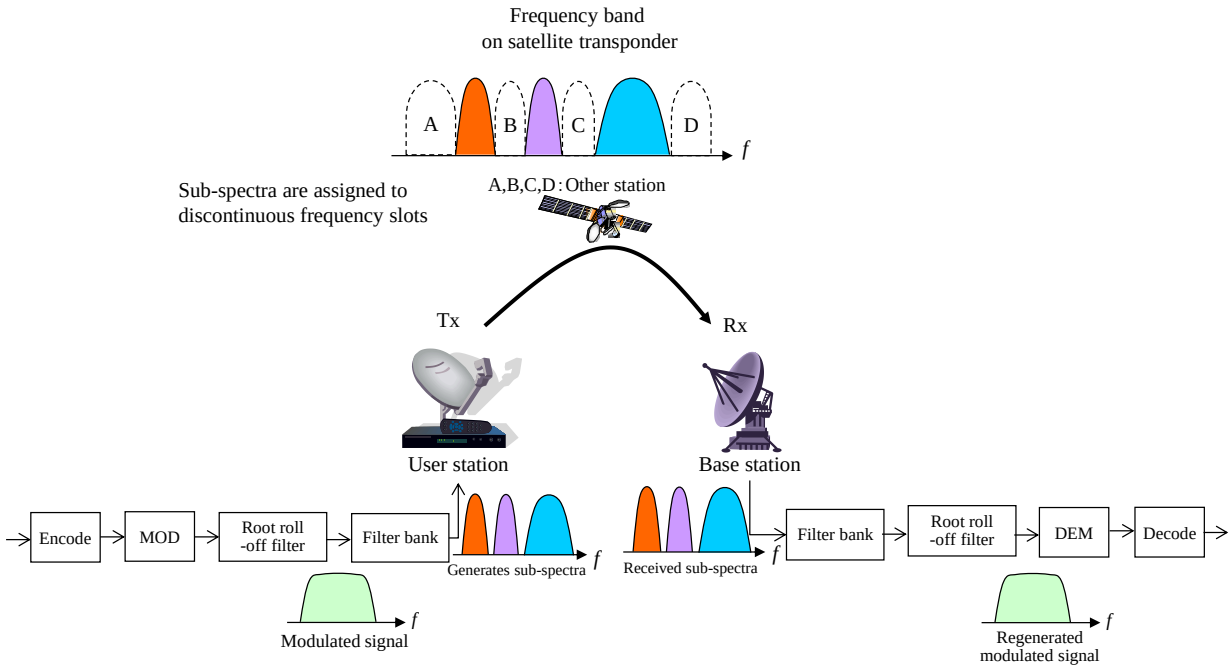
The performance of the DSDT for a satellite communication system is investigated in this section.

10.2.1 System model

Figure 40 shows the concept of communications by the DSDT via satellite. In the transmitter (Tx) side, the single-carrier modulated signal is converted into the frequency domain. Its spectrum is divided into multiple sub-spectra and assigned to unused discontinuous frequency slots. The multiple sub-spectra are converted into the time domain signal and transmitted. In the receiver (Rx) side, the received signal is converted into a frequency domain signal, and the sub-spectra are shifted and combined. This retrieves the Tx signal.

FIGURE 40

System model of DSDT via satellite



10.2.2 Test results

To evaluate the basic transmission characteristics of the DSDT, several computer simulations were carried out. Table 5 shows the simulation parameters. Figure 41 plots an example of multiple sub-spectra divided by the DSDT. Since each sub-spectrum is shaped by a root roll-off filter, the multiple sub-spectra created by the DSDT are equivalent to the spectra of multi-carrier FDMA signal transmission. Figure 42 plots the CCDF (complementary cumulative distribution function) of PAPR with the DSDT. For reference, the PAPR of multi-carrier FDMA signals is also plotted. These results show that the PAPR of the DSDT is smaller than that of multi-carrier FDMA for any division number. In particular, PAPR of the DSDT is about 2.0 dB superior to that of multi-carrier FDMA with 32 carriers at CCDF=10^{-4}.

To confirm the fundamental principle of the DSDT, the BER performance was simulated. In this simulation, the maximum number of spectrum divisions is 32. Figure 43 plots the BER vs E_b/N_0 performance of our proposal. To verify its fundamental attributes, the BER performance was evaluated while changing the number of sub-spectra. Regardless of the spectrum division scenarios, the resulting BER performance is close to that of the non-divided case. This means the spectrum division conserves signal quality.

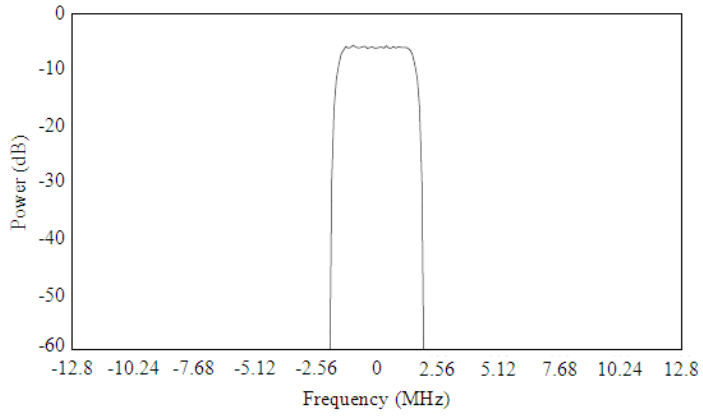
TABLE 5

Simulation parameters

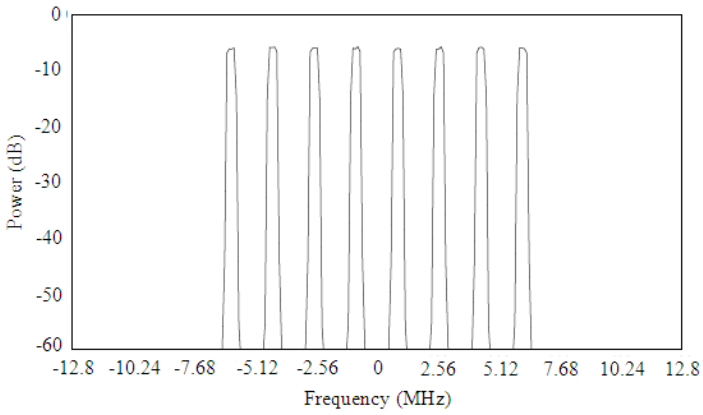
Modulation	**QPSK**
FEC	*R*=1/2 Convolutional code / Viterbi decoding
Sampling rate	25.6 MHz
Symbol rate	3.2 MHz
Type of spectrum division	1,2,4,8,32
FFT size	2048

FIGURE 41

Example of direct spectrum division



(a) Single carrier modulated signal



(b) Spectrum divided signals by the DSDT

FIGURE 42

PAPR of direct spectrum division transmission

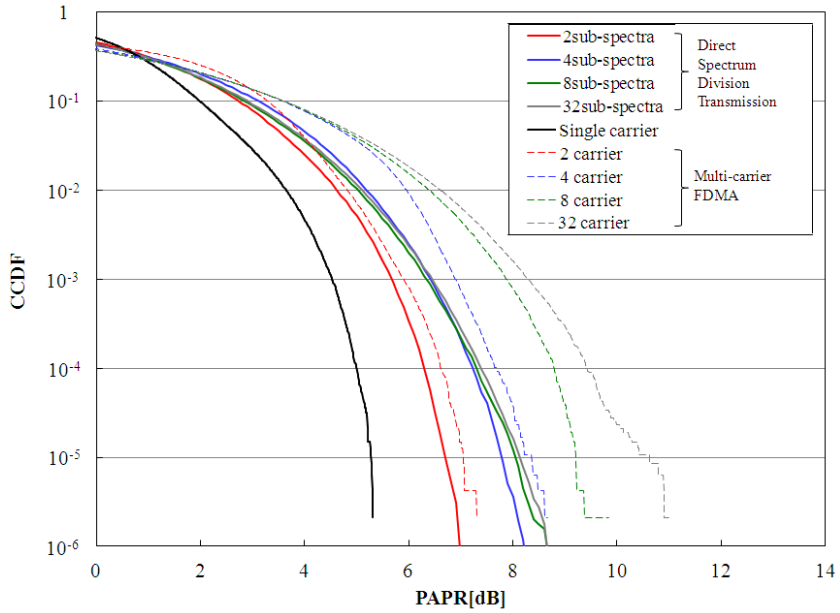
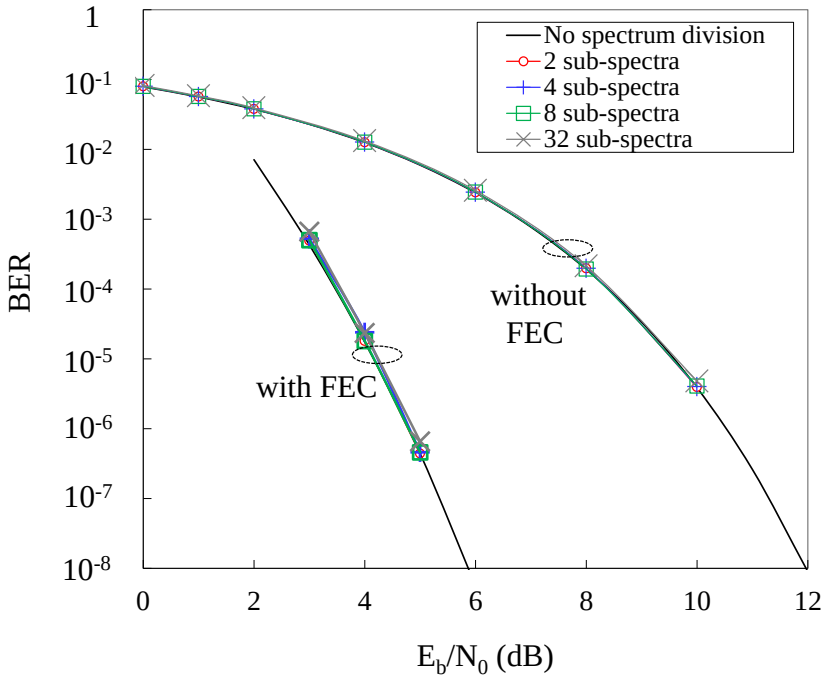


FIGURE 43

BER vs E_b/N_0 performance of direct spectrum division transmission



To evaluate the proposal, several signal transmission experiments were carried out by using a prototype DSDT modem. Table 6 shows the experimental parameters. Figure 44 shows the picture of prototype modem with hardware implementation of DSDT and an example of multiple sub-spectra divided by the DSDT. Since each sub-spectrum is shaped by a root roll-off filter, the multiple sub-spectra created by the DSDT are equivalent to the spectra of multi-carrier FDMA signal transmission.

TABLE 6

Experimental parameters

Modulation	**QPSK**
FEC	*R*=1/2 Convolutional code / Viterbi decoding
Sampling rate	25.6 MHz
Symbol rate	3.2 MHz
Type of spectrum division	1,2,4,8,32
FFT size	2048

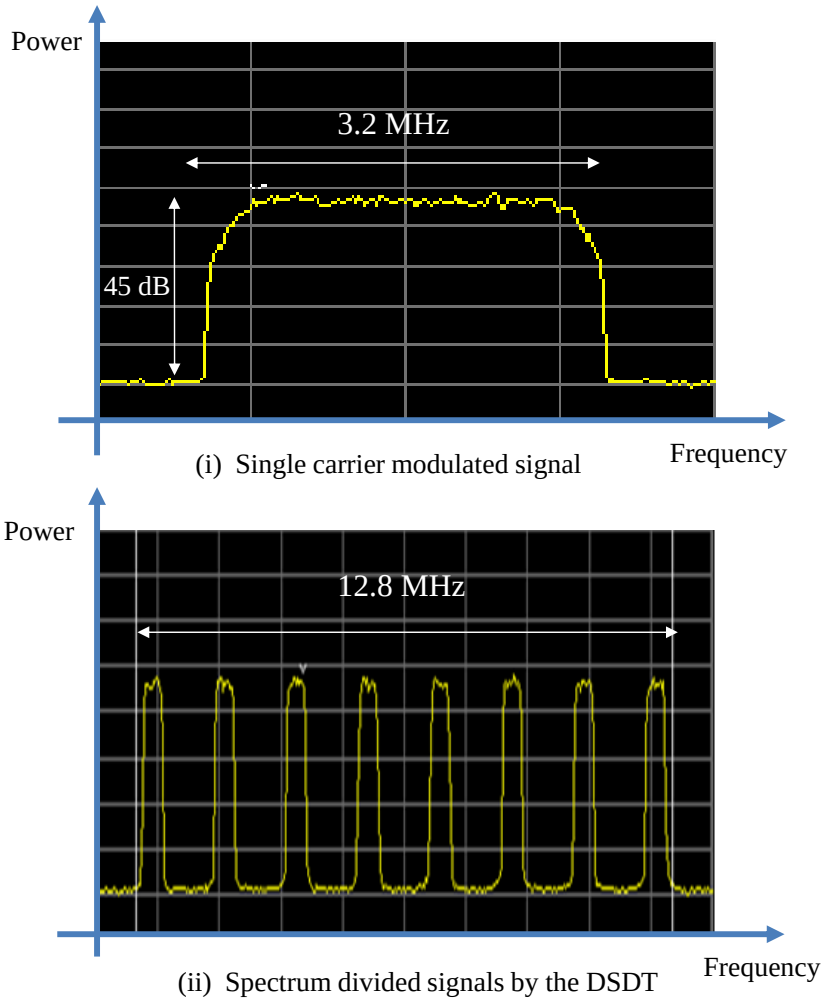
Figure 45 shows the evaluation on PAPR performance by using a prototype DSDT modem and a multi-carrier FDMA modem. As well as the simulation results in Figure 42, the results of hardware experiments show that the PAPR of the DSDT is smaller than that of multi-carrier FDMA for any division number. In particular, PAPR of the DSDT is about 1.5 dB superior to that of multi-carrier FDMA with 32 carriers at CCDF=10^{-4}.

Figure 46 shows the BER performance of a prototype DSDT modem. The resulting BER performance is close to that of the non-divided case. From these results of hardware experiments, it is concluded that the DSDT is a practical transmission technique.

