Technical Advances in Digital Audio Radio Broadcasting

CHRISTOF FALLER, BIING-HWANG JUANG, FELLOW, IEEE, PETER KROON, FELLOW, IEEE, HUI-LING LOU, MEMBER, IEEE, SEAN A. RAMPRASHAD, MEMBER, IEEE, AND CARL-ERIK W. SUNDBERG, FELLOW, IEEE

Invited Paper

The move to digital is a natural progression taking place in all aspects of broadcast media applications from document processing in newspapers to video processing in television distribution. This is no less true for audio broadcasting which has taken a unique development path in the United States. This path has been heavily influenced by a combination of regulatory and migratory requirements specific to the U.S. market. In addition, competition between proposed terrestrial and satellite systems combined with increasing consumer expectations have set ambitious, and often changing, requirements for the systems. The result has been a unique set of evolving requirements on source coding, channel coding, and modulation technologies to make these systems a reality.

This paper outlines the technical development of the terrestrial wireless and satellite audio broadcasting systems in the U.S., providing details on specific source and channel coding designs and adding perspective on why specific designs were selected in the final systems. These systems are also compared to other systems such as Eureka-147, DRM, and Worldspace, developed under different requirements.

Keywords**—***Audio coding, channel coding, digital sound broadcasting.*

I. INTRODUCTION

A century ago, Marconi pioneered transmission of information across the Atlantic Ocean using electromagnetic (EM) wave radiation instead of electric current over conducting wires. Marconi's transmission took the form of Morse code, which is obviously a discrete expression of information, and thus could be considered the first digital wireless electronic communication in existence. The use of EM waves for broadcasting, however, did not come about until two decades later. In 1919, Frank Conrad founded a broadcasting venture in a small red brick garage behind his home in Pittsburgh, PA, spawning the term "radio," as he used EM radiation as Marconi did. Public broadcast of radio was finally realized in 1922, leading to a new era of mass communication based on electronic medium. Since then, broadcast radio has been an important source of information, powerful politically both in peace time and in wartime and informative and influential culturally both at work and in the household. Today, the average household in the United States has 5.6 radio receivers, totaling 580 million units in use nationwide. Every week, radio programs reach 96% of people over 12 years old who on the average listen over 3.2 h daily. These programs are being transmitted from over 11 700 radio stations in the U.S. alone.

The 20th century has been a century of communications with the advent of telephone, radio, and television technologies at the juncture of the 19th and 20th centuries to facilitate information sharing between people hundreds of miles apart or across the continents. For over 60 years, however, the transmission technology was mostly based on analog techniques, such as *amplitude modulation* (AM), *frequency modulation* (FM), *phase modulation* (PM), or their derivatives. Even in wired telephony, AM was used to achieve multiplexing in military carrier systems as early as World War I. The majority of public radios today operate in three modes, AM, FM, and stereo FM (some with stereo AM) over a spectrum suitable for terrestrial propagation, including via ionospheric or sky waves.

Broadcast transmission via EM waves is subject to degradation which defines a station's coverage area. A coverage or service area is defined by two contours: the *interference-limited* contour and the *noise-limited* contour. The noise-limited contour is largely defined by the transmission power of the

Manuscript received December 17, 2001; revised April 30, 2002.

C. Faller, P. Kroon, and S. A. Ramprashad were with Bell Laboratories, Murray Hill, NJ 07974 USA. They are now with Agere Systems, Berkeley Heights, NJ 07922 USA.

B.-H. Juang was with Bell Laboratories, Murray Hill, NJ 07974 USA. He is now with Avaya Labs Research, Basking Ridge, NJ 07920 USA.

H.-L. Lou was with Bell Laboratories, Murray Hill, NJ 07974 USA. He is now with Marvell Semiconductors, Sunnyvale, CA 94089 USA.

C.-E. W. Sundberg was with Bell Laboratories, Murray Hill, NJ 07974 USA. He is now with Ibiquity Digital Corporation, Warren, NJ 07059 USA.

Publisher Item Identifier 10.1109/JPROC.2002.800718.

station, and reception of the signal becomes negligible beyond this contour. The interference-limited contour is largely defined by the interference from colocated stations, i.e., stations having the same carrier frequency but geographically separated by a certain minimum distance. Outside this contour, the level of the interference signal supersedes the broadcast signal, resulting in either no or very poor reception of the original station. A radio station requires a license from the Federal Communications Commission (FCC) [1] of the U.S. Government to operate its broadcasting. Regulatory procedures set up by the FCC ensure proper selection of the antenna location and power management to define a service area map for each broadcasting licensee [1], [2].

Within the contour defined coverage area, there are still several causes of signal degradation, such as fading and shadowing. Fading is due to multiple reflections of the signal from the terrain (e.g., hills and mountains) or large buildings. Shadowing refers to blockage of the signal by terrain or buildings. These various causes of degradation result in poor sound quality, which is difficult to mitigate with typical analog transmission schemes.

Note that, in AM and FM systems, the transmitter power can be increased to improve the SNR of the received signal. Thus, the noise-limited contour is increased. However, the interference-limited contour is decreased at the same time. Also, in FM systems, the frequency deviation ratio can be increased to improve the fidelity of the received signal. However, this uses bandwidth, which in turn makes less spectrum available for other users or for digital transmission, ultimately decreasing the interference-limited contour.

The problems with analog transmission are quite well understood among communications engineers. During and after World War II, substantial research effort was spent on developing the basis of digital communication technologies, ranging from Shannon's information theory and pulse-coded modulation (PCM) to the theory of digital filtering and signal processing. A comprehensive comparison of the pros and cons between analog and digital transmission can be found in [3]. In essence, digital communication allows incorporation of safeguarding measures (i.e., channel coding) to insure the fidelity of the received digital representation of the source signal (i.e., the result of source coding) and regeneration of the signal without accumulative degradation. Coupled with the progress in digital computing and microprocessor technologies, digital communication has been a part of the digital revolution since the 1960s. As a matter of fact, the telephone network backbone, which forms the so-called trunk lines, have become virtually all digital since the 1980s in the U.S. and possibly worldwide. Today, most media signals are also represented, stored, or transmitted in digital forms (for example, the compact disc (CD) for music, high-definition television (HDTV), etc.). These formats also provide improved quality, in terms of audio bandwidth and picture resolution, over the analog formats. The traditional terrestrial radio is the last communication and broadcasting service to become digital, at least in the North America region. The drive to all-digital radio broadcasting thus gained momentum in the late 1980s as CD music became ubiquitous and audio compression techniques demonstrated ever-increasing efficiency due to the introduction of perceptual audio coding [4], [5].

In the early 1990s, progress toward the digital broadcast of audio programs took place along several directions. In Europe, the European Union (EU) attempted to unify broadcasting across the national boundaries by supporting a development effort called Eureka-147 [64], [6]. This plan realized a new business model similar to that of the cable TV industry in the U.S. In the new model, a station is responsible for the transmission of program ensembles, each of which consists of six channels, over an authorized frequency spectrum. A channel refers to a programming entity and is likely to be associated with a media production company. The notion of a station and that of a channel are thus separated, unlike the traditional model in which the two are synonymous. (For example, a station presently may have a designation such as WOR710-AM, where WOR is the station name and the numeric 710 refers to the carrier frequency in AM mode.) A standard carrying the name Eureka-147 was adopted by the European Community in 1995 [64]. A number of countries have since announced plans to test and adopt the system for future digital audio broadcasting. At the time when this paper is written, live broadcasting using Eureka-147 is taking place in several countries on a daily basis.

The Eureka-147 system was designed to operate in several frequency bands, most commonly in the *L*-band (1500 MHz). Sometimes it is also referred to as "new-band" radio. These spectral bands were allocated by the EU and approved in 1992 by the World Administrative Radio Conference (WARC) for the new digital audio radio service. However, these spectral bands are not immediately available in the U.S. due to prior usage authorization. Adoption of Eureka-147 in the U.S., although strongly supported by the Consumer Electronics Manufacturers Association (CEMA), was met with difficulty in spectrum allocation. Another hurdle to the adoption of the European system is that it entails a new broadcast licensing campaign, which can unpredictably change the landscape of the entire broadcast industry in the U.S. The National Association of Broadcasters (NAB) in the U.S. thus favored a technology called *in-band, on-channel* (IBOC). This technology allows a station to smoothly migrate into digital broadcasting without having to seek a new operating license from the FCC or abruptly discontinuing its analog transmission. This is the case for both AM (510–1710 kHz) and FM (88–108 MHz) bands. Since 1994, NAB has worked with several key technology teams to promote IBOC (see Section IV for details). The terrestrial U.S. digital audio radio systems will first be introduced as hybrid IBOC systems where digital transmission is added to existing analog FM and analog AM. These systems will then evolve to all-digital IBOC systems where the analog signals are replaced by additional digital transmission. A draft recommendation for a world standard for digital audio radio below 30 MHz has been recognized by the International Telecommunication Union (ITU) [7]. Part of this standard is being developed by Digital Radio Mondiale (DRM) [8], [9]. For the medium-wave AM band, the U.S. hybrid IBOC and all-digital IBOC are also part of this world standard.