FIGURE 44

Experimental results of direct spectrum division by using the prototype modem

(a) Picture of the prototype modem with hardware implementation of DSDT



(b) Measured spectrum of DSDT with the prototype modem

FIGURE 45

Performance evaluations on PAPR with prototype modems

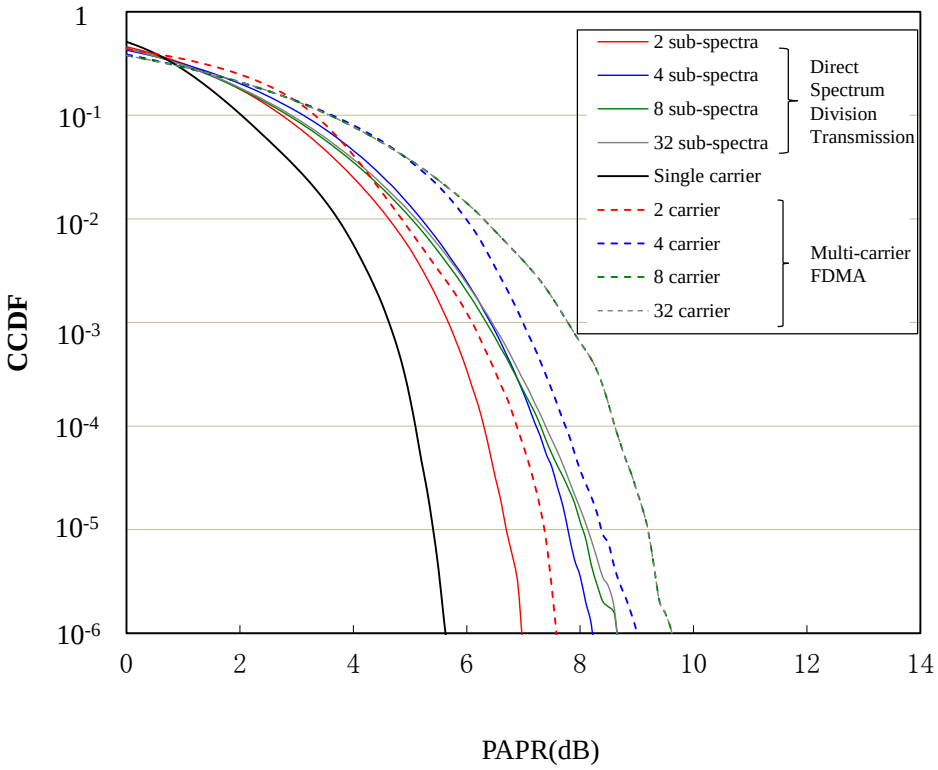
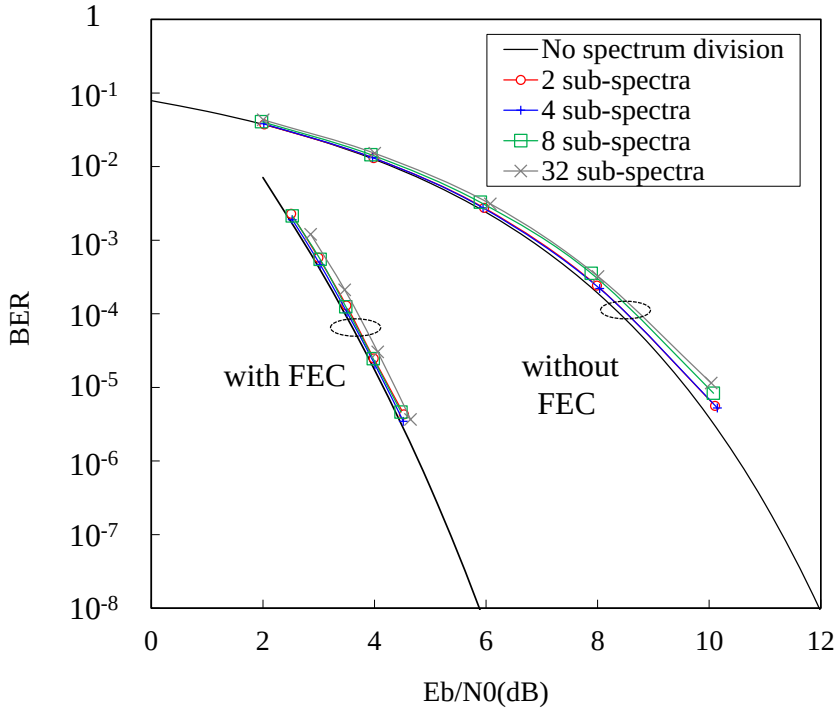


FIGURE 46

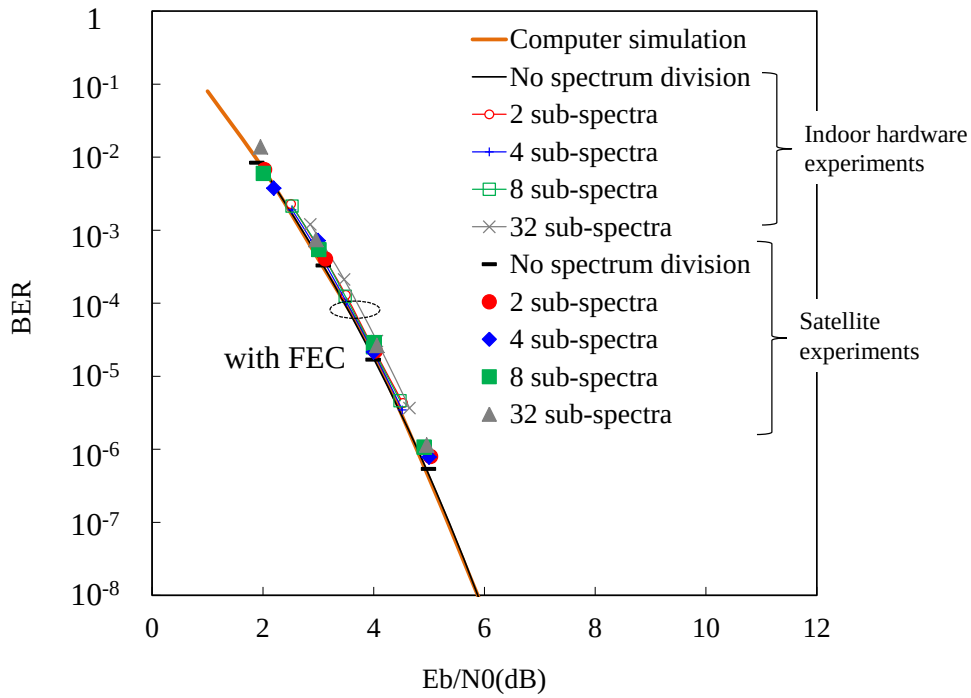
BER vs E_b/N_0 performance of direct spectrum division transmission (indoor hardware experiments)



To evaluate the overall performance of the DSDT technique, several satellite experiments were conducted in the 14/12 GHz bands. The experimental parameters are the same as those in Table 6. Figure 47 shows the practical BER performance via satellite. As shown in Figure 47, the degradation in required E_b/N_0 was negligible compared to the performance of indoor hardware experiments and computer simulations. No significant performance degradation was observed during the satellite experiments.

FIGURE 47

BER vs E_b/N_0 performance of direct spectrum division transmission (satellite experiments)



10.3 Evaluation of CI-OFDM transmissions in a non-linear satellite channel

The performance test results of CI-OFDM transmission over a satellite link is investigated in this section.

10.3.1 System model

Figure 48 shows the configuration of the test system. The transponder can use a 36 MHz reverse/forward link bandwidth in 12/14 GHz of the Ku-Band, 20/30 GHz of the Ka-Band, and 4/6 GHz of the C-Band. A vector signal generator (VSG) and vector signal analyser (VSA) are used as ideal CI-OFDM modulators. OFDM modulation with the same frame format as CI-OFDM is used for comparison.

The frame structure of CI-OFDM used by the test system is shown in Fig. 49. Each frame is composed of a block of pilot signals (Pilot) for channel and noise estimation; a unique word (UW) for frame synchronization; and the modulated data for transmission. The Pilot signal is generated using a regular OFDM signal generator. The UW and the data block – transmitted following the Pilot signal – are generated using the CI-OFDM signal generator shown in Fig. 50.

FIGURE 48

Configuration of performance test system for CI-OFDM transmission over a satellite link

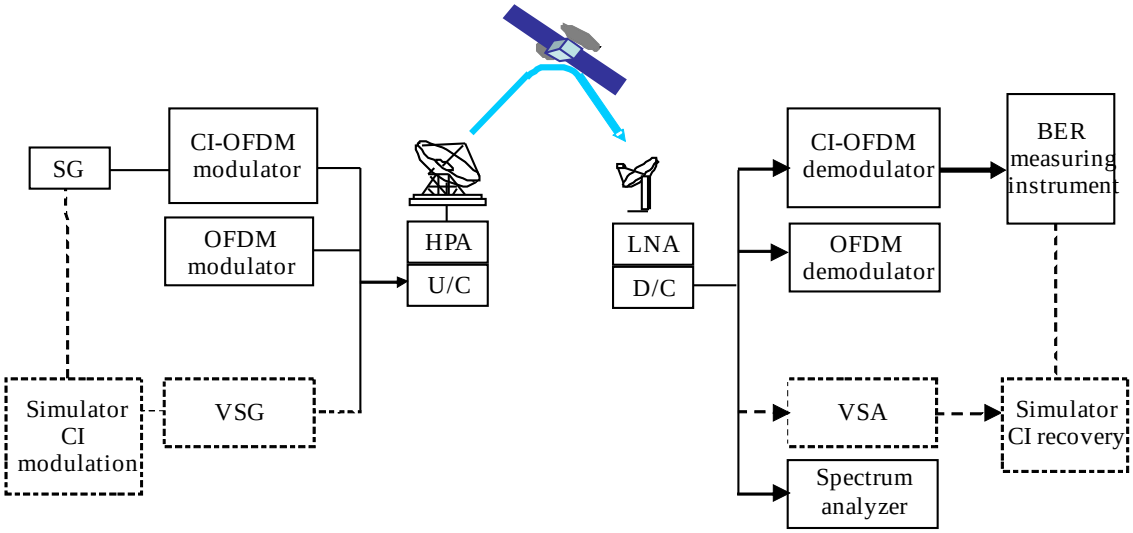


FIGURE 49

CI-OFDM frame structure

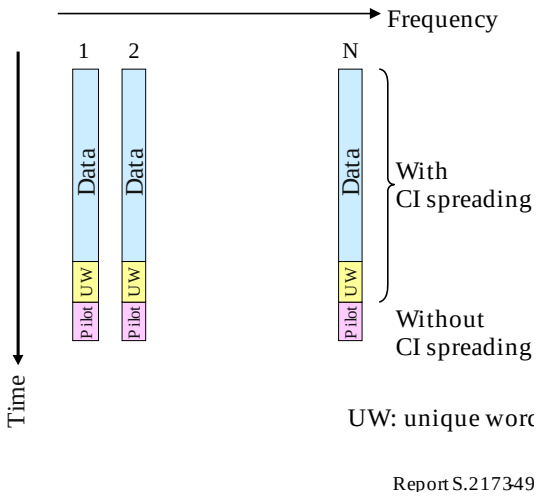
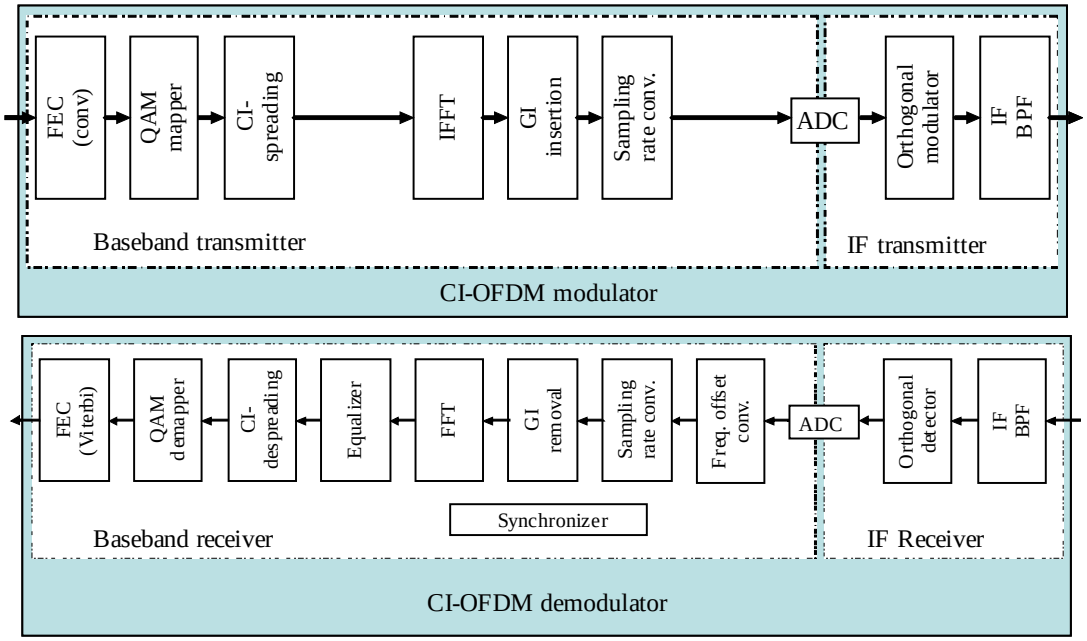


Figure 50 shows the block diagrams of CI-OFDM modulator and demodulator. The modulator consists of a baseband and intermediate frequency (IF) transmitter that converts the baseband signal into an equivalent analogue multi-carrier IF signal. The baseband transmitter is composed of an encoder for a convolutional code (FEC (Conv)), which converts message-word inputs into codewords; a QAM mapper; a CI spreading IFFT module; GI insertion denoted as GI Insertion; and a module that does oversampling, denoted as sampling rate converter.

The demodulator consists of a baseband receiver and IF receiver that converts the analogue IF signal to an equivalent multi-carrier baseband signal. The demodulator has a Frequency Offset Conversion module that attempts to correct any frequency synchronization offset. The baseband signal is then converted into a received data stream by passing it through a chain that includes a Sampling Rate Converter; a GI removal module; a FFT module; a CI-despreading module; a QAM Demapper and a Viterbi decoder (FEC (Viterbi)).

FIGURE 50

Block diagrams of CI-OFDM MODEM



Report S.217350

The CI-OFDM modem can employ several combinations of convolutional codes and QAM modulation, where the 1/1 convolutional code represents an uncoded system. The parameters and information rates of the several combinations of convolutional codes and QAM modulation are shown in Table 7.

TABLE 7

Information rate of ACM in each transmission band (Mbit/s)

BW (MHz)	Mod	QPSK			16-QAM			64-QAM		
	Code rate of FEC / No. of subcarriers	1/1	1/2	3/4	1/1	1/2	3/4	1/1	3/4	7/8
1	32	1.47	0.74	1.11	2.95	1.47	2.21	4.42	3.32	3.87
2	32	3.32	1.66	2.49	6.64	3.32	4.98	9.96	7.47	8.71
5	32	8.85	4.42	6.64	17.7	8.85	13.2	26.5	19.9	23.2
10	64	18.3	9.17	13.7	36.6	18.3	27.5	55.0	41.2	48.1
20	128	37.3	18.6	28.0	74.7	37.3	56.0	112.0	84.0	98.0
27	256	50.7	25.3	38.0	101.5	50.7	76.1	152.2	114.2	133.2
36	256	67.8	33.9	50.8	135.6	67.8	101.7	203.4	152.5	177.9

10.3.2 Test results

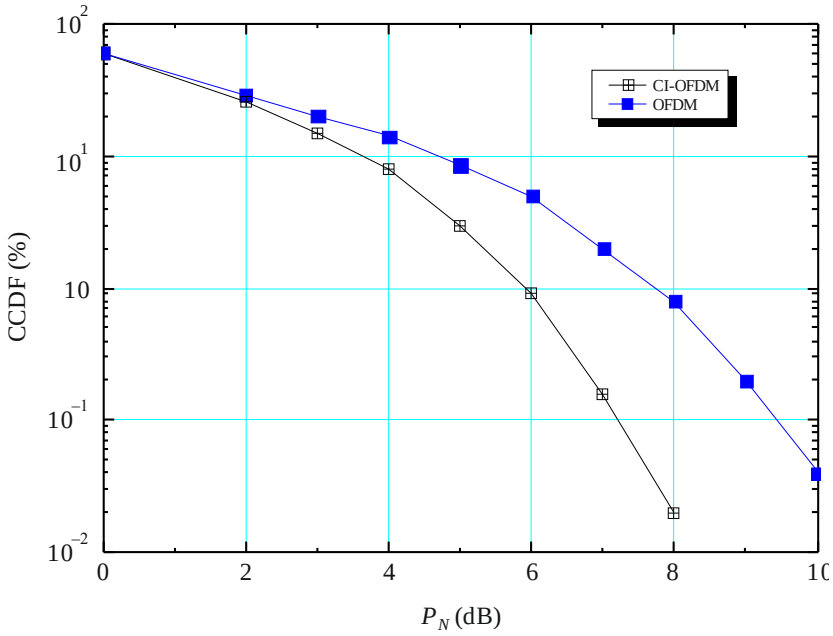
Figure 51 shows the PAPR complementary cumulative distribution function (CCDF) of 256-subcarrier OFDM and CI-OFDM systems operating with a 36 MHz bandwidth and transmitting uncoded 64-QAM constellations. The x-axis indicates the normalized instantaneous power, PN, relative to the average power level. In this case, the information rate is 203.4 Mbit/s. At a CCDF of 0.01%, a gain of about 2.5 dB is observed for the CI-OFDM over the OFDM system.

Figure 52 shows the TD for different OBOs of a 256-subcarrier OFDM and CI-OFDM system operating using a 36 MHz bandwidth and transmitting rate 3/4 FEC coded 64-QAM constellations. In the non-linear region of the TWTA, it is observed that the CI-OFDM system operates about 4 dB better than the OFDM system. In this case, the information rate is 152.5 Mbit/s.

These results show that not only does CI-OFDM have a better PAPR characteristic, but it also shows more tolerance to distortion in the TWTA non-linear region than OFDM.

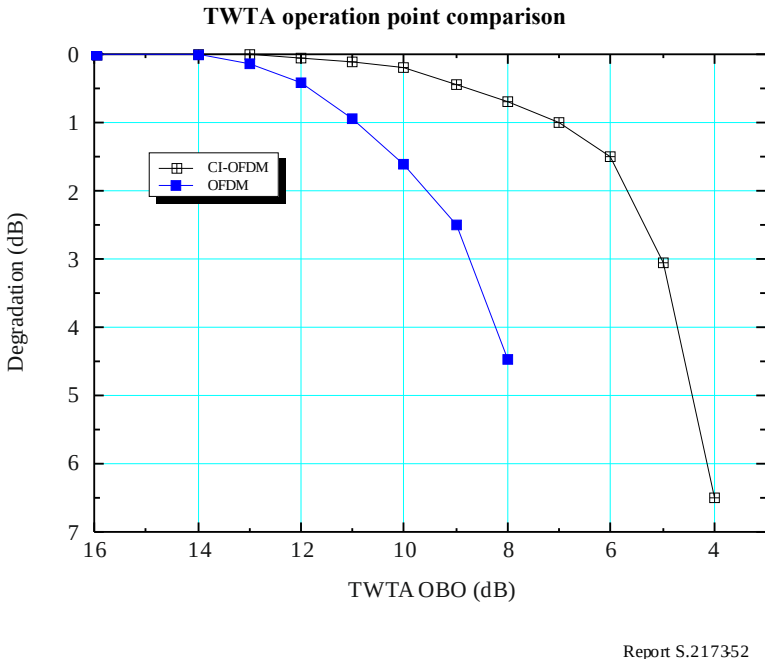
FIGURE 51

PAPR performance comparison



Report S.2173-51

FIGURE 52



10.4 Performance and spectral efficiency of CI-OFDM in a non-linear satellite channel

An implementation of a MCSS which uses CI-OFDM is studied in the section. This implementation is studied for the case where a SSPA is used in the ground terminal and a linearized TWTA (L-TWTA) is used in the satellite. The results of this implementation are then compared with those presented in § 10.1.

10.4.1 System model

Figure 53 shows the system model for a multi-carrier satellite system. Input data is buffered and passed on to a DVB-S2 encoder, whose rate is set by the ACM controller. The encoded data is then passed on to a bank of N symbol mappers, where N is the number of subcarriers used by the multi-carrier signal generator (MSG). The MSG is composed of two blocks for simulation purposes: An OFDM signal generator and a CI-OFDM signal generator. Only one MSG block is used during simulation. The output of the MSG is oversampled by a factor of 4 so as to get an accurate representation of the modulated signal and is passed on to the ground-station HPA. This HPA is assumed to be linearized and sufficiently backed-off as to avoid clipping noise due to driving the amplifier beyond saturation. The HPA output is then passed on to an analogue signal up-converter (U/C) that creates an analogue signal from the digital baseband signal at a desired carrier frequency and transmits it on the channel. The satellite, which is assumed to be bent-pipe, receives the noise-corrupted transmitted signal, amplifies and re-transmits it. A TWTA is often used for satellite transponders. Symbol predistortion is used by the multi-carrier satellite system to linearize the output of the TWTA. Note that many modern satellites are now being manufactured using L-TWTAs, and that the combination of a symbol precoder with a TWTA is essentially a L-TWTA. The amplitude-to-amplitude (AM/AM) and amplitude-to-phase (AM/PM) characteristics of the TWTA used in simulation are:

$$F_A(\mathrm{x}) = \frac{2\,|\,\mathrm{x}\,|}{(1+|\,\mathrm{x}\,|^2)} \tag{20}$$

$$F_{\varphi}(x) = \varphi_0 \frac{2|x|^2}{(1+|x|^2)}, \varphi_0 = \frac{\pi}{6} \quad (21)$$

where $F_A(x)$ and $F_{\varphi}(x)$ are the AM/AM and AM/PM characteristics – based on the Saleh model [78] – respectively. The symbol predistorter is characterised using a least-squares polynomial, obtained using 20 samples and a degree-10 polynomial [79]. This large degree polynomial essentially makes the L-TWTA an ideal limiter.The receiver receives the transmitted analogue signal, corrupted by Gaussian noise, passes it through a LNA and down-converter (D/C), and passes it on to either a signal sampler or a channel estimator. The received signal is passed on to the channel estimator if pilot signals are transmitted. The channel estimator estimates the instantaneous CNR, selects an appropriate MODCOD and relays the MODCOD information back to ACM controller at the transmitter. When data is received by the receiver, the signals are passed on to the signal sampler, which creates a set of samples, sampled at the Nyquist rate, for the multi-carrier processing unit (MPU). The MPU is composed of two modules for simulation: an OFDM processing unit and a CI-OFDM processing unit. The receiver uses the MPU module corresponding to the MSG module used by the transmitter. The MPU output is then passed on to a bank of symbol demappers, whose constellation is set by the channel estimator during MODCOD selection. The bank of symbol demappers use the average received constellations of each modulation and their respective error vector magnitudes to create log-likelihood ratios for each transmitted bit which they pass on to the DVB-S2 decoder. The decoder outputs a decision on the transmitted data after 100 decoding iterations. In simulation, error-free MODCOD feedback and channel estimation is assumed.

10.4.2 Test results

OFDM signals are widely known to have high PAPR, which can be irreparably distorted by a satellite HPA. In this section, the MCSS model described in Fig. 53 is evaluated with and without CI-OFDM technology.

It is not sufficient to evaluate PAPR mitigation techniques solely based on their ability to reduce the PAPR of generated signals [80]. Some PAPR reduction techniques, in fact, introduce more distortion into the signal while reducing the PAPR. An example of this is signal clipping. A fairer way to evaluate the performance of a PAPR reduction technique is total degradation (TD) [80]. TD is the difference, in dB, between the CNR required for a system with a nonlinear HPA to obtain a particular PER and that of a system with an ideal linear HPA – henceforth referred to as a linear HPA[17]– to obtain the same PER performance, plus the IBO of the nonlinear HPA. Mathematically this is:

$$TD \text{ (dB)} = CNR_{nonlinear} \text{ (dB)} - CNR_{linear} \text{ (dB)} + IBO \text{ (dB)} \quad (22)$$

where CNR_{linear} and $CNR_{nonlinear}$ are the CNRs required to obtain a particular PER for the linear and nonlinear HPA respectively.

17 Note that a distinguishing difference between a linear HPA and a L-TWTA is that the linear HPA has no saturation point at which it clips the signal.

FIGURE 53

System model for MCSS

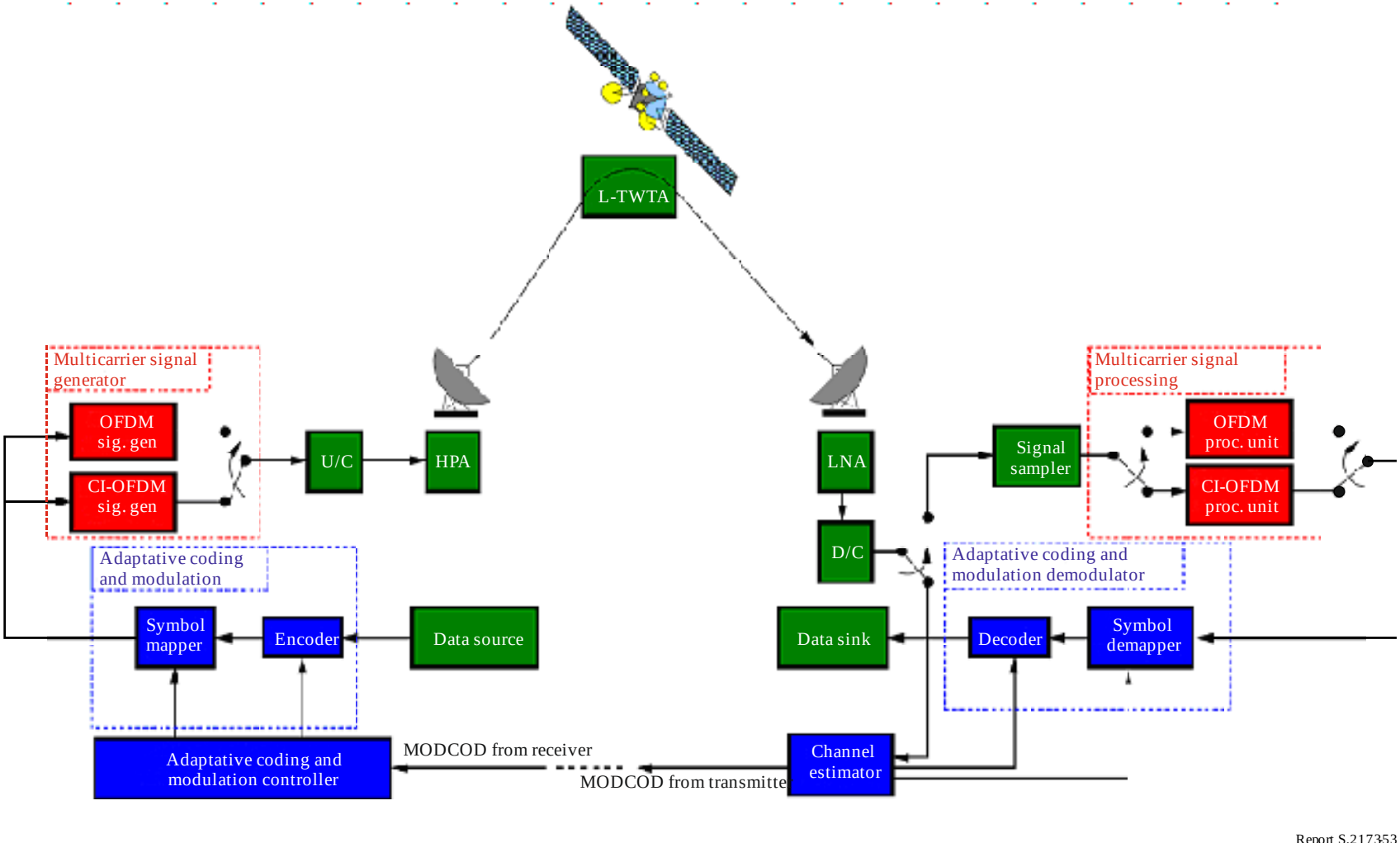
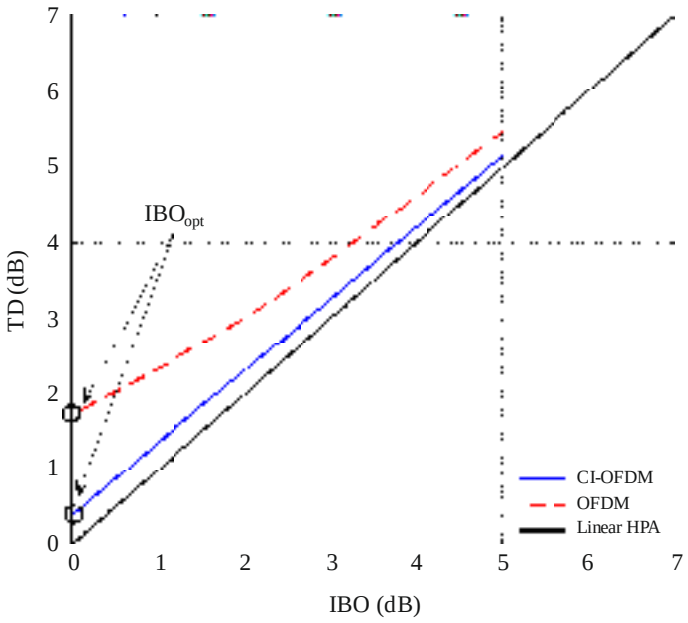


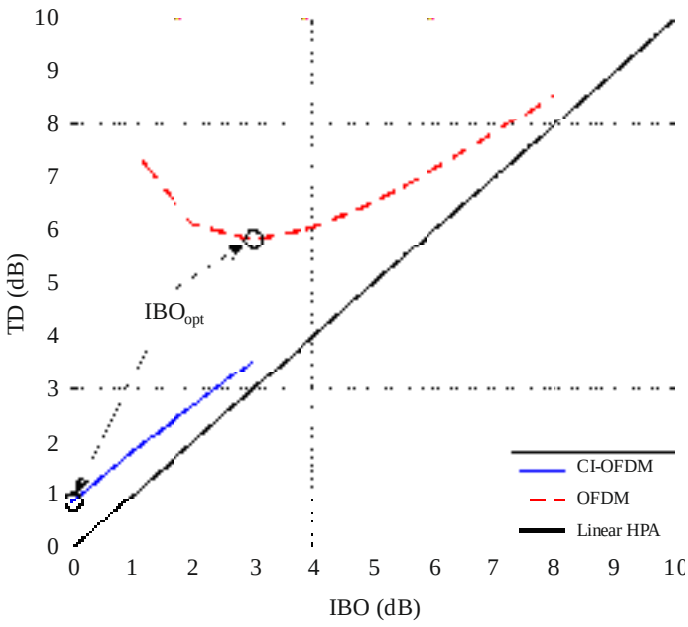
Figure 54 shows the TD performance of different MODCODs for 64-subcarrier OFDM and CI-OFDM systems with a PER of 10-3. Figure 54 makes clear that the CI-OFDM system has much better performance than the OFDM system due to its PAPR reduction capability; with gains of 1.35 dB, 4.97 dB, 4.96 dB and 3.05 dB for the QPSK 5/6, 8-PSK 9/10, 16-APSK 9/10, 32-APSK 9/10 CI-OFDM systems over the OFDM systems respectively (in terms of the difference between their minimal TDs). Using curves such as those produced in Fig. 54, an optimal IBO can be approximated. The optimal IBO (IBO_{opt}) is the IBO value that minimizes the TD [80]. For example, from Fig. 54(b), IBO_{opt} = 0 dB for the CI-OFDM system and IBO_{opt} = 3 dB for the OFDM system.