Another push to digital audio broadcast in North America came from proponents of direct satellite transmission. Direct satellite broadcast (DSB) for television has been in service since early 1990s. It has, however, not been extended to audio services, which, according to market studies, are quite attractive in mobile applications. Drivers of automobiles and trucks had expressed desire to subscribe to high quality audio broadcast throughout the North America region. Proponents of the plan convinced the FCC to release two bands of spectrum, 12.5 MHz each, around 2.3 GHz (*S*-band) for such a satellite-based digital audio broadcast service. Subsequently, the allotted spectra were auctioned in 1997 and two spectrum licensees (Sirius [10] and XM [11]) thus set out to develop the systems, with target broadcast launch date sometime in the later part of the year 2001. This is often referred to as *satellite digital audio radio services* (SDARS).

Service in SDARS is subscription based; a subscriber pays a monthly fee to receive the digitally protected broadcast signal. With the allotted spectrum, each broadcast company is able to provide about 100 channels of audio programs, some mostly music while others mostly voice-oriented talk shows. The two broadcasters, however, employ different satellite technologies; one uses a geosynchronous system and the other uses a geostationary system. These two systems require different signal relay plans (the so-called gap-fillers) in order to provide proper coverage for areas that may be blocked by terrain or buildings. A distinct feature of SDARS compared to terrestrial systems is that a listener can stay with a particular program throughout the entire North America region without having to switch channels due to the nature of the satellite coverage.

Global radio [12] is a potential future provider of satellite digital radio in Europe. It is set for an early 2005 launch and aims at providing 200 channels of audio. Three satellites in a 24-h highly elliptic orbit will be used. One main beam and seven spot beams over Europe are planned. Thus, local programming in separate languages is possible.

Finally, it is worth mentioning that the ubiquity of Internet and multimedia capabilities of personal computers (PCs), both in software and hardware, have given rise to an entirely new paradigm in radio broadcast, i.e., the so-called "webcast" or "Internet radio." Using media streaming technologies, instead of an EM wave receiver, a PC can download a "radio" or TV program from a server, i.e., the "webcaster," and allow the user to listen and watch without being limited by typical wireless constraints, e.g., contour limits and spectrum availability. Proper streaming technologies coupled with efficient audio coding techniques, plus the virtually unlimited reach of the Internet, make webcast a new favorite of many listeners who "tune" to stations that are continental apart and are otherwise unreachable via traditional radio waves. According to BRS Media, over 3000 radio stations worldwide are webcasting as of April 2000, among which nearly 300 are broadcasting over the Internet only [13]. The statistics include 58 radio networks. Several stand-alone receivers, the so-called "Internet radios," which have IP access without requiring a PC, are being offered on the market. With ever increasing Internet access and active user population, webcasting and Internet radio undoubtedly are reshaping the traditional radio industry.

In short, digital audio radio services, whether it is over terrestrial transmission, relayed by satellite, or in the form of media streaming via Internet, is taking place at this turn of century, after nearly eighty years of operation in analog modes [14]. The advance is primarily due to the progress in digital audio coding and several key innovations in transmission technologies. The purpose of this paper is to present, according to our involvement and insights, the technology and system components that are behind this important evolution. Our presentation will focus on the terrestrial and the satellite systems as they represent the most profound depth of complexity and technical challenges. Also, we believe that the advance in audio coding played a critical role in making digital audio broadcasting possible given the current spectral and regulatory constraints. Hence, a large section of this paper is a discussion on recent progress in audio coding. Specific coding schemes designed for various broadcast services will be covered when details on individual systems are presented.

II. Audio Coding Algorithms

As mentioned in the Introduction, there are many advantages of using digital transmission and digital representations of audio including an increased robustness to channel conditions and the ability to regenerate signals without accumulative degradation.

The digital nature of the links also increases the flexibility of the underlying audio format of the signals being broadcast. For example, the audio signals transmitted can have different sampling rates and different multichannel formats and can even include differentiated levels of quality targeting different receivers. Digital representation also allows the systems to transmit data, e.g., stock quotes and messages, to use encryption algorithms and to manage access to subscription-based services at receivers.

A. General Requirements for the Source Coders

The advantages of digital systems just mentioned do have corresponding requirements on the source coding algorithms used to encode the digital audio source signals. These considerations include:

- the compression rate of the raw information;
- the format represented in the compressed bitstream;
- the algorithm's robustness to channel errors;
- the audio quality, possibly as a function of the signal type (e.g., music or speech) and/or station;
- the delay introduced by source coding;
- the complexity of the source encoder and decoder.

Many of these considerations are specific to the broadcast environment and differ greatly from those of speech communication and storage/retrieval (e.g., MP3 player or audio CD) type applications.

The first requirement (the compression of the raw digital information) is probably the most obvious of the considerations. Within the design constraints, the systems presented have effective average (source) transmission rates of no greater than 64–128 kb/s for each audio program. In fact, some systems, such as the AM systems discussed in Section VI, may have even lower average rates, e.g., 24–32 kb/s. Other systems, such as the FM system, may have a variety of limits operating simultaneously depending on a receiver's location in the coverage area. In fact, satellite systems can have a single aggregate bound sharing a *cluster* of programs with each program potentially having a different target bit rate, e.g., 20–96 kb/s.

In contrast, the raw digital information rate of a single stereophonic source at 16 b/sample and a sampling rate of 44.1 k-samples/s per channel is a fixed bit rate of 1.411 Mb/s. To meet the bit-rate requirement range of 20–128 kb/s, compression ratios of up to 70 are required. At bit rates below 64 kb/s, compression will inevitably involve compromises on the audio quality, acoustic bandwidth, and the source format, e.g., monophonic rather than stereophonic formats.

Related in part to compression is the second requirement of having formats within the compressed source bitstream. A rudimentary example is the simple separation of stereophonic and monophonic information, as happens in current analog FM systems. Digital systems, however, have the potential to take advantage of more elaborate multistream, multidescription, and layered-description schemes. Such schemes can be matched to the specific properties of the channels in a given broadcast application. For example, the IBOC systems described in Sections V and VI have possibly unequal and dynamic error characteristics on either side of the transmission band. This can create diversity in the channel conditions seen by different areas of the source bitstream. Similarly, satellite systems have diversity provided by the use of multiple satellites. The source coding algorithm, as well as the channel coding algorithm, can be used to produce bitstream formats well suited to these applications.

The use of diversity does improve the performance of the systems. Despite this, bit errors do occur during transmission and create the third set of error robustness requirements that must be taken into account by the source coder designs. The systems presented use an error-detecting channel code to detect bit errors in blocks (frames) of the received bitstreams. The designs target low undetected (residual) bit error rates (BERs) on the order of $\leq 10^{-5}$. This is in contrast to the higher rates tolerated by cellular systems which are on the order of $\sim 10^{-3}$. When an error is detected by the error-detecting code, the corresponding frame of bits is discarded entirely, creating another error situation termed a *frame erasure.* The probability of such an erasure, the frame-erasure rate, can be several percent in these systems. The source decoders used in these systems therefore include frame erasure mitigation strategies that fill in the missing audio information when a frame erasure occurs (see Section III-D). The source decoders deployed are designed to be robust to the target residual BERs and frame-erasure rates of these systems.

The fourth set of requirements focuses on the decoded audio quality. Under this set of considerations, the robustness to source material is probably the most challenging aspect of the source coder designs. This is due mainly to two factors: 1) the wide variation in audio content in broadcast applications and 2) the fact that the ultimate judge of quality is the human listener. The material of most audio programs includes a broad range of speech, music genre, recording conditions, and mixes of different source inputs. Acoustic bandwidths and noise conditions can vary within a program, e.g., in situations where listeners or reporters call in via the telephone network, and gross characteristics of the music material can vary from station to station, e.g., a popular music station versus a classical music station. In addition, the material found in many broadcast environments can be heavily preprocessed, e.g., gain equalized or compressed/decompressed with another source coder. The variability in the audio material, and the low source bit rates used in many of the digital audio systems, present a significant challenge to the source coding technology. This will be discussed further in Section II-B.