FIGURE 54

Total degradation of OFDM and CI-OFDM for different MODCODs @ PER = 10^{-3}

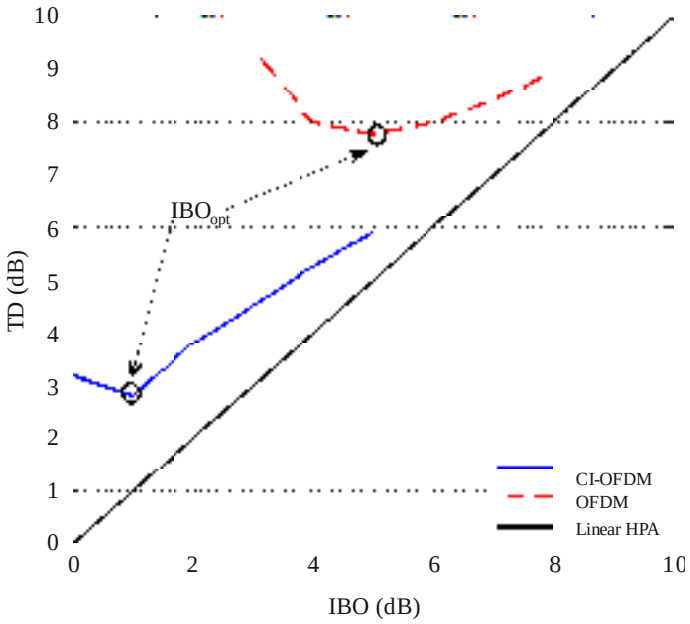


a) 64-subcarrier QPSK 5/6

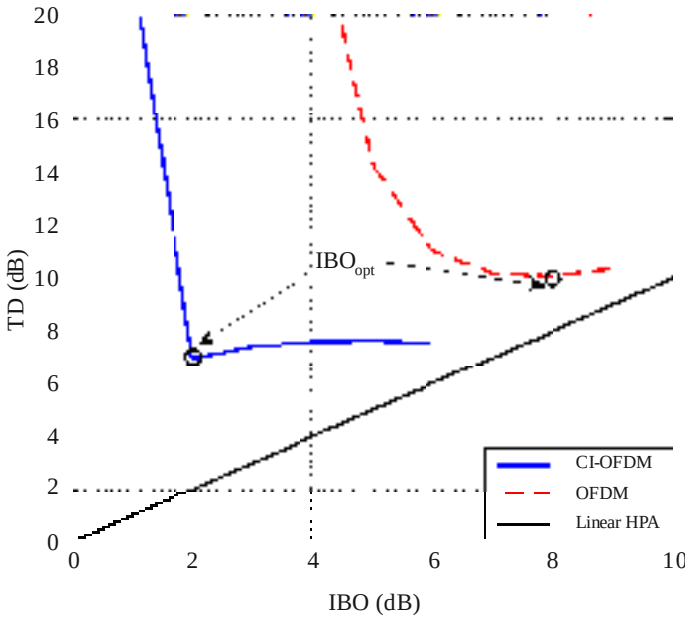


b) 64-subcarrier 8-PSK 9/10

Report S.2173-54a,b



c) 64-subcarrier 16-APSK 9/10



d) 64-subcarrier 32-APSK 9/10

Report S.2173-54c-d

Tables 8 and 9 compare the TD and CNR performance of a 64-subcarrier SCSS with linear HPA; SCSS with L-TWTA; MCSS with L-TWTA OFDM transmissions and MCSS with L-TWTA and CI-OFDM transmissions. These results are obtained for a PER of 10^{-3} operating using different DVB-S2 MODCODs, where the system was operated at IBO_{opt}.

TABLE 8

Degradation due to L-TWTA for a satellite system using various combinations of DVB-S2 MODCOD

MODCOD	Spectral efficiency (Bit/s/Hz)	Linear amplifier	L-TWTA	
		CNR_{eq} (dB) @ PER = 10^{-3}	CNR_{eq} (dB) @ PER = 10^{-3}	TDL-TWTA (dB)
QPSK 1/4	0.49	–2.96	–2.96	0
QPSK 2/5	0.79	–0.64	–0.64	0
QPSK 1/2	0.99	1.13	1.13	0
QPSK 5/6	1.65	5.05	5.05	0
8-PSK 3/5	1.78	5.61	5.61	0
8-PSK 3/4	2.23	7.84	7.84	0
8-PSK 5/6	2.48	9.31	9.31	0
8-PSK 9/10	2.68	10.84	10.84	0
16-APSK 3/4	2.96	10.14	10.21	0.07
16-APSK 4/5	3.16	10.92	11.00	0.08
16-APSK 5/6	3.30	11.53	11.63	0.10
16-APSK 8/9	3.52	12.76	12.88	0.12
16-APSK 9/10	3.56	12.99	13.13	0.14
32-APSK 3/4	3.70	12.80	13.48	0.68
32-APSK 4/5	3.95	13.61	14.45	0.84
32-APSK 5/6	4.12	14.26	15.20	0.94
32-APSK 8/9	4.39	15.50	16.70	1.20
32-APSK 9/10	4.45	15.75	16.98	1.23

TABLE 9

Comparison of TD performance for MCSS with OFDM and CI-OFDM transmissions using various combinations of DVB-S2 MODCOD

MODCOD	Spectral efficiency (Bit/s/Hz)	OFDM			CI-OFDM		
		CNR_{eq} (dB) @PER = 10^{-3}	IBO_{opt} (dB)	TD (dB)	CNR_{eq} (dB) @PER = 10^{-3}	IBO_{opt} (dB)	TD (dB)
QPSK 1/4	0.49	–2.29	0	0.67	–2.78	0	0.18
QPSK 2/5	0.79	0.16	0	0.80	–0.44	0	0.20
QPSK 1/2	0.99	1.73	0	0.60	1.23	0	0.10
QPSK 5/6	1.65	6.78	0	1.73	5.43	0	0.38
8-PSK 3/5	1.78	8.12	0	2.51	6.01	0	0.40
8-PSK 3/4	2.23	11.17	0	3.33	8.29	0	0.45

TABLE 9 (*end*)

MODCOD	Spectral efficiency (Bit/s/Hz)	OFDM			CI-OFDM		
		CNR_{eq} (dB) @PER = 10^{-3}	IBO_{opt} (dB)	TD (dB)	CNR_{eq} (dB) @PER = 10^{-3}	IBO_{opt} (dB)	TD (dB)
8-PSK 5/6	2.48	13.93	1	4.62	9.95	0	0.64
8-PSK 9/10	2.68	16.69	3	5.85	11.72	0	0.88
16-APSK 3/4	2.96	15.41	2	5.27	11.53	0	1.39
16-APSK 4/5	3.16	16.79	3	5.87	12.59	0	1.67
16-APSK 5/6	3.30	18.08	3	6.55	13.56	0	2.03
16-APSK 8/9	3.51	20.04	5	7.28	15.42	1	2.66
16-APSK 9/10	3.56	20.76	5	7.77	15.81	1	2.82
32-APSK 3/4	3.70	20.40	5	7.60	16.26	2	3.46
32-APSK 4/5	3.95	22.05	6	8.13	17.47	2	3.55
32-APSK 5/6	4.12	23.16	6	8.9	18.55	2	4.29
32-APSK 8/9	4.39	25.43	8	9.93	21.81	2	6.31
32-APSK 9/10	4.45	25.81	8	10.06	22.75	2	7.00

Table 8 demonstrates the TD caused by passing a different DVB-S2 modulation through a L-TWTA, obtained at a PER of 10^{-3}. Note that to properly compare the CNR of the linear HPA with the CNR of the system with L-TWTA, the equivalent CNR is:

$$CNR_{eq}\ (\text{dB}) = CNR\ (\text{dB}) + IBO_{opt}\ (\text{dB}) \tag{23}$$

This conversion must be done to fairly compare the performance of both systems, operating at their maximal output power. The linear HPA always operates at 0 dB IBO (HPA saturation), whereas the L-TWTA is not necessarily operated at saturation. Simulation results for the SCSS with L-TWTA specify that IBO_{opt} at which to operate the L-TWTA is 0 dB. For constant-envelope modulation such as M-ary PSK there is no degradation; however the degradation for 16-APSK is negligible, while there is a noticeable degradation for 32-APSK. Table 8 demonstrates that a SCSS can operate using the DVB-S2 with very little loss when compared to the theoretical system with linear amplifier.

Table 9 shows that the change in TD for the MCSS systems is far more dramatic than when compared to the SCSS systems. This is because of the high PAPR of multi-carrier signals. It can also be observed that the MCSS with CI-OFDM transmissions has between 0.5 and 4.5 dB gain in terms of TD over the MCSS with OFDM transmissions depending on the MODCOD employed.

Figure 55 demonstrates this behaviour by plotting TD with respect to the spectral efficiency (in bits per second per hertz (bit/s/Hz)) of the DVB-S2 ACM scheme. Note that the results are presented in terms of CNR_{eq} – as calculated in (23) – for each MCSS system. Also note that the curves are plotted using the maximum spectral efficiency generated by all MODCODs at each CNR_{eq} for a particular system. That is, if MODCOD *x* has higher spectral efficiency than MODCOD y, and MODCOD x has lower CNR_{eq} than MODCOD y, then MODCOD y is omitted from Fig. 55. MODCODs not included in Fig. 55 are underlined in Tables 8 and 9. It can be observed that the curve representing the MCSS using OFDM transmission has a much steeper ascent than the MCSS using CI-OFDM transmissions. In fact, up to a spectral efficiency of 3.6 bit/s/Hz, the MCSS with

CI-OFDM transmissions has a TD less than 3 dB. This means that the MCSS with CI-OFDM transmissions could be employed for spectral efficiencies of up to 3.6 bit/s/Hz, at no more than double the required transmission power.

Figure 56 gives the energy efficiency of the SCSS with linear HPA, SCSS with L-TWTA, MCSS with OFDM transmissions and MCSS with CI-OFDM transmissions, by plotting the spectral efficiency of these systems – at a PER of 10^{-3} – versus CNR_{eq} using Tables 8 and 9. Each step in a curve represents the use of a new MODCOD with higher spectral efficiency. Note that as was explained with Fig. 55, only the MODCODs giving a maximum spectral efficiency are used to plot Fig. 56. The results in Fig. 56 reflect those of Fig. 55, with the SCSSs having better energy efficiency than the MCSSs, especially for MODCODs having higher spectral efficiency. In particular, MCSSs using MODCODs with 32-APSK modulation have very poor energy efficiency when compared with the SCSS with L-TWTA. However, it is clear that the CI-OFDM greatly improves the energy efficiency of the MCSS when compared with the MCSS using OFDM transmissions. The spectral efficiency that can be used for a MCSS using OFDM transmissions, without more than doubling the transmission power, is 1.7 bit/s/Hz. This is almost 2 bit/s/Hz lower than it is for the MCSS using CI-OFDM transmissions.

FIGURE 55

Total degradation of SCSS and MCSS versus spectral efficiency

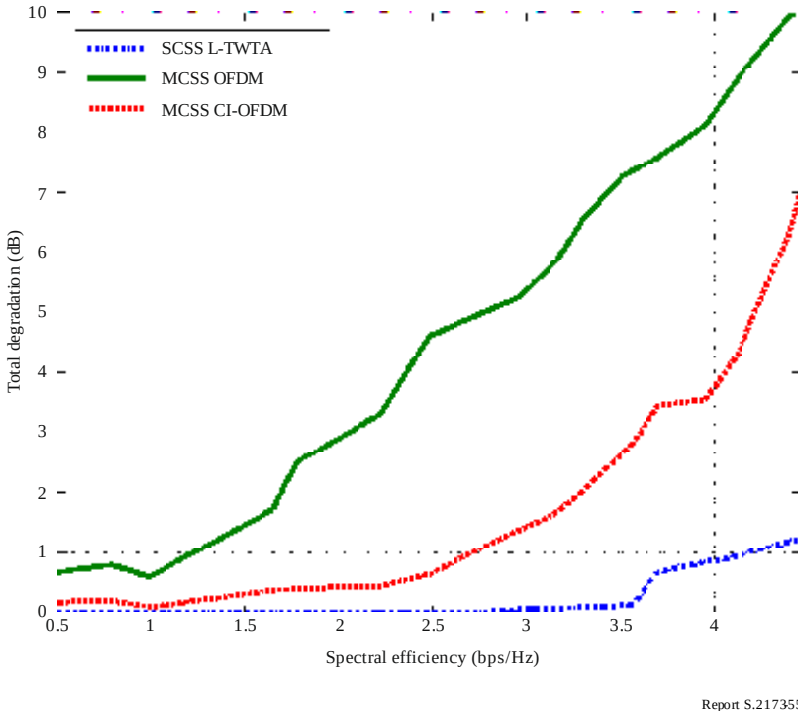
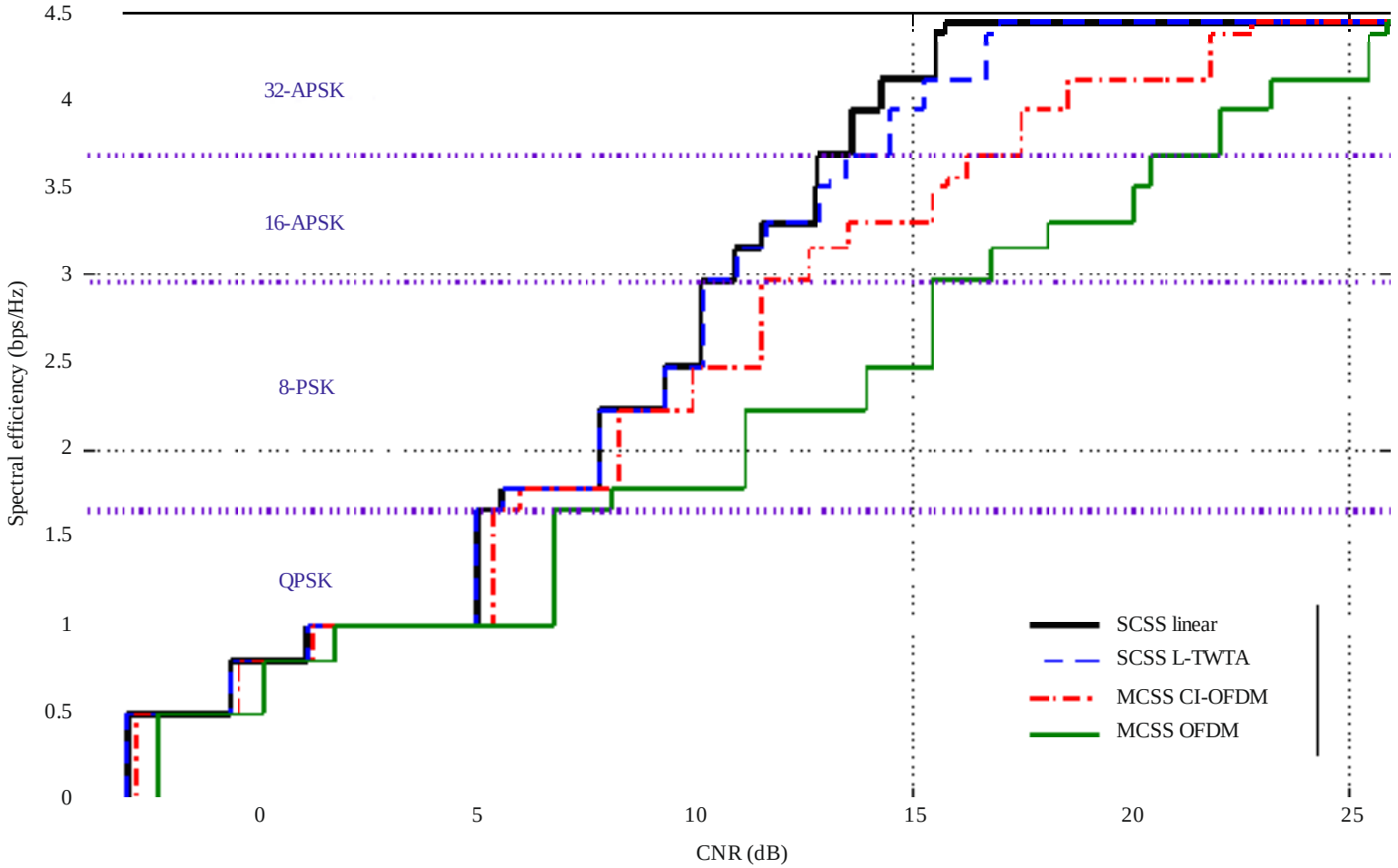