Finally, all these requirements have to be achieved under the practical constraints of the application, system, and hardware. There are two main such considerations which impact source coders: complexity and delay. The first constraint is mainly due to limitations placed on the receiver hardware. These limitations are defined in terms of measures such as the rate of algorithmic operations, the memory (RAM and ROM) requirements, and the size and number of chips. These quantities directly impact consumer-sensitive concerns such as the size, price, and power consumption (battery life) of receivers. The exact numbers in terms of MIPS, ROM, and RAM depend on the coder used, its sampling rate, and the target hardware, but are well within reach of low-cost consumer applications. Encoder complexity also has to be managed but is less of a concern since encoders are only deployed at the (limited number of) broadcasting installations. The increased costs and complexity associated with more advanced (better performing) source-encoding algorithms can therefore often be absorbed by the broadcasting companies.

The second major system constraint of that of *latency* or *delay*. In simple terms, this is defined as the interval in time between which the encoder system first receives input audio samples and the time the receiver produces the corresponding output audio. Although the delay constraints do not need to be as stringent as in two-way communications, they are bounded for several reasons. The first reason is a cost issue since some of the sources of delay translate proportionally into the need for memory (RAM/ROM) and processor resources at the receiver. A second reason focuses on impacts of delays at the receiver which affect the user interface. Here, the main concern is referred to as the *tuning problem*. This problem can occur if receiver-specific components of the end-to-end delay adversely affect the time required to tune into a station, i.e., the time from when the user selects or changes a station to the time the required audio output is played. Users need to be able to scan the material of different stations in a timely and convenient fashion.

The sources of delay introduced by the source coder are mainly due to the block-wise (or frame-wise) processing of input samples and the use of *signal lookahead*. Block-wise processing allows coders to take advantages of correlations in the signal, thereby improving coding efficiency [15]–[17]. Signal lookahead involves the use of future signal information to influence and improve the encoding of the present block of signal. Signal lookahead is also used in lapped-transforms such as the *modified discrete cosine transform* (MDCT) [18], [19]. The combined effect of the coder's frame size and signal lookahead is often termed the *algorithmic delay* [20].

Another less obvious way source coders introduce delay is through the use of variable-bit-rate schemes, i.e., using varying numbers of bits for different encoded frames. In systems with a fixed source transmission bit rate, buffers for bits awaiting transmission and bits awaiting decoding are used to absorb the bit-rate variability, thereby minimizing underflow and overflow problems in the system.

Finally, it is worth mentioning that the channel coding algorithm can also be a source of delay. One main factor is due to bit-interleaving. This interleaving delay is the period over which transmitted bits are randomly ordered before transmission to reduce correlations between bit errors. If block channel encoders are used such as in cyclic redundancy check (CRC) or Reed–Solomon (RS) codes, they can, depending on their block-length, also introduce delay. This again creates tradeoffs in the bit-error characteristics and, therefore, the resulting requirements and tradeoffs in source coder designs.

From the above discussion, it should be clear that a source coding design represents a tradeoff in multiple source related factors and potentially a tradeoff with a channel coding algorithm. For example, one can minimize source coder delay but this may come at the expense of compression efficiency for a fixed quality level. Similarly, minimizing the delay introduced by channel coding can be at the expense of error correction performance and the bit-error characteristics seen by source coders. The choice of source coder is therefore a compromise between the issues specific to the application, system, and even audio program.

B. Source Coding Paradigms and Quality Tradeoffs

One of the major factors enabling the deployment of digital audio broadcast systems is the advances in audio compression technology. With these advances, source bit rates for *transparent* stereophonic CD quality (perceptually indistinguishable from uncompressed CD quality) are now below the 128-kb/s bound required by these broadcast systems. There is even evidence for transparent quality at rates as low as 96 kb/s and "CD-like" quality, the quality at which most untrained listeners cannot tell the difference between compressed and uncompressed quality, at rates as low as 64 kb/s, depending on the music material.

The audio broadcast application differs from many other applications in that there are stringent requirements and expectations on both speech quality and music quality. This importance not only reflects requirements from the broadcasters and listeners themselves, but also the expectations of artists, talk radio hosts, and advertisers who create the broadcast content.

The joint speech and audio requirement is not of great concern at higher bit rates (above 96 kb/s), where transparent stereophonic CD quality is possible on most signal types. However, as mentioned in the introduction to this section, at lower bit rates, compromises in quality have to be made. Complicating the matter further is the fact that at low bit rates there are significant differences in the attributes of the compression technologies available. Selecting an appropriate coding technology often involves a balance between performance attributes as a function of signal type as well as on the hardware requirements different technologies impose. There is no ideal technology satisfying all concerns at low bit rates.

To understand why this is so, and the choices that are reflected in later sections, it is good to briefly review the attributes of the two main categories of coding technologies that are available to various digital audio broadcast applications, i.e., speech and audio coding technologies. To begin, speech coding technologies in general use model-based or waveform-based techniques that take advantage of the redundancies in the production and perception mechanism of speech [20]. Sources considered are generally single-channel signals with a primary sample rate of 8 k-samples/s [20]. More recently, the compression of wide-band speech sampled at 16 k-samples/s has also seen significant advances. [20, ch. 8], [21]. Interest in stereophonic speech is also emerging [22]–[24], but for different reasons and with different technical challenges than those of broadcasts in music recording industries [23], [25]. Speech coders are also designed under constraints of low algorithmic delay (e.g., less than 50 ms) and with bounds on both encoder and decoder complexity, making them useful for two-way communication applications. By default, they are also useful for broadcast applications.

Speech coding designs often achieve high speech quality at rates less than 1 b per input sample, which is quite notable especially at lower sample rates of 8 and 16 k-samples/s, where up to 21 of the 25 critical bands of hearing are covered in the acoustic bandwidth [26]–[29]. However, for robust performance across all signal classes, in particular music, bit rates closer to 2 b per input sample are generally required. State-of-the-art technology for high-quality speech coding at bit rates of $\leq$16 kb/s for narrow-band sources and $\leq$32 kb/s for wide-band sources have been standardized by bodies such as the ITU-T and ETSI [21], [30]–[32].

Audio coders in contrast rely less on speech-specific attributes and more on the statistical redundancy common in many audio signals and general principles on the human auditory perception [5], [26], [29], [33]–[38]. Common audio coder designs include transform or filter-bank signal decompositions combined with perception models and/or lossless coding techniques such as Huffman coding [33], [27], [28], [38] (see Section II-C).

At higher sample rates, from 32 to 48 k-samples/s, CD-like quality and transparent quality can be achieved with many popular coding technologies between 1.0 and 1.5 b per input sample per channel [39]. The MPEG-2 AAC [35], [36] and perceptual audio coding (PAC) [38], [40] coders claim to have transparent CD quality below 128 kb/s and nearly CD-like quality at 64 kb/s for stereophonic signals.

At lower sample rates and acoustic bandwidths, e.g., 8 and 16 k-samples/s and acoustic bandwidths of less than 4 or 8 kHz, robust performance of audio coding designs usually requires bit rates closer to 2 b/input sample, similar to the results obtained with speech coding technology. The increase

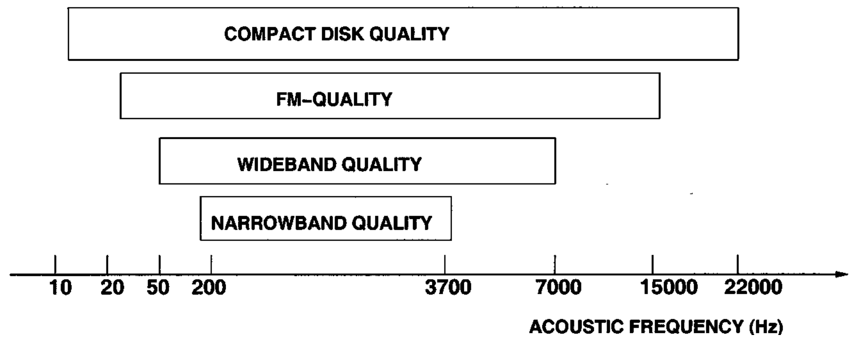

Fig. 1. Acoustic bandwidth and quality.