FIGURE 56

Energy efficiency of SCSS and MCSS for various combinations of DVB-S2 MODCOD



10.5 Performance of MC-CDMA in a non-linear satellite channel

The performance of the MC-CDMA for a satellite communication system is investigated in this section.

10.5.1 System model

Figure 57 shows the block diagram for the transmitter and receiver of an adaptive MC-CDMA satellite system. Adaptive MC-CDMA systems can be used as a countermeasure for atmospheric fading. The adaptive transmitter – as shown in Fig. 57 – changes the coding, modulation, and spreading schemes by a control command determined according to the satellite channel conditions. The channel encoder in the adaptive transmitter changes its encoding parameters using the aforementioned control command. The symbol spreader for each user consists of a symbol mapper, a symbol repeater, a Walsh spreader, and also changes its parameters using the control command. The chip interleaver in the transmitter is for dual mode MC-CDMA schemes, which are especially useful for changing the utilized spreading scheme according to the channel and traffic conditions. Adaptive MC-CDMA systems can also utilize techniques for reducing the PAPR. Additionally, a predistorter is used to linearize the HPA. Since this is a satellite communications system, a TWTA is assumed.

According to the adaptive operation of the transmitter, the receiver in Fig. 57 should also change its decoding, demodulating, and de-spreading schemes using the control command in synchronization with the command utilized by the transmitter. Multi-carrier demodulation is accomplished by a simple FFT, and all other operations performed by the receiver are applied in reverse-order with respect to those operations performed by the transmitter.

10.5.2 BER performance of non-linear MC-CDMA satellite system

The system parameters in Table 10 are used to investigate the signal distortions appearing in the symbol constellations due to the nonlinear transfer function of the TWTA in the MC-CDMA satellite system. For simulation purpose, the transfer function of the TWTA in equations (20) and (21) is used. Figure 58 shows signal distortions for various numbers of users and OBO values. It is assumed that the MC-CDMA system uses 128 subcarriers and a Walsh-Hadamard (WH) code of length 16. There are up to 16 active users in the system, and the 128 subcarriers are evenly allocated to each active user. In Fig. 58, the small red circles in the constellation represent the signal without distortion; the outer blue dots represent the signal when number of current users, $K = 16$; the green dots represent the signal when $K = 8$, and the inner black dots represent the signal when $K = 1$.

The effect of nonlinear distortion by the TWTA increases as the number of users increases. It is also evident that the signal distortion increases as the OBO value of TWTA decreases [29]. The signal distortion degrades the BER performance of the system.

FIGURE 57

Block diagram of the adaptive MC-CDMA satellite system

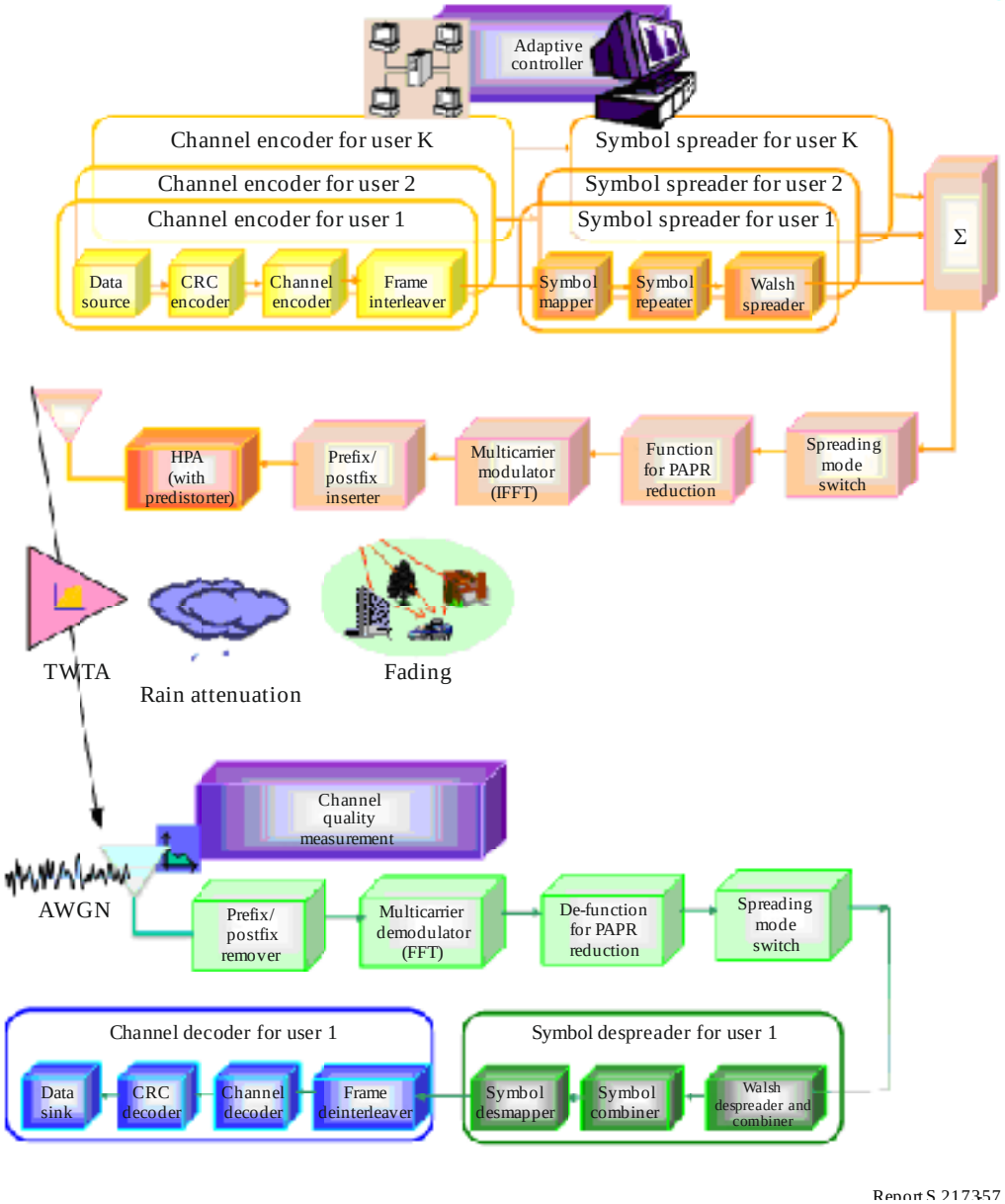


TABLE 10

Parameters for MC-CDMA satellite system

Signal constellation	QPSK, 16-QAM
Spreading sequence	Walsh-Hadamard
Processing gain	16
Number of symbol per frame (M)	4, 8, 16
Number of active user (K)	1-16
Number of subcarrier (N)	64, 128, 256
Scrambling code	Random code

FIGURE 58

Signal constellation distorted by the TWTA

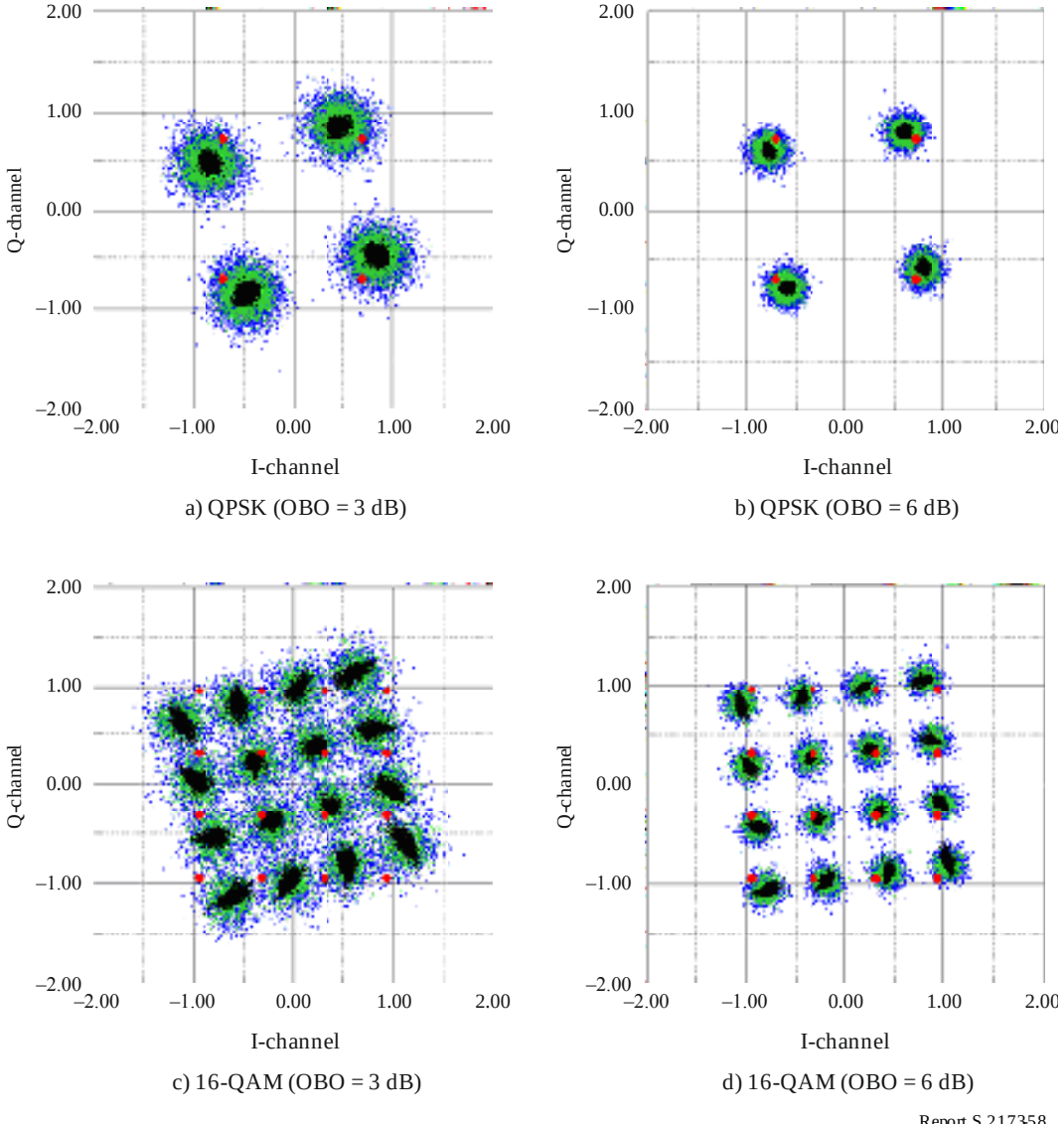
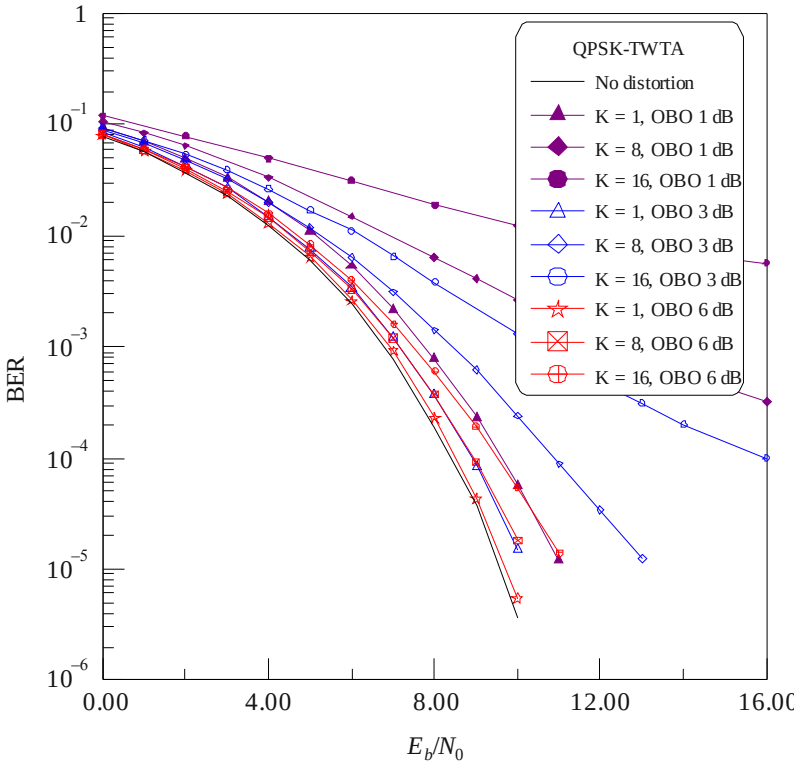


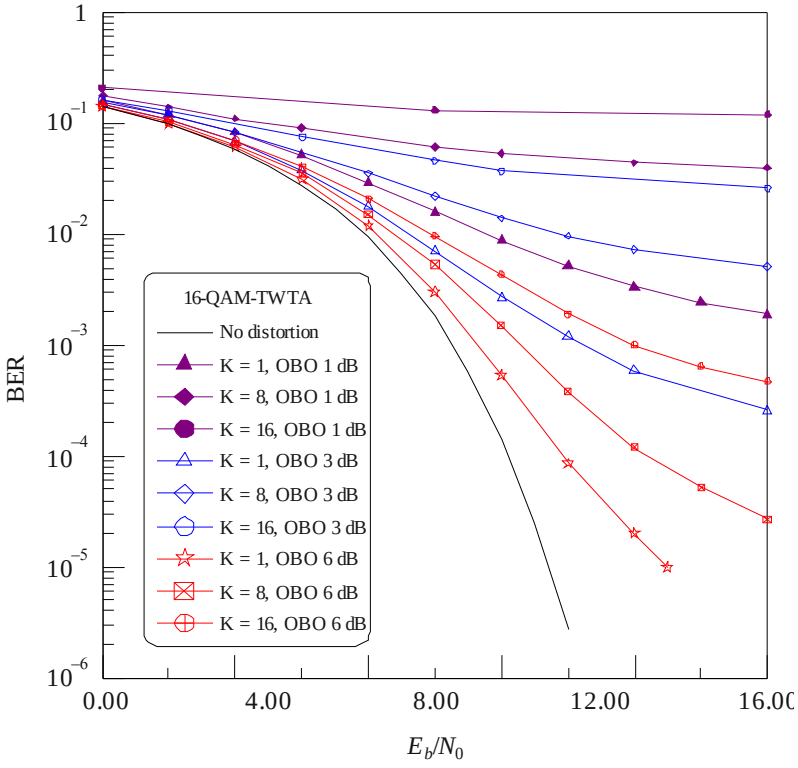
Figure 59 examines the BER performance of a MC-CDMA satellite system with various numbers of users and OBO values. The simulation results in Fig. 59 show that the nonlinear distortion has a larger impact on the 16-QAM constellation than the QPSK constellation. It is also noticeable that the effect of nonlinear distortion on the BER performance of the MC-CDMA satellite system increases as the number of users increase and the OBO values of TWTA decrease – as is suggested by Fig. 58.

FIGURE 59

BER performance of the MC-CDMA satellite system



a) QPSK



b) 16-QAM

Report S.2173-59

Using a symbol predistorter can mitigate the performance degradations caused by the nonlinear distortion of the TWTA. Figure 60 shows the uncoded BER performance of the MC-CDMA satellite system when combined with a L-TWTA, which is the combination of a TWTA and an ideal predistorter – henceforth referred to as an ideal L-TWTA. The results presented demonstrate considerable improvement in the BER performance of the MC-CDMA satellite system with ideal predistortion; especially for the case where larger OBO values are observed. The improvement in BER performance is greater with the 16-QAM constellation because 16-QAM is experiences greater nonlinear distortion than QPSK. The improvement in performance is the result of phase distortion compensation and linearizing effect of the predistorter in the regions before HPA saturation. Although simulations results are presented using an ideal predistorter, the high-peak problems of MC-CDMA cannot be fully resolved. This is because the ideal predistorter linearizes signals only in the regions before saturation, and the saturation region is determined by the OBO value. Therefore, it is very important to reduce the PAPR by using an efficient PAPR reduction method. Doing so can reduce the OBO value and results in efficient power usage.

10.5.3 Adaptive operation of MC-CDMA satellite system

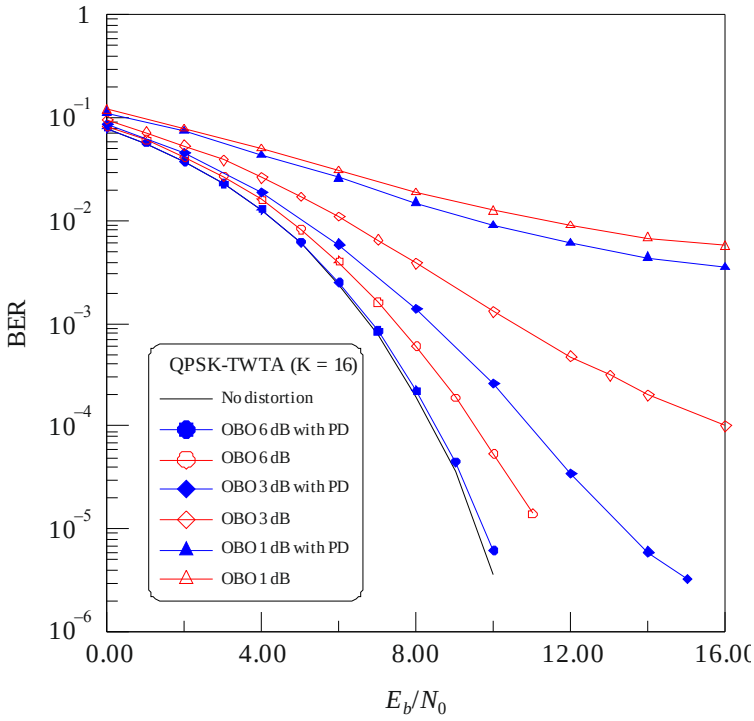
In this section, the MC-CDMA satellite system is investigated when implemented in a mobile satellite channel model for suburban Ka-Band [81]. Various BTC MODCODs from Table 4 are used for adaptive transmission. E_s/N_0 is fixed at 20.5 dB which corresponds to the required E_s/N_0 value of 18.5 dB for BTC MODCOD of 64-QAM $(63,56)^2$ to produce a BER of 10^{-6} as shown in Table 3, plus an additional 2 dB power margin. The results are shown in Table 11, which compares the performance of an adaptive MC-CDMA satellite system with that of a conventional non-adaptive MC-CDMA satellite system. In both simulation systems, an ideal L-TWTA with IBO of 3 dB is used.

It is clear that the performance of the adaptive MC-CDMA satellite system is superior to that of the non-adaptive MC-CDMA system. For example, the adaptive scheme can achieve a spectral efficiency[18] of 1.97 bit/s/Hz, which corresponds to roughly the same spectral efficiency as the 8-PSK (31,25)2 MODCOD, but with far better BER performance for the same E_s/N_0.

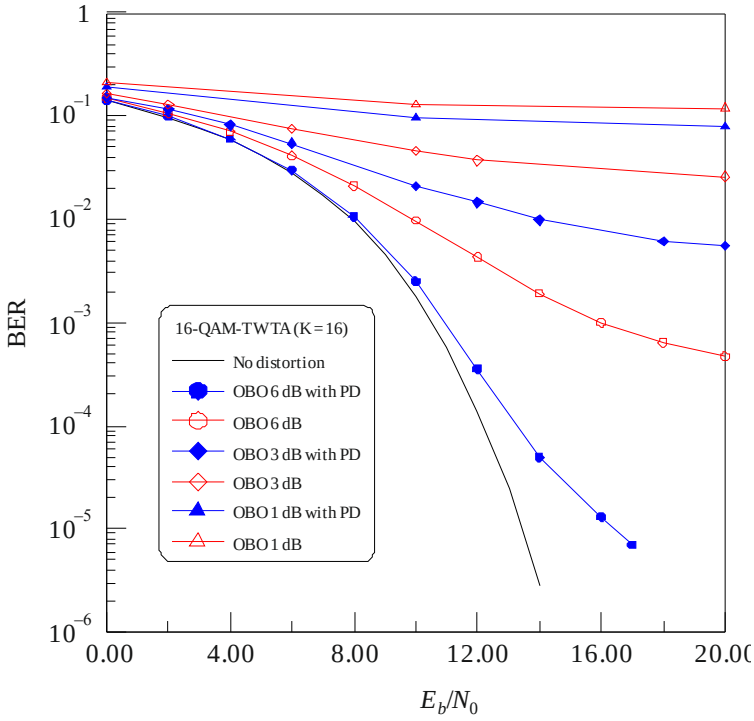
[18] Where the definition of spectral efficiency in this context is the average spectral efficiency found by averaging the results from Table 11 over all instantaneous channel realizations generated using the channel model from [81]. Note that this does not consider packet errors or retransmission.

FIGURE 60

BER performance of the MC-CDMA with the L-TWTA with an ideal predistorter



a) QPSK (K = 16)



b) 16-QAM (K = 16)

Report S.2173-60

TABLE 11

The performance of the adaptive MC-CDMA in mobile satellite channel

	Adaptive	**8-PSK (31,25)[2]**	**8-PSK (15,10)[3]**	**BPSK (15,10)[3]**
BER	8.91×10^{-6}	1.15×10^{-1}	8.78×10^{-2}	2.09×10^{-2}
PER	3.10×10^{-4}	4.59×10^{-1}	2.81×10^{-1}	8.31×10^{-2}
Spectral efficiency (bit/s/Hz)	1.97	1.95	0.89	0.30

11 Future trends (on-board processing)

11.1 Introduction

Processing capability on-board satellites is a technology that is getting more and the increasing adoption of packet-switched networks. Traditional satellite communication systems are circuit-switched systems that require a link – allocated with a specific amount of bandwidth – established between a HUB and user terminal before communication can take place. Often this link is underutilized as it runs at the data rate required by that particular application, even though several services could utilize the same bandwidth simultaneously. Packet-switched networks allows for better utilization of resources as bandwidth can be assigned bit-by-bit, rather than Hz-by-Hz. The use of on-board processing also allows for dynamic IP-routing, which can allow for packets to be routed more efficiently over the network, possibly avoiding unnecessary latency-inducing hops over multiple links. This section describes the advantages brought by on-board processing, as well as, some implementation considerations and examples.

11.2 Signal regeneration

The use of on-board processing can enable the use of signal regeneration, which can be of varying benefits depending on how complex the on-board processor is. The simplest example of signal regeneration is to include a baseband demodulator and re-modulator on the satellite, as depicted in Fig. 61. The uplink transmitted symbol, s_u, is received by the satellite as a received sample, r_u, which comprises a faded copy of s_u corrupted by noise. The satellite then regenerates a copy of s_u (s_d) by choosing the constellation point from the employed modulation that has the minimum Euclidean distance from the received sample. Figure 62 compares the BER performance of a signal regeneration satellite (SRS) with that of a traditional bent-pipe satellite (BPS), where both systems use QPSK and have FEC at the receiver, for various downlink and uplink noise scenarios.

Note that CNR_u is the uplink CNR and CNR_d is the downlink CNR. The FEC used by the receiver is the rate-compatible 3GPP2 LDPC code, which is fixed to be a (768, 384) LDPC code and gives a coding rate of 1/2. Figure 62 demonstrates that the use of a simple signal regeneration scheme can provide as much as roughly 0.75 dB (when CNR_u = 8 dB) of downlink CNR gain at a BER of 10^{-4}. Note that the effective CNR, CNR_{eff}, of the channel (uplink and downlink), as defined by the ITU [82] for a satellite link is

$$CNR_{eff} = (1/CNR_u + 1/CNR_d)^{-1} \quad (24)$$

Through manipulation of equation (24) it can be found that the 0.75 dB gain in downlink CNR translates into a gain of roughly 0.55 dB in terms of the effective CNR. It should be noted that the gains of SRS systems with respect to BPS systems are usually evaluated for system without FEC at

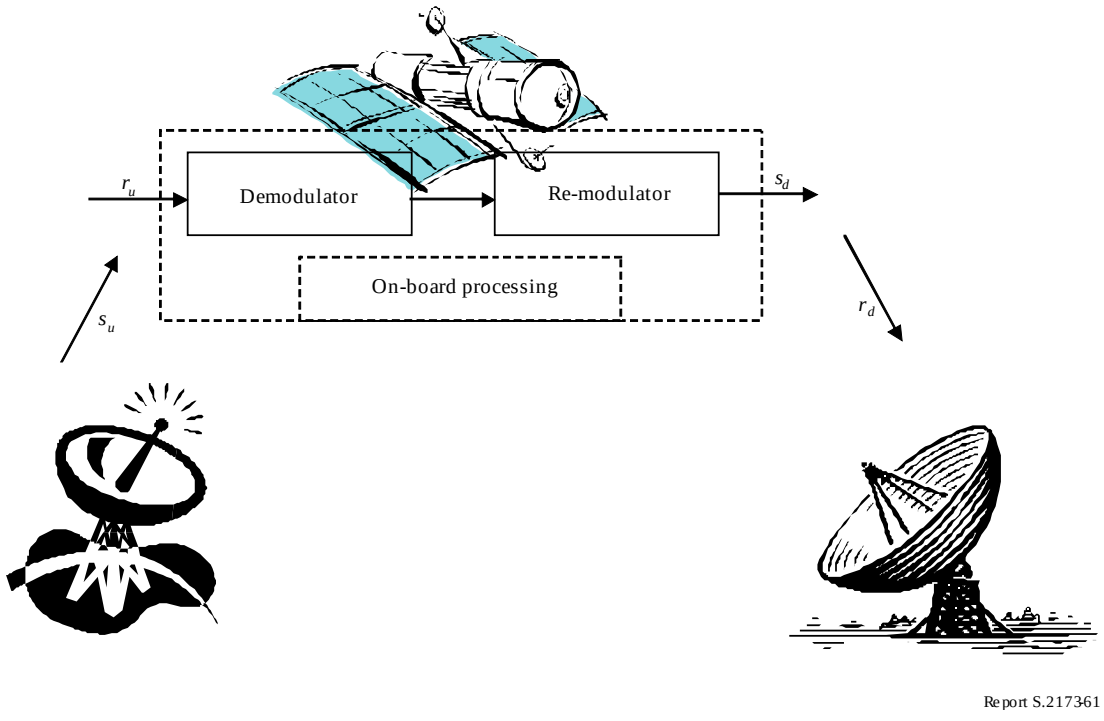
the receiver. A 3 dB gain in the SNR_{eff} has been claimed for the uncoded SRS with respect to the uncoded BPS [83, 84]. This is misleading, as all digital satellite standards include the use of FEC in both the forward and return link. However, the benefits of a SRS are still clear from Fig. 62, even if not as pronounced as it is for the uncoded systems. Figure 62 also demonstrates that there is an inherent error-floor for a SRS system (e.g. CNR_u < 8 dB in Fig. 62) when the uplink CNR is not high enough. This is due to the fact that the on-board processor is also prone to making errors in the same way that the receiver would. As a result, if the uplink CNR is not high enough, the wrong constellation point may be re-modulated, resulting in error-propagation through to the receiver. An error-floor is also introduced into the BPS system when the CNR_u is too low. However, its introduction is less sharp as there is no error-propagation due to incorrect re-modulation. This explains why the BPS system begins to outperform the SRS system for higher SNRs (for SNR_u < 8 dB in Fig. 62). Yet, it should be noted that there are still error-floors for BPS system. In fact, both the SRS and BPS systems are limited by CNR_u:

$$\lim_{CNR_d \to \infty} CNR_{eff} = \left(1/CNR_d + 1/CNR_u\right)^{-1} = \left(1/\infty + 1/CNR_u\right)^{-1} = CNR_u \tag{25}$$

However, it should be noted that uplink power control is used to establish a CNR_u that will be sufficient as to avoid unnecessary error-floors, meaning that a satellite systems is limited by CNR_d.