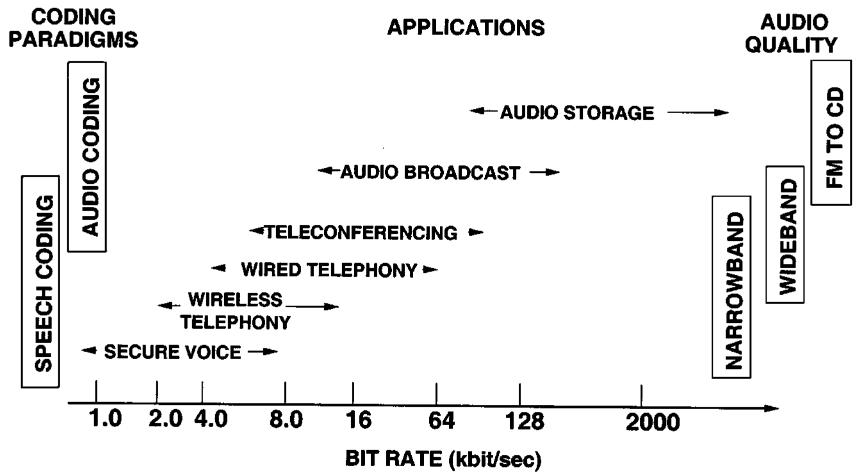

Fig. 2. Application space versus bit rate, paradigms, and acoustic bandwidth.

in the required number of bits per sample per channel for both technologies at the lower sample rates is due to a number of factors, including: 1) the general statistical structure of many audio signals which have higher energy in the lower frequencies; 2) the move to monophonic signals; and 3) the human hearing mechanism which has a greater frequency selectivity at lower frequencies [26]–[29].

Even at 2 b/sample, it is important to stress that the difference between speech and audio technologies are still apparent. The classic (and expected) tradeoffs in performance is speech coders outperforming audio coders on speech and audio coders outperforming speech coders on music and general audio. Therefore, even in the situation where coders are considered to be robust to source material, the choice of coder can be heavily influenced by the program material of the broadcast application.

To summarize, a general picture of the audio quality as a function of acoustic bandwidth is shown in Fig. 1. A general summary of the application areas, bit rates, coding paradigms, and the nominal audio quality is shown in Fig. 2. The overlap in speech and audio coding technologies is clearly visible. More will be said on the potential quality, bit rates, and technology tradeoffs in Sections III-A and III-G.

Matching the tradeoffs of the different paradigms to the source material, transmission channels, and hardware requirements is the challenge faced by source coding technology in digital audio broadcasting systems. Some of these challenges have resulted in new advances in the area of speech and audio coding, including ideas on statistical multiplexing of multiple programs in a perceptually meaningful way [41], using diversity in the source stream in both embedded and multidescriptive fashions, improving quality of audio coders on speech signals and using multiple paradigms within a single coding structure. These will be discussed in Sections II-C and III. Section II-C discusses the primary coding technology used in satellite and terrestrial systems.

Fig. 3. Generic audio encoding/decoding diagram.

C. Perceptual Audio Coding

Fig. 3 shows the general scheme for audio coding. General source coding algorithms maximize objective measures such as the SNR for a given bit rate. Perceptual audio coders explore factors of human perception with the aim of minimizing the perceived distortion for a given bit rate. Compression in a perceptual audio coder involves two processes: *redundancy reduction* and *irrelevancy reduction*. The filter bank of a perceptual audio coder yields a high degree of *redundancy reduction* due to the statistical nature of audio sources, e.g., the energy of many audio sources is often concentrated in a few subbands of the entire signal bandwidth. The efficiency of the coder is further improved without impairing the audio quality by shaping the quantization noise

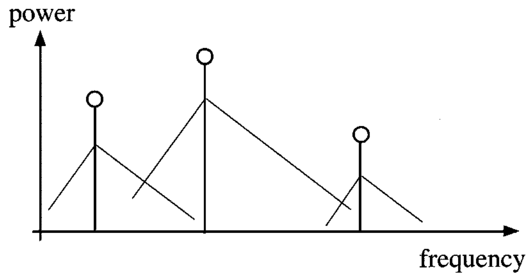

Fig. 4. The masked threshold is computed by considering the masking effect of each spectral component of the audio signal.

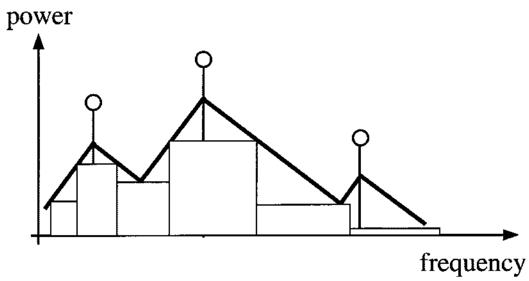

Fig. 5. The spectral coefficients are divided into coding bands. Each coding band is quantized such that the error is just below the masked threshold.

according to perceptual considerations. This is the basis for irrelevancy reduction.

One way irrelevancy reduction is achieved is by taking masking effects of the human auditory system into account. Masking describes the phenomenon in which one signal (in this case, quantization noise) becomes inaudible in the presence of another signal (in this case, the coded version of the input signal). Such masking happens in both the time and frequency domains. In the frequency domain, the level below which the masked signal becomes inaudible is termed the *masked threshold*. This threshold is a function of the masking signal and is often computed by considering the masking effect of each component of the audio signal [42], [4]. Fig. 4 shows how, for each component of the audio signal, the masking spreading function is considered separately for obtaining the net masked threshold for the audio signal. During the encoding process, the spectral coefficients of the filter bank of a perceptual audio coder are grouped into coding bands. Each of these coding bands is quantized separately such that the resulting quantization error is just below the masked threshold, as shown in Fig. 5.

The structure of a generic perceptual audio encoder [43], [44] is shown in Fig. 6. The four main functions are as follows.

- The input samples are converted into a subsampled spectral representation using a filter bank [18].
- A perceptual model estimates the signal's masked threshold [42]. For each spectral coefficient, this gives the maximum coding error that can be allowed in the audio signal while still maintaining perceptually transparent signal quality.
- The spectral values are quantized such that the error will be just below the masked threshold. Thus, the quantization noise is hidden by the respective transmitted signal. The resulting quantizer indices are coded with a lossless coder.
- The coded spectral values and additional side information are packed into a bitstream and transmitted to the decoder or stored for future decoding.

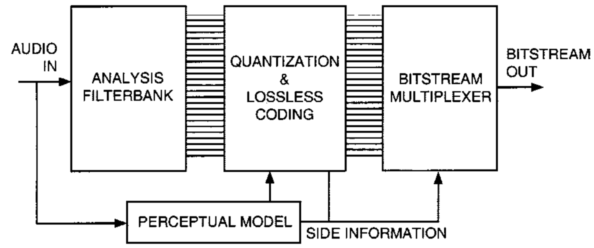

Fig. 6. Generic perceptual audio encoder (monophonic).

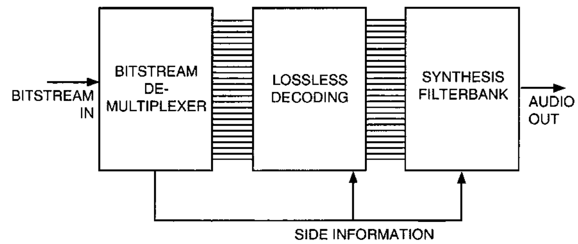

Fig. 7. Generic audio decoder (monophonic).

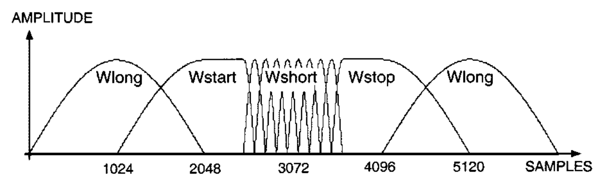

Fig. 8. Example for an adaptive window switching sequence in PAC.

The decoder reverses this process (Fig. 7). The three main functions are the following.

- The bitstream is parsed, yielding the coded spectral values and the side information.
- The lossless decoding of the spectral indices is performed, resulting in the quantized spectral values.
- The spectral values are transformed back into the time domain.

The filter banks used in perceptual coders such as PAC and MPEG-2 AAC are lapped transforms with adaptive window sizes [45]. These coders use a 1024-band modified discrete cosine transform (MDCT) [18] filter bank with a 2048-sample transform window. The size of the transform is chosen such that a high-frequency resolution is obtained. However, the corresponding time resolution will be low. Hence, during transient areas, e.g., when there is a signal onset within the frame (1024 samples), the coder switches to a shorter transform window of 256 samples for a 128-band MDCT to better track the signal changes. Thus, a frame is either encoded with a 1024-band MDCT or eight 128-band MDCTs. An adaptive window switching sequence is shown in Fig. 8. The long transform windows before and after switching to short windows have a different shape (Fig. 8) and are called transition windows. Some versions of perceptual coders use wavelet transforms instead of short MDCT transforms for increased coding efficiency [38].

III. Design Advances Matched to Application-Specific Requirements

A. Matching Coding Technology to Broadcast Material

Given the wide variety of program material in radio broadcasting, the choice of source coding technology is an important consideration. The broadcast environment includes virtually all types of acoustic material one can imagine, from music to synthetic sound effects to speech to noise. The environments of these sources may include carefully controlled studio productions as well as live productions such as sports and outdoor concert events.

As an illustration, even those categorized in the limited class of "speech-only" signals in broadcasting do not necessarily behave as speech signals considered in communication applications. In communications applications, signals are usually carefully acquired, bandwidth-limited, level equalized, and filtered in known fashions. In the broadcast environment, audio bandwidth, signal levels, and equalization can vary unpredictably. In addition, many digital radio stations use classic nonlinear preprocessing techniques from the legacy analog systems [46]. These nonlinear techniques are used to give stations a perceptual distinction, i.e., the so-called *signature sound*.

It therefore follows that traditional audio coders that do not make assumptions on the production (source model) mechanism are a better match for audio broadcasting applications. However, as mentioned, at lower bit rates, weaknesses become apparent in audio coders, in particular for speech. Common problems with applying audio codecs to speech include distortions such as pre-echos and post-echoes. In addition, there is a reverberant distortion (ghost image) produced when speech is compressed at very low rates by transform type coders. Some *expert* listeners can also perceive a loss in the "fullness" of the decoded compressed speech.

The audio coders can be modified to improve quality on speech at low rates. For example, several new technologies have been incorporated into the PAC audio coder. A new enhanced algorithm for pre-echo control reduces the spread of quantization effects in time, thereby decreasing the pre-echos or reverberation effects with speech. Additionally, the short block mode in an audio coder can be enhanced for stationary signals by new Huffman coding schemes and more accurate parameterization of the masked threshold. In this case, the coders use the short-block mode more frequently without harming the quality for stationary signals while improving quality for quasi-stationary signals such as speech.

Stereo coding is also a problem at the lower bit rates. In the current perceptual audio coders, the stereo coding scheme is largely designed for encoding audio signals at transparent audio quality, i.e., when the quantization noise is below both the left and right channel masked thresholds. The left and right masked thresholds are computed by a binaural

Table 1
Coding Paradigms and Audio Formats

Bitrate (kb/s)	Audio BW (Hz)	Material	Coding Technique
8-16	100-4k (mono)	traffic stock info	speech
24-32	50-7k (mono)	talk radio AM-Stations	speech or audio
32-48	20-10k (mono)	talk radio/ FM or AM	audio
48-64	20-16k (stereo)	music and speech	audio* (See III-G)
64-72	20-16k (stereo)	music and speech	audio
96-128	20-20k (stereo)	music and speech	audio

perceptual model which takes into account reductions in masking level when left and right signals are correlated; this is known as *binaural masking level difference* (BMLD). However, when operating at nontransparent quality, the effects of quantization noise on the stereo image are less well understood. Often heuristic techniques that balance multiple considerations are used.

Given the tradeoffs in the technologies, it is important to carefully match coding paradigms to the general formats that will be seen in the audio broadcasting applications. Table 1 outlines a general summary of the formats, bit rates, and the potential source coding technologies.

Within each bit-rate range, further optimization of the bit-rate/bandwidth/distortion tradeoff can be made. Fig. 9 shows a contour plot of the long-term average bit rate as a function of the distortion and the audio bandwidth for a stereo PAC implementation. The estimates are made over a representative mix of audio signals (both music and speech) used in audio broadcasting. The conclusions are therefore general and may differ for a specific piece of audio. For high-quality stereo music radio channels, typically, bit rates within the range of 56–96 kb/s are used. For each bit rate, different tradeoffs between the audio bandwidth and the amount of perceived distortion can be chosen using Fig. 9. For example, for a bit rate of 56 kb/s, a stereo signal could be encoded with bandwidth B and distortion D (derived from Fig. 9): $B = 10$ kHz, $D = 0$; $B = 11$ kHz, $D = 20$; and $B = 12$ kHz, $D = 38$. The tradeoff is chosen such that the impairments of the coded audio signal resulting from reduced bandwidth and added coding distortion are about the same. It has to be noted that the perception of tradeoff between bandwidth and coding distortion is a highly subjective matter. It also depends on the listening environment. For example, in a noisy car, more distortion can be tolerated than in a quiet listening environment.

B. Variable Bit-Rate Coding Versus Constant Bit-Rate Transmission

Typical nonstationary signals such as audio signals have a varying amount of inherent perceptual entropy as a function of time [47]. Variable bit-rate compression techniques are therefore natural means of approaching the compression limit of audio signals (i.e., perceptual entropy for transparent

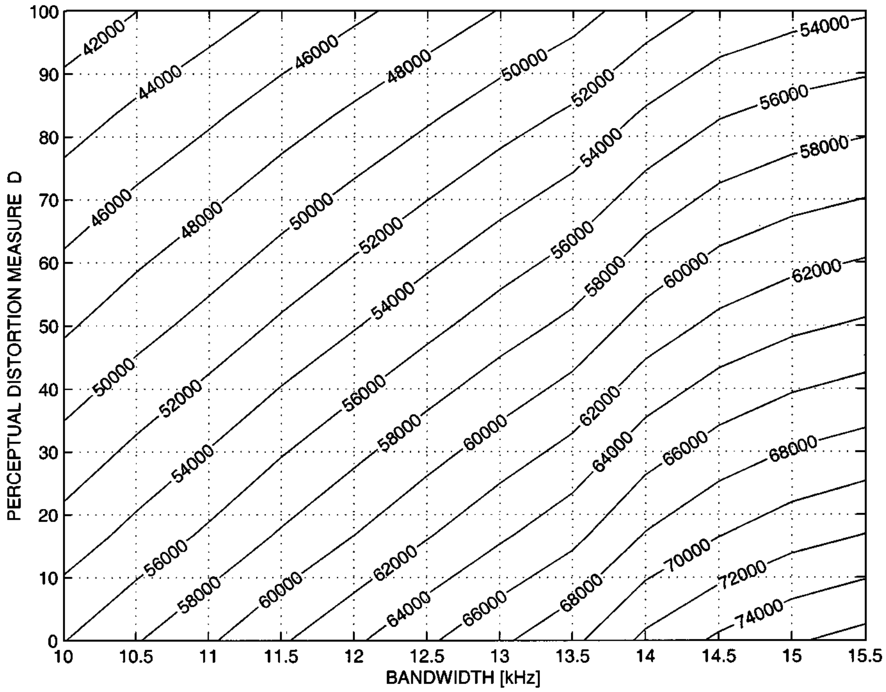

Fig. 9. An example of the dependency of the bit rate (indicated on contours in b/s) on the distortion and bandwidth for the PAC audio coder for stereo music signals. Transparent quality corresponds to a value of $D = 0$, and "Annoying" quality corresponds to a value $D = 100$.

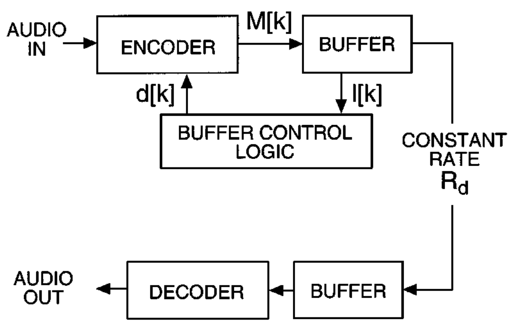

Fig. 10. An audio encoder and decoder with a constant bit-rate transmission channel.

audio coding). However, most broadcasting applications require a constant bit-rate transmission. When a variable bit-rate source coder is used together with a constant bit-rate transmission channel, the output of the source coder needs to be buffered to absorb the variations in the bit rate.