FIGURE 61

Satellite with on-board demodulator and re-modulator



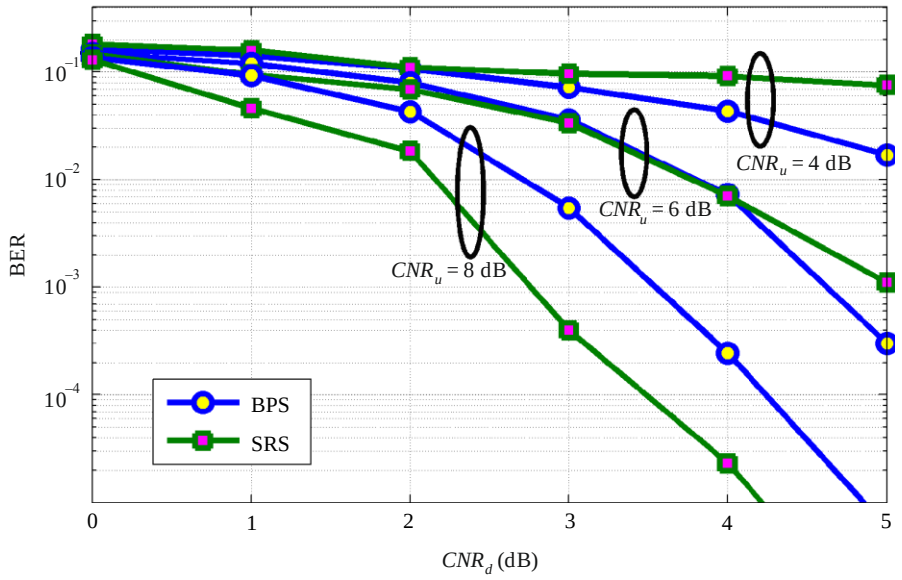
Report S.2173-61

Signal regeneration schemes that simply re-modulate the samples received by the satellite ignore the potential improvements that could be obtained by including FEC as part of the signal regeneration scheme. That is, the addition of a FEC decoder on-board the satellite creates an even more reliable satellite link. This is achieved by replacing the on-board processor of Fig. 61 with the on-board processor in Fig. 63. For the on-board processor of Fig. 63 the set of samples, r_u, representing the uplink symbols, s_u, are demodulated to create either a vector of inputs, λ_u, which are either hard-decisions on or soft-estimates of the uplink data. This information is then passed to a decoder, which could be of any type, as long as it corresponds to the code employed by the satellite

communication system. That is, the decoder could be an algebraic decoder, a soft-decision decoder, or even an iterative decoder, depending on the whether λ_u are hard-decisions or soft-estimates. The decoder output, $\mathbf{c}_u$, is a vector representing a hard-decision on the codeword[19], which is passed to the re-modulator to regenerate the symbol vector representing the uplink codeword.

FIGURE 62

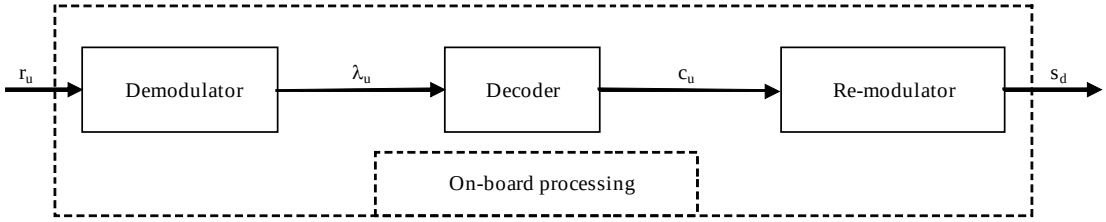
BER performance comparison of SRS and BPS systems decoded by 3GGP2 LDPC code and transmitting QPSK



Report S.217362

FIGURE 63

On-board signal regenerator with FEC



Report S.217363

Figure 64 demonstrates the additional gain that can be obtained by using the on-board processing described by Fig. 63. The results in Fig. 64 are obtained where the receiver used by both satellite systems (SRS and BPS), as well as, the on-board processor of the SRS uses a LDPC decoder; and QPSK is used as modulation. All decoders use 50 iterations and the same 3GPP2 LDPC code that was used to generate the results in Fig. 62. Note that "SRS with FEC" denotes a SRS system with on-board FEC, as depicted in Fig. 63. Additionally, "SRS no FEC" is the simple SRS system depicted in Fig. 61. A downlink CNR gain of 0.7 dB can be observed for a BER of 10^{-4} for the SRS

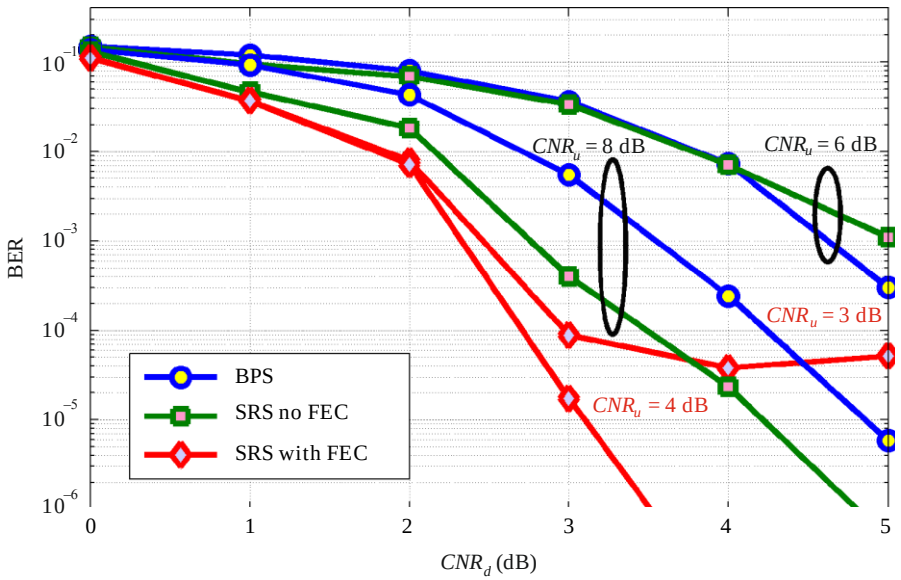
[19] Note that if the decoder output must be a hard-decision on the original message-word, an encoder can be placed after the decoder to create the hard-decision on the codeword, $\mathbf{c}_u$.

system with on-board FEC with respect to the SRS system without on-board FEC; however, this comparison is unfair due to the fact that these curves are obtained for largely different values of CNR_u. A more fair comparison, which uses the effective CNR, reveals a gain of 1.82 dB for the SRS with FEC over the SRS without FEC. This translates into a gain of 2.37 dB for the SRS with FEC over the BPS system. This example suggest that the use of an on-board FEC scheme, as depicted in Fig. 63, can significantly increase the power efficiency of a satellite system.

Again, there is a noticeable error-floor in Fig. 64 for lower values of CNR_u. As was stated formerly in this section, if CNRu is not large enough, errors will be made in decoding and the wrong constellation points will be re-modulated, leading to error-propagation; although, it should be emphasized that the system would not be operated for such low values of CNR_u. Additionally, for the forward link, uplink power control can be used to ensure that the transmitted signal has high enough CNR.

FIGURE 64

BER performance of SRS system with FEC



Report S.217364

11.3 Reducing latency: IP-routing and caching

On board processors can include the use of IP-routing. The use of IP-routing on satellite can greatly reduce the latency incurred using satellite systems. This is extremely important as satellite service providers attempt to enable interactive next-generation communication services, such as video conferencing, as well as, entice possible customers with service comparable to those offered by terrestrial-based competitors.

The use of IP-routing can avoid unnecessary additional hops over the satellite. Take for example the scenario of an oil rig in the Gulf of Mexico trying to do video conferencing with headquarters in South Carolina over a traditional BPS. The ship's satellite dish would have to transmit up to the satellite, and the satellite would send the signal to the teleport (which could be in Louisiana in the US) through a particular transponder servicing that region. From there the signal is routed back up through the satellite and down to headquarters via a different transponder, establishing a link. Since the network is circuit-switched, the established "two-hop" path is used by the satellite network to maintain the link, which introduces a more delay in the video feed. In the former scenario, two hops over the same satellite were required. If IP-routing were available on-board the satellite, the satellite

system could have routed the call directly to headquarters single hop, reducing the latency by one half.

A broadband satellite with on-board memory could make use of caching, whereby previously requested or popular content is stored indefinitely in the satellite's memory. In the event that information in the satellite's cache is requested, the satellite could transmit this information directly to the user, cutting link latency in half and avoiding any additional delay incurred from data access at the HUB station (e.g. Internet access by the HUB to be returned via satellite to the end-user). The use of an on-board cache may be attractive for news applications delivered via satellite, where service providers could pre-load updated news content from media outlets periodically, giving their users faster access to media outlets that are willing to pay a small premium.

11.4 Flexible signals, flexible design: variable data rates, cross-layer optimization and software-defined radio

On-board processing can allow for better spectral efficiency and the use of quality of service (QoS) protocols. Since on-board processing enables the use of packet-switching, all data transmitted to the satellite can be flexibly allocated to a particular beam, at a particular data-rate. This means that the data-rate for a particular packet on the uplink may not be the same on the downlink. The satellite with on-board processing attempts to make the best use of its resources (e.g. spot beams, frequencies, time), which means that if, for example, satellite resources are underutilized, the satellite can assign more capacity to a given user. However, it also means that if there is congestion, the data-rate can be appropriately downscaled such that links between the HUB and user terminals are maintained. Additionally, the capability to flexibly assign data-rates allows the satellite to use QoS to assign the appropriate resources, such as capacity, required to run a particular application. This QoS flexibility could also be used to offer customers different grades of service; where if a customer is willing to pay more, more capacity can be allocated to them.

A satellite with on-board processing could facilitate the use of cross-layer optimization protocols on-board the satellite. Layers within the satellite system's protocol stack could communicate and share information in a non-hierarchical manner (as is dictated by traditional networking models such as OSI) in such a way that improves the system's overall performance. For example, the use of an ARQ scheme on a satellite that has on-board FEC could prevent the satellite from re-transmitting erroneous packets. In the event that an erroneous packet is forwarded by the satellite, an additional 2τ of delay is incurred on that particular packet; where τ includes the propagation and processing delay from the HUB to the satellite and the satellite to terminal. A satellite with on-board ARQ protocol can alert the sender, reducing the re-transmission delay of that particular packet by τ. If an H-ARQ scheme is employed on-board the satellite, this could potentially further reduce this delay.

On-board routers and modems that are designed using software defined radios (SDR)s have the flexibility to change their radio interface anytime. This means that a simple firmware upgrade can be done to make use of the latest transmission standard. This allows satellite networks to adapt more quickly to innovations. It should be noted that although software is prone to errors, simple firmware upgrades during testing would ensure that a radio failure due to improper programming, discovered after deployment of the satellite, could easily be rectified.

11.5 Implementation considerations and examples

Although there are many advantages listed in §§ 11.2-11.4, there are some sizable obstacles holding back the implementation of on-board processing for satellites. For example, there is higher risk involved with the deployment of new satellites with components that are not yet market tested. Although, this technology has promise, it may take time for its adoption given that service providers are content with the services that they can provide for a cheaper BPS. Although, it should be noted that the capital expenditures sunk into a satellite with on-board processing would greatly reduce the operational expenditures (cost/bit/s) of the satellite network, given that a packet-switched network can be operated in a more spectrally efficient-manner.

Another consideration is that satellites cannot be maintained in the same fashion as terrestrial systems, as they are inaccessible from space. Traditionally satellites are built with backup transponders so that if one fails, the system can maintain coverage by switching to the alternate transponder. A satellite with on-board processing would require a failsafe mechanism, which could prove more costly and act as an additional deterrent to the deployment of satellites with on-board processing.

The addition of on-board processors on the satellite could also open up new vulnerabilities within the network. As a simple bent-pipe architecture, no higher layer networking protocols were available on-board, which means that the satellites in the network remained impossible to access through security attacks. However, with on-board processing such as routers, it would be possible for a hacker to infiltrate the satellite with a variety of security attacks, such as denial-of-service attacks, for example. It is important that any implementation of IP protocols on-board satellites include detection- and encryption-based protocols to protect satellites with on-board processing.

On November 23rd, 2009, a radiation-tolerant Cisco router – the first-ever deployment of an IP-router aboard a commercial GEO satellite – was sent into space on-board the Intelsat-14 satellite [85]. The router is part of a hosted payload for the US military to demonstrate Internet routing in space (IRIS). Amongst the features and benefits listed by Cisco Systems in [86] are:

– On-board detection- and prevention-based security that allow for uplink and downlink termination.
– Dynamic IP routing.
– QoS capability.
– Software defined radio that can be upgraded from a ground site.
– Compatible with both IPv4 and IPv6.
– On-board routing between transponders to eliminate unnecessary double-hop scenarios.

On January 18, 2010 Cisco's IRIS router was tested and shown to be operating successfully [87]. The demonstrations showed that the router could route data to multiple ground stations while eliminating double-hops to teleports and that the on-board can be upgraded dynamically from the ground. At the end of three-months of testing the router will be opened up to commercial use.

12 Conclusions

With regard to the increasing importance of multi-carrier-based transmission techniques over wireless channels, and considering a number of reasons which make these techniques attractive for satellite applications, this Report describes various aspects on multi-carrier-based transmissions over satellite links. This comprehensive Report covers a wide range of topics including basic operational principles, application scenarios and the performance of multi-carrier-based transmissions analyzed through computer simulation.

13 References

[1] FCC, "Connecting America: the national broadband plan", ch. 5, March, 2010.

Available Online: http://www.broadband.gov/download-plan/

[2] Motorola, "Motorola's Guide to 3DTV", 2010.

Available Online: http://www.motorola.com/staticfiles/Business/_Documents/static%20files/Guide_To_3D_TV.pdf

[3] DVB, "DVB shifts to top gear for 3DTV", 2010.

Available Online: http://dvb.org/news_events/press_releases/press_releases/DVB_pr190-DVB-Gets-Ready-For-3DTV.pdf

[4] P. Chini, G. Giambene, and S. Kota, "A survey of mobile satellite systems", *International Journal of Satellite Communications and Networking*, vol. 28, no. 1, pp. 29-57, February 2010.

[5] ITU, "Birth of broadband – frequently asked questions", June 2007.

Available Online: http://www.itu.int/osg/spu/publications/birthofbroadband/faq.html

[6] OECD – Working Party on Communication Infrastructures and Services Policy, "Indicators of broadband coverage", January 2010.

Available Online: http://www.oecd.org/dataoecd/41/39/44381795.pdf

[7] FCC, "Wireline Competition Bureau presentation of the section 706 report and broadband data gathering order", March, 2008.

Available Online: http://www.fcc.gov/WCB_031908_open_meeting_slides.pdf

[8] Industry Canada, "Broadband Canada: Connecting Rural Canadians - FAQ", October, 2009.

Available Online: http://www.ic.gc.ca/eic/site/719.nsf/eng/h_00004.html#BPQ5

[9] FCC, "Connecting America: the national broadband plan", March, 2010.

Available Online: http://www.broadband.gov/download-plan/

[10] Sastri L. Kota, Kaveh Pahlavan, Pentti Leppanen, *Broadband satellite communications for internet access*, Kluwer Academic Publishers, Norwell, MA, October 2003.

[11] O3b Networks, http://www.o3bnetworks.com/, 2010.

[12] ITU, *Handbook on satellite communications*, 3rd ed. John Wiley & Sons Inc, New York, NY, 2002, ch. 1, p. 7.