Fig. 10 shows an audio encoder and decoder with a buffered bitstream to enable a constant bit-rate transmission. In this scenario, at each frame, k bits from the audio encoder are put into a first-in–first-out (FIFO) buffer at a variable bit rate of $M[k]$ b per frame from the source coder, and bits are removed from the FIFO buffer at a constant bit rate of R_d b per frame where R_d is equal to the rate of the transmission channel. The number of data bits in the buffer after the processing of frame k, $l[k]$, can be expressed iteratively as

$$l[k] = l[k-1] + M[k] - R_d \tag{1}$$

assuming some initial buffer level of $l[0]$ bits.

The buffer itself represents an interesting tradeoff influencing the source coder design. The larger the buffer size, the more variations in bit rate can be absorbed and the less the impact is to the audio quality due to constant bit-rate transmission. However, as mentioned in Section II-A, the size of the buffer is restricted by constraints on tuning delay and cost. In such a system, *buffer control logic* is necessary. This mechanism monitors the buffer level $l[k]$ and influences the encoding process to make sure the buffer does not overflow. Buffer underflow is less severe and can be always prevented by injecting additional bits into the frame.

The ultimate goal of the buffer control is to provide the best possible perceptual quality for a given buffer size restriction. To influence the encoding process and $M[k]$ in a perceptually meaningful way, the buffer control logic determines a level of quantization distortion in frame k through a perceptual criterion $d[k]$. The distortion criterion $d[k]$ determines how much noise is added above the masked threshold. If $d[k] = 0$, then frame k is encoded with coding distortion just below the masked threshold. If $d[k] > 0$, the coding distortion is allowed to exceed the masked threshold. In general, the larger the value of $d[k]$, the smaller the number of bits that will be required to encode frame k. The criterion $d[k]$ therefore regulates the bit rate coming out of the source encoder.

To select the required value of *d[k]*, many buffer control schemes for audio coders typically use two processing loops [5], [48], [35]. The *outer loop* determines for each frame k a bit rate $M_d[k]$ at which the frame should be encoded. The bit rate $M_d[k]$ is computed as a function of the buffer level $l[k-1]$ and the perceptual entropy or a related measure [47] of the frame. The *inner loop* then iteratively reencodes the frame at different levels of distortion $d[k]$ until the bit rate

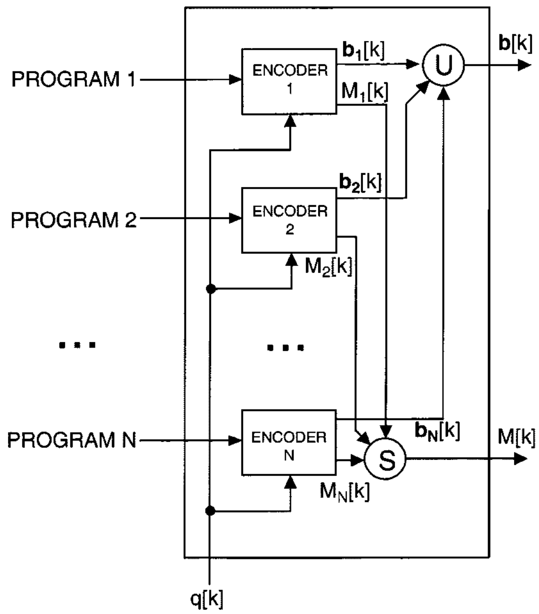

Fig. 11. The bitstreams of the N encoders are combined. The bit rate of a joint frame is $J[k]$. A single common distortion criterion is used.

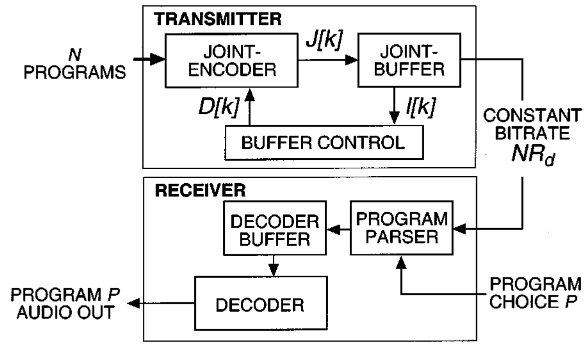

Fig. 12. The joint encoder is treated as a single encoder. The bitstream parser at the receiver extracts the bitstream of a specific radio program P.

of the encoded frame $M[k]$ is sufficiently close to $M_d[k]$, keeping $d[k]$ to a minimum.

Typically, the outer loop determines the bit rate of each frame $M_d[k]$ using a strategy that keeps a fairly low buffer level (a low buffer level means that many bits are available). This is a done in anticipation of critical frames such as transients which may have locally high bit demands. This approach is largely heuristic and may not explicitly reduce the local variation in distortion. A more efficient approach is to reduce the variations in distortions due to buffer control by using statistical bit-rate estimations. This approach is described in detail in [49]. In addition to reducing the variation of distortions over time, this approach is also significantly less complex than iterative schemes.

C. Joint Bitstream Transmission

Satellite digital radio services (see Section VII) broadcast a large number of radio programs (up to 100) simultaneously. In these situations, better performance (i.e., a larger number of programs and/or better audio quality of the programs) can be achieved if $N > 1$ radio programs are encoded jointly with a shared bitstream. That is, it is better if N channels share a common stream at NR_d kb/s than if each program is encoded individually, each with a single bitstream at R_d kb/s. To achieve this, a buffer-control scheme for joint coding is used which dynamically allocates the channel capacity between the audio coders sharing the common bitstream.

Fig. 11 shows how N audio encoders are connected to form a joint encoder with a joint bitstream. The bit rate of each joint frame $J[k]$ is the sum of the bit rates of the frames of the individual encoders $M_n[k]$ $(n \in \{1, 2, \ldots, N\})$

$$J[k] = \sum_{n=1}^{N} M_n[k]. \tag{2}$$

A distortion criterion $D[k]$ common to all encoders is used since it is simpler than dealing with a separate distortion criterion for each encoder. In addition, by having the same *perceptual* distortion criterion, the buffer control has the same average quality/bit-rate impact on each audio encoder. Note that it is also possible to consider different criteria for each encoder.

Except for the use of multiple audio inputs, the operation of the joint encoder of Fig. 11 is similar to a single audio encoder. A buffered joint encoding scheme with a receiver is shown in Fig. 12. The joint frames of the joint encoder are put into the FIFO *joint buffer*. A *buffer-control* scheme determines $D[k]$ such that the buffer level does not overflow. The bits in the joint buffer are transmitted to the receiver with a constant bit rate NR_d. Once a joint frame arrives at the receiver, the bits of the desired radio program P are extracted and placed into the *decoder buffer* by the *program parser*.

One reason the joint scheme is preferred is that the statistics of the joint bit rates $J[k]$ are much more favorable than those of the average individual channel. For example, assume that the bit rates of the single audio coders $M_n[k]$ $(n \in \{1, 2, \ldots, N\})$ are independent random variables with means $m_i = m$ and variances $\sigma_n^2 = \sigma^2$. It then follows that the mean and variance of the joint bit rate $J[k]$, as in (2), is Nm and $N\sigma^2$, respectively. Assume also that the average bit rate available for one audio coder is R_d and, therefore, that the average bit rate available for the N audio coders is NR_d. The standard deviation of the bit rate normalized by the desired bit rate R_d for one audio coder is σ/R_d, whereas the standard deviation of the joint encoder bit rate normalized by the total available bit rate NR_d is only $\sigma/(\sqrt{N}\,R_d)$. Similarly, in cases where the bit rates of the audio coders have different statistics, one can still expect a reduction in the normalized standard deviation of the bit rate for the joint encoder. As a result, for the same performance, the joint buffer can be either smaller than N times the buffer size of a single audio coder or, for the same relative buffer size, better performance can be achieved by allowing for more variation in the instantaneous source bit rate.

A second important advantage of joint coding is that the different audio coders can operate at different average bit rates according to the individual demands of their audio inputs. The dependence of the perceived quality of the decoded audio on each channel's program material is greatly reduced [50].

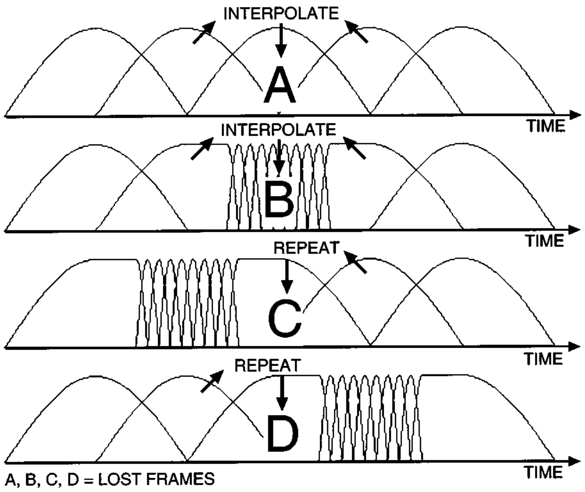

Fig. 13. An example of four different cases for mitigating a lost frame.

D. Error Mitigation

Some channel conditions (such as Rayleigh fading) will introduce bursts of residual errors which cannot be corrected by the channel codes. In those cases where these errors affect critically sensitive bits, the best course of action is to declare all the information for a given frame to be lost or erased and to mitigate the error by a frame-erasure concealment strategy. These concealment algorithms estimate of the missing portions of the waveform and in general can be quite effective once erasures occur at relatively low rates (a few percent) and do not span large intervals of time ($\geq$60 ms).

In some coders, e.g., speech coders, information about the waveform is represented by parameters representing the time and spectral structure of the signal. Such structures usually change in predictable ways from frame to frame. Concealment in such coders is often therefore done by making estimates of the missing parameters and using these estimates in the source decoder, possibly with minor modifications such as attenuation, to generate the output waveform [30], [31], [51].

Other coders, such as perceptual audio coders, do not explicitly represent structure through parameters. Due to the nonparametric interpretation of the decoded information, it is much more difficult to come up with a good mitigation strategy for these coders. On the other hand, because of the potential flexibility in delay constraints in a broadcast application, it is possible to recover lost information based on past and future information, i.e., by interpolation.

Fig. 13 shows four examples of five successive frames of an audio coder such as PAC or MPEG-2 AAC. Either one long transform window is used or eight short transform windows for encoding one frame. Long MDCT windows are used for encoding stationary parts of an audio signal and short MDCT windows are used to encode transients. In case A of Fig. 13, a frame with a long window is lost. In this case, the lost frame and its adjacent frames represent a stationary signal, and good results can still be achieved by substituting the lost frame with a frame obtained by interpolating the spectral content of the adjacent frames. In case B of Fig. 13, a frame with short windows is lost. The lost frame contained a transient but its adjacent frames are stationary. Therefore, good results can be achieved by substituting the lost frame with a long window frame obtained by interpolating the spectral content of the adjacent frames. In case C of Fig. 13, a frame with a long window is lost. The lost frame is preceded by a transient. Repeating the transient of the preceding frame would likely be perceived as an artifact. Therefore, the future frame is used to predict the present frame. In case D of Fig. 13, a frame with a long window is lost. The lost frame is followed by a transient. Repeating the transient of the following frame would introduce an echo and likely be perceived as an artifact. Therefore, the previous frame is repeated instead.

Finally, it is worth noting that, since frames are coded with a variable number of source bits, the fixed block length of the error-detecting codes may flag errors in subsets of a single source coding frame or may flag a group of source coding frames simultaneously. This opens the possibility of having partially decodable frames and/or bursts of frame erasures. Some of the aforementioned techniques can be applied to these cases with minor modifications.

E. Embedded and Multistream Audio Coding Schemes

In embedded and multidescriptive (stream) audio coding, the bitstream of the source coder is divided into a number of subsets that can be transmitted over independent channels. The subsets can be combined into various subbitstreams, each of which can be independently decoded.

In multidescriptive coding, each subset is a subbitstream that can be decoded independently. Multiple subsets can also be combined and decoded together to get higher quality. In the case of embedded coding, these subsets, or layers, have a hierarchy. The first layer, the “core” layer, is essential to all descriptions (i.e., subsequent layers of the bitstream) and can be used on its own to produce a decoded output. All other “enhancement” layers can be combined with the core and then decoded to produce output with increased quality. The enhancement layers may be themselves ordered though, like multidescriptive coding, the layers may be combined in various ways.

In the example of Fig. 14, the bitstream is divided into the following.

- *CORE*: This is the core part of the bitstream. It is self-sufficient and can be decoded independently of the other substreams.
- *Enhancement Layer 1*: This consists of encoded high-frequency spectral coefficients. This subbitstream enhances the audio bandwidth of the core.
- *Enhancement Layer 2*: This consists of encoded left–right difference spectral coefficients. Given these, the core can be enhanced from mono to stereo.
- *Enhancement Layer 3*: This consists of encoded high-frequency left–right difference spectral coefficients. Given 1 and 2 and this subbitstream, the core is enhanced to high audio bandwidth stereo.

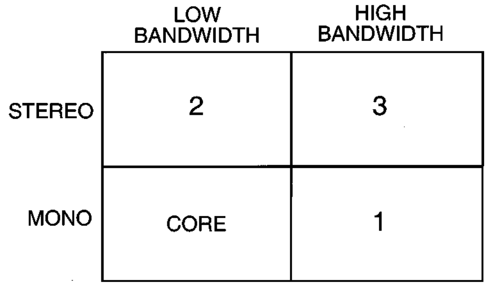

Fig. 14. The core bitstream provides basic audio quality. Bitstreams 1–3 enhance the audio quality.

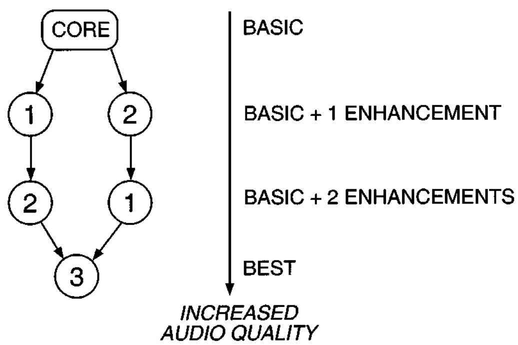

Fig. 15. The core bitstream can be enhanced in different ways for better audio quality: embedded audio coding.

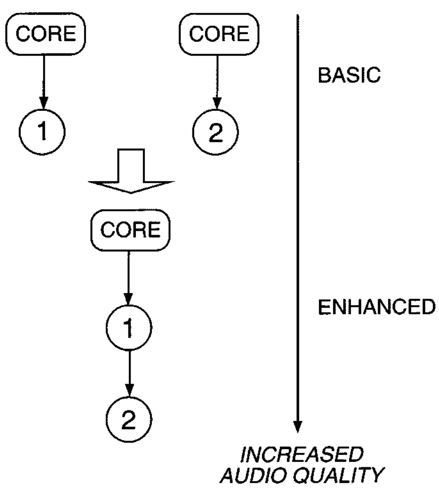

Fig. 16. Multistream audio coding: two independent bitstreams combined yield a bitstream with enhanced quality.

For embedded audio coding, the subbitstreams of Fig. 14 can be used as shown in Fig. 15. The core can be combined with the different subbitstreams to enhance the audio quality. For multistream audio coding, several independent bitstreams are formed, given the building blocks of Fig. 14. For example, Fig. 16 shows two independent bitstreams ($\mathrm{CORE}+1$, $\mathrm{CORE}+2$) which, when combined, yield enhanced audio quality ($\mathrm{CORE}+1+2$). Another possibility for multistream audio coding is encoding the audio signal using complementary quantizers [52] and sending the information from each quantizer in different streams. If information from both quantizers are received, then the quantizers are combined and the audio signal is decoded with less distortion.

F. Unequal Error Protection

The embedded bitstream formats just mentioned imply a hierarchy in bits in terms of each bit's influence on the decoded quality. For example, some bits add high-frequency information, while others add stereo information, etc. It is also well known that, in most source bitstreams, both nonembedded and embedded, individual bits also have an unequal profile in terms of *bit-error sensitivity*, i.e., the degree of quality loss when a particular bit is decoded with an error. This unequal sensitivity among bits can be exploited by the channel coding by using unequal error protection (UEP).

To implement a UEP channel coding scheme, the source bits are divided into different classes. The source coder can implicitly specify these classes by simply ordering the positions of the bits in a stream accordingly [53], [54]. Each class is protected by a different error-correcting code with the more sensitive bit classes protected by the stronger (e.g., higher rate) channel codes. This is in contrast to an equal error protection (EEP) scheme which uses a single channel code to protect all bits equally.