[13] Annex 12 to Working Party 4B *Chairman's Report on the twenty sixth meeting*, Document 4B/85, May 2009.

[14] Chairman, Coordination Committee for Vocabulary (CCV), *Liaison statement to Study Group 6 (copy to Working Party 4B) – Proposed definitions for "Integrated MSS System" and "Hybrid Satellite/Terrestrial System"*, Document 4B/131, May 2010.

[15] Annex 4 to *Working Party 4B Chairman's Report on the twenty seventh meeting*, Document 4B/109, September 2009.

[16] J. Snyder and S. Patsiokas, "XM Satellite Radio-Satellite Technology Meets a Real Market", *Proceedings of the 22nd AIAA International Communications Satellite Systems Conference & Exhibit*, May 2004, Monterey, CA, USA.

[17] R. D. Briskman and S. Sharma, "DARS Satellite Constellation Performance", *Proceedings of the 20th AIAA International Communications Satellite Systems Conference & Exhibit*, Montreal, Canada, May 2002.

[18] S.-J. Lee, SangWoon Lee, K.-W. Kim and J.-S. Seo, "Personal and Mobile Satellite DMB Services in Korea", *IEEE Transactions on Broadcasting*, Vol. 53, No. 1, p. 179-187, March 2007.

[19] ITU-R, "Systems for digital satellite broadcasting to vehicular, portable and fixed receivers in the bands allocated to BSS (sound) in the frequency range 1 400-2 700 MHz", Recommendation ITU-R BO.1130-4, 2001.

[20] N. Chuberre, O. Courseille, P. Laine, L. Roullet, T. Quignon and M. Tatard, "Hybrid Satellite and Terrestrial Infrastructure for Mobile Broadcast Services Delivery: An outlook to the 'Unlimited Mobile TV' System Performance", *International Journal of Satellite Communications and Networking*, Vol. 26, No. 5, p. 405-426, September 2008.

[21] ETSI EN 302 583 V1.1.1, *Digital video broadcasting (DVB); Framing Structure, channel coding and modulation for Satellite Services to Handheld devices (SH) below 3 GHz*, 2008.

[22] ITU-R, "Technical characteristics of air interfaces for global broadband satellite systems", Recommendation ITU-R S.1709-1, 2007.

[23] R. D. Gaudenzi, A. G. i Fabregas, and A. Martinez, "Performance analysis of turbo coded APSK modulations over nonlinear satellite channels", *IEEE Transactions on Wireless Communications*, Vol. 5, p. 2396–2407, September 2006.

[24] Y.-J. Song, P.-S. Kim, D.-G. Oh, S.-I. Jeon and H.-J. Lee, "Development of mobile broadband interactive satellite access system for Ku/Ka band", *International Journal of Satellite Communications and Networking*, Vol. 24, No. 2, p. 101-117, March 2006.

[25] S. Egami, "A Power-Sharing Multiple-beam Mobile Satellite in Ka Band", *IEEE Journal on Selected Areas in Communications*, Vol. 17, No. 2, p. 145-152, February 1999.

[26] K. Lim, S. Kim, and H.-J. Lee, "Adaptive Radio Resource Allocation for a Mobile Packet Service in Multibeam Satellite Systems", *ETRI Journal*, Vol. 27, No. 1, p. 43-52, February 2005.

[27] A. Ginesi and F. Potevin, "OFDM Digital Transmission Techniques for Broadband Satellites", *Proceedings of the 24th AIAA International Communications Satellite Systems Conference (ICSSC)*, San Diego, California, June 2006.

[28] J. Krause and H. Stadali, "ETSI technical standards for satellite digital radio", *International Journal of Satellite Communications and Networking*, Vol. 26, No. 5, p. 463-474, September 2008.

[29] K. Kang, S. Kim, D. Ahn, and H. J. Lee, "Efficient PAPR reduction scheme for satellite MC-CDMA systems", *IEE Proceedings-Communications*, Vol. 152, No. 5, p. 697-703, October 2005.

[30] K. Choi, K. Kang, and Sooyoung Kim, "Peak power reduction scheme based on subcarrier scrambling for MC-CDMA systems", *IEE Proceedings-Communications*, Vol. 151, No. 1, p. 39-43, February 2004.

[31] D. A. Wiegandt, C. R. Nassar, and Z. Wu, "Overcoming peak-to-average power ratio issues in OFDM via carrier-iterferometry codes", *in IEEE 54th Vehicular Technology Conference*, Vol. 2, Atlantic City, NJ, USA, p. 660–663, October 2001.

[32] K. Anwar and H. Yamamoto, "A new design of carrier interferometry OFDM with FFT as spreading codes", *in IEEE Radio and Wireless Symposium*, San Diego, CA, USA, p. 543-546, January 2006.

[33] J. Abe, F. Yamashita, K. Nakahira, and K. Kobayashi, "Direct Spectrum Division Transmission for Highly Efficient Frequency Utilization in Satellite Communications", *IEICE Transactions on Communications*, Volume E95.B, Issue 2, pp. 563-571, February 2012.

[34] Viterbi, A. J., "Error Bounds for Convolutional Codes and an Asymptotically Optimum Decoding Algorithm" *IEEE Transcations on Information Theory*, Vol. IT-13, p. 260-269, February 1967.

[35] D. G. Forney, Jr., Concatenated Codes. Cambrige, MA, MIT Press, 1966.

[36] T. K. Moon, Error correction coding: mathematical methods and algorithms. Hoboken, NJ: John Wiley and Sons Inc., 2005, ch. 10, p. 432-433.

[37] ETSI EN 300 421 V1.1.2, *Framing structure, channel coding and modulation for 11/12 GHz satellite services*, 1997.

[38] ETSI EN 302 307 V1.2.1, *Second generation framing structure, channel coding and modulation systems for broadcasting, interactive services, news gathering and other broadband satellite applications*, 2006.

[39] S. Lin and D. J. Costello Jr., *Error control coding: fundamentals and applications*, 2nd ed.Upper Sadle River, NJ: Pearson Prentice Hall, 2004, ch. 15, p. 743-748.

[40] C. Berrou and A. Glavieux, "Near optimum error correcting coding and decoding: turbo-codes", *IEEE Transactions on Communications*, Vol. 44, No. 10, p. 1261-1271, October 1996.

[41] T. K. Moon, *Error correction coding: mathematical methods and algorithms*. Hoboken, NJ: John Wiley and Sons Inc., 2005, ch. 14, p. 585.

[42] P. Sweeney, *Error control coding: from theory to practice*, John Wiley & Sons, Ltd., 2002.

[43] ETSI EN 301 390 V1.5.1, *Digital video broadcasting (DVB); Interaction channel for satellite distribution systems*, 2009.

[44] R. Pyndiah, A. Picart, and S. Jacq, "Near optimum decoding of product codes", *IEEE GLOBECOM'94*, Vol. 1, San Francisco, CA, p. 339–343, November 1994.

[45] R. Pyndiah, "Near-optimum decoding of product codes: block turbo codes", *IEEE Transactions on Communications*, Vol. 46, No. 8, p. 1003–1010, August 1998.

[46] S. Kim, W. S. Yang, and H.-J. Lee, "Trellis-based decoding of high-dimensional block turbo codes", *ETRI Journal*, Vol. 25, No. 1, p. 1-8, February 2003.

[47] G. C. Clark, Jr. and J. B. Cain, "Error-Correction Coding for Digital Communications," *Plenum Press*, New York, 1981.

[48] D. Chase, "A class of algorithms for decoding block codes with channel measurement information", *IEEE Transactions on Information Theory*, Vol. IT-18, p. 170–182, January 1972.

[49] J. Hagenauer and P. Höher, "A Viterbi algorithm with soft decision outputs and its application", *Proceedings of the IEEE Global Communications Conference*, GLOBECOM 1989, p. 47.1.1-47.1.7, November 1989.

[50] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate", *IEEE Transactions on Information Theory*, Vol. IT-20, p. 284-287, March 1974.

[51] L. Papke and P. Robertson: "Improved decoding with the SOVA in a parallel concatenated (turbo-code) scheme", *International Conference on Communications (ICC'96)*, Dallas, USA, p. 102-106, June 1996.

[52] L. Lin and R. S. Cheng: "Improvements in SOVA-based decoding for turbo codes", *International Conference on Communications (ICC'97)*, Montreal, Canada, p. 1473-1478, June 1997.

[53] M. C. Fossorier, F. Burket, S. Lin and J. Hagenauer: "On the equivalence between SOVA and Max-Log-MAP decodings", *IEEE Communication Letters*, Vol. 2, No. 5, p. 137-139, May 1998.

[54] A. Picart and R. M. Pyndiah, "Adapted iterative decoding of product codes", GLOBECOM'99, Rio de Janeiro, Brazil, p. 2357-2362, December 1999.

[55] R. E. Gallager, "Low density parity-check codes", *MIT Press*, 1963.

[56] R. M. Tanner, "A recursive approach to low complexity codes", *IEEE Transactions on Information Theory*, Vol. IT-27, p. 553–547, September 1981.

[57] D. J. C. MacKay and R. Neal, "Good codes based on very sparse matrices", *5th IMA Conference on Cryptography and Coding*, no. 1025. C. Boyd, Ed. Berlin, Germany: Springer Lecture Notes in Computer Science, December 1995, p. 100–111.

[58] S. Sesia, I. Toufik and M. Baker, *LTE – The UMTS long term evolution*. Chichester, West Sussex, United Kingdom: John Wiley and Sons Inc., 2009, ch. 9, p. 226–227.

[59] 3GPP2, 3GPP2: *Physical Layer for Ultra Mobile Broadband* (UMB) *Air Interface Specification*, 2007.

[60] M. Dorrance and I. Marland, "Adaptive Discrete-Rate MIMO Communications with Rate-Compatible LDPC Codes", Submitted to *IEEE Transactions on Communications*, September 2009.

[61] R. J. McEliece, D. J. C. Mackay, and J.-F. Cheng, "Turbo decoding as an instance of Pearl's 'Belief Propagation' algorithm", *IEEE Journal on Selected Areas in Communications*, Vol. 16, p. 140-152, February 1998.

[62] J. Pearl, *Probabilistic reasoning in intelligent systems*. San Mateo, CA, USA: Morgan Kaufmann, 1988, ch. 4.

[63] J. R. Barry, "Low-Density Parity Check Codes", Georgia Institute of Technology, October 2001. Available Online: http://eref.uqu.edu.sa/files/Others/Ldpc/LDPC%20intro%20and%20tutorial/Notes%20on%20LDPC%20codes.pdf

[64] S. Lin and P. S. Yu, "A hybrid-ARQ scheme with parity retransmission for error control of satellite channels", *IEEE Transactions on Communications*, vol. COM-30, p. 1701–1719, July 1982.

[65] Shu Lin, D. Costello, and M. Miller, "Automatic-repeat-request error-control schemes", *IEEE Communications Magazine*, Vol. 22 , No. 12 , p. 5-17, December 1984.

[66] S. Kallel, "Complementary punctured convolutional codes and their applications", *IEEE Transactions on Communications*, Vol. 43, p. 2005-2009, June 1995.

[67] S. Kim, S. Ryoo, and D. S, Ahn, "Evaluation of rate compatible block turbo codes for multimedia application in satellite communication network", *International Journal of Satellite communications and Networking*, Vol. 24, No. 5, p. 419-435, September 2006.

[68] D. J. Costello, Jr., J. Hagenauer, H. Imai, and S. B. Wicker, "Applications of error-control coding," *IEEE Transactions on Information Theory*, Vol. 44, No. 6, p. 2531–2560, October 1998.

[69] D. Chase, "Code combining-A maximum-likelihood decoding approach for combining an arbitrary number of noisy packets", *IEEE Transactions on Communications*, Vol. COM-33, p. 385–393, May 1985.

[70] D. M. Mandelbaum, "*An adaptive-feedback coding scheme using incremental redundancy*", *IEEE Transactions on Information Theory*, Vol. IT-20, p. 388–389, May 1974.

[71] DVB, "DVB Fact Sheet – 2nd generation satellite: The most advanced satellite broadcasting system in the world", July 2009. Available Online: http://www.dvb.org/technology/fact_sheets/DVB-S2-Fact-Sheet.0709.pdf

[72] ETSI TR 101 198 V1.1.1, *Implementation of Binary Phase Shift Keying (BPSK) modulation in DVB satellite transmission systems*, 1997.

[73] ITU-R, "Digital satellite broadcasting system with flexible configuration (television, sound and data)", Recommendation ITU-R BO.1784 1, 2006.

[74] S. Lin and D. J. Costello Jr., *Error control coding: fundamentals and applications*, 2nd ed. Upper Sadle River, NJ: Pearson Prentice Hall, 2004, ch. 1, p. 18–21.

[75] DVB, "DVB Fact Sheet – Return Channel Satellite: The open standard for two-way broadband VSAT systems", July 2009. Available Online: http://www.dvb.org/technology/fact_sheets/DVB-RCS-Fact-Sheet.0509.pdf

[76] DVB, "DVB Fact Sheet – Satellite services to handhelds: Mobile TV over advanced hybrid satellite/terrestrial networks", July 2009. Available Online: http://www.dvb.org/technology/fact_sheets/DVB-RCS-Fact-Sheet.0509.pdf

[77] ETSI EN 302 304 V1.1.1, *Digital video broadcasting (DVB)*; *Transmission system for hanheld terminals*, 2004.

[78] A. A. M. Saleh, "Frequency-independent and frequency-dependent nonlinear models of TWT amplifiers", *IEEE Transactions on Communications*, Vol. COM-29, p. 1715-1720, November 1981.

[79] A. N. D'Andrea, V. Lottici, and R. Reggiannini, "RF power amplifier linearization through amplitude and phase predistortion", *IEEE Transactions on Communications*, Vol. 44, p. 1477-1484, November 1996.

[80] S. C. Thomson, "Constant envelope OFDM phase modulation", Master's thesis, University of California, 2005.

[81] F. P. Fontán, M. Vázquez-Gastro, C. E. Cabado, J. P. García, and E. Kubista, "Statistical modeling of the LMS channel", *IEEE Transactions on Vehicular Technology*, Vol. 50, No. 6, p. 1549-1567, November 2001.

[82] ITU, *Handbook on satellite communications*, 3rd ed. John Wiley & Sons Inc, New York, NY, 2002, ch. 2, pp. 75.

[83] G. Maral and M. Bousquet, *Satellite Communications Systems: Systems, Techniques and Technology*, John Wiley & Sons Ltd, Chichester, England, 2003.

[84] Cisco Systems, "Demo: Internet Routing in Space Capabilities", November, 2009. [Video Available Online: http://www.cisco.com/web/strategy/government/space-routing.html]

[85] Cisco Systems, "Cisco router sent into space aboard Intelsat satellite", 2009. Available Online: http://www.cisco.com/web/strategy/docs/gov/Cisco_Intelsat_ IRIS_release.pdf

[86] Cisco Systems, "Cisco space router: Increase utilization, reduce latency", November, 2009. Available Online: http://www.cisco.com/web/strategy/docs/gov/Cisco_Space_Router.pdf

[87] Cisco Systems, "Cisco router successfully operates in orbit", 2010. Available Online: http://newsroom.cisco.com/dlls/2010/prod_011910b.html