In general, a true EEP is rarely used since there is often a subset of critical bits (for example, the coding mode, the transform length, and the framing information) that need to have a higher level of protection. This subset is often further protected by an error-detecting code. If an error is detected, a frame erasure is invoked. However, even considering these enhanced "EEP" schemes, it has been shown that PAC and other coders perform better using true multiclass UEP schemes that take into account more information on differences in bit-error sensitivity [54], [55]. A further discussion of UEP is given in Section V.

G. Further Developments in Coding Algorithms

It is worth noting that there are continued, less traditional, developments in audio compression technology that are receiving more attention in the broadcast environment, in particular because of the high compression ratios which are required in some applications. Traditional audio coding designs have often focused on minimizing the bit rate while maintaining perceptual considerations focused on transparent audio quality. The tradeoff, as outlined in Section III-A, is often made between the bit rate, the acoustic bandwidth, and/or the number of channels (e.g., stereophonic vs monophonic) coded. The newer tradeoffs considered allow a greater degree of flexibility, allowing designers to further reduce bit rate while at the same time maintaining good "nontransparent" audio quality.

One such technique is the use of *bandwidth extension* techniques [56], [9]. These techniques try to synthesize or "fill in" the higher frequency acoustic information (not transmitted) based on received information in lower frequency acoustic bands. While such techniques can never ensure that the higher frequency information is similar to

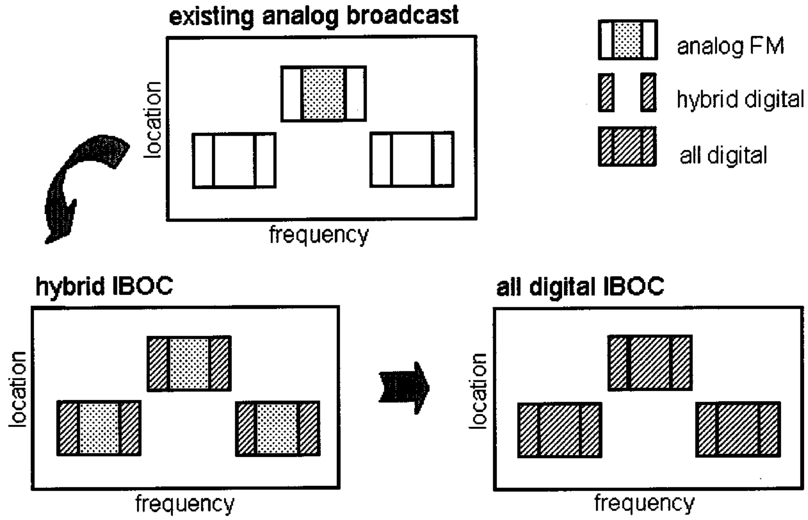

Fig. 17. IBOC transition from analog to digital.

that in the original uncompressed source material, the technique does create a "natural" impression of higher acoustic bandwidth. This allows designers to maintain targets of perceived acoustic bandwidth while saving bits to improve quality in the lower, more important acoustic bands.

Another interesting technique is that of *binaural cue coding* (BCC), [57]–[60]. This technique makes compromises in the multichannel format by explicitly coding spatial cues between pairs of audio channels. In this way the signal transmitted consists of a single audio signal (sum of all input channels) and very low-rate BCC side information. As a result, the decoded signal will not be transparent in terms of spatial image. However, it does produce a very natural spatial impression while allowing the majority of bits to be spent to improve the quality of the single audio channel transmitted. Such a technique is promising for achieving high-quality CD-like acoustic bandwidths at bit rates of approximately 40–50 kb/s while maintaining some, though not all, stereophonic properties.

The last set of techniques to mention comes from the area of speech coding technology. Here, wide-band coders such as the Multimode Transform Predictive Coder [61] and ITU-T Rec. G.722.1 [21] allow systems to make explicit compromises that improve speech quality while maintaining acceptable audio quality at bit rates of 16–32 kb/s. Such coders can be considered for programs in which the primary source material is speech.

IV. Terrestrial Systems

A. Introduction

In North America, terrestrial radio commonly refers to broadcast in the FM band (88–108 MHz) and the AM band (510–1710 kHz). To circumvent the difficulty in allocating a new spectrum for digital audio broadcasting over terrestrial channels and to allow current analog radio stations to migrate into digital transmission without causing disruption in consumer adaptation, the NAB has been supporting the development of IBOC technology. The argument for supporting the technology is mostly based on a migration plan that the NAB deems sensible and acceptable. In the plan, the migration to all-digital audio broadcasting will take two steps. The first step is to move from today's analog transmission to a hybrid system, which inserts digital signals along the two sidebands of the host analog signal. The second and final step is to totally replace the analog host signal with digital signals, which may carry additional services, as the market adapts gradually to the new system. Fig. 17 depicts such a strategy. In the following sections, we summarize the requirements and the progress in the past decade in the area of terrestrial digital audio broadcast.

B. Requirements

In promoting a digital radio system, one needs to set up the requirement of the system with clear enunciation of the potential benefit (and possible sacrifice) that the new technology shall bring about. The requirements for the IBOC system can be addressed along several dimensions.

Coverage: The coverage of existing AM and FM stations, in reference to the contours limited by interference and by noise, shall not be compromised due to the digital signal in both hybrid and all-digital modes. In other words, the digital system must provide a service area that is at least equivalent to the host station's analog service area while simultaneously providing suitable protection in cochannel and adjacent channel situations. Such a requirement ensures market stability in the service areas.

Service Quality: Audio quality in both hybrid and all-digital modes shall be significantly better than that of existing analog AM and FM modes. In fact, an original appeal in moving to digital systems was the improvement in audio quality, potentially to the level of CD quality in FM systems and to the level of analog FM quality in AM systems.

Spectral Efficiency: Spectral efficiency provided by IBOC shall be better than existing AM and FM bands in both hybrid and all-digital modes. Spectral efficiency refers to the ratio between the source signal bandwidth and the transmission signal bandwidth at given audio quality.

Feature Set: Both the hybrid and the all-digital modes shall support a substantial set of new features such as auxiliary data channel and an automated public safety infrastructure (emergency alarm system, weather alerts, and traffic conditions).

Compatibility: Deployment of IBOC in either hybrid or all-digital mode shall not impact existing analog stations or analog receivers. Insertion of digital signals shall not create additional interference to the existing analog signal. The hybrid transmission mode shall be backward compatible with current analog receivers already in use (i.e., without mandatory upgrade on listeners' equipment if they are not prepared to receive digital programs), and the all-digital mode shall be backward compatible with hybrid IBOC receivers. In short, the system shall afford a smooth transition from analog to digital services. The IBOC migration plan discussed above is a commonly accepted plan.

These requirements provide a design guideline in the development of the hybrid system and the eventual goal of an all-digital system.

C. Evolution of IBOC in the USA

In the early 1990s, in light of the development of the Eureka-147 system in Europe, the Consumer Electronics Manufacturer's Association (CEMA) and proponents of Eureka-147 urged the National Radio Systems Committee (NRSC), jointly formed by the Consumer Electronics Association (CEA) sector of the Electronics Industry Association (EIA) and the National Association of Broadcasters (NAB), to consider a plan for digital audio services. A call for proposals was issued in 1991 to lay out possible technical approaches and a plan to test the proposed systems. Several systems were proposed, including the *L*-band Eureka-147 system at two different bit rates, an *S*-band satellite system, an in-band, adjacent-channel (IBAC) system, and various IBOC systems. The key idea of an IBAC system is to find vacant channels in the current AM and FM bands for digital broadcasting. Table 2 lists all the systems that participated in the test, some in the laboratory only and some in the field. The field test was conducted in 1994 in the city of San Francisco. It was determined that the current AM and FM bands are too "crowded" to accommodate a new digital channel for each station license holder as was done in the transition to digital in the TV band. The IBAC system was thus deemed unsuitable. The NRSC also concluded in 1995 that the technology had not yet progressed to a viable point and, in 1996, subsequently suspended its activity until sufficient progress could be shown to warrant renewal of activities.

We must note that the unsatisfactory performance of early digital audio radio systems is mostly due to the relatively high bit rates needed for audio coding. The lowest audio-coding rate attempted in these systems was 128 kb/s, which could not be supported by the digital transmission scheme. Clearly, given the power and interference requirements dictated by the coverage map authorized by the FCC, a much more efficient audio coding algorithm would have to be developed before IBOC digital radio services could become viable.

Table 2
Submitted Systems for the 1994 IBOC Test

System	Band
Eureka-147 (224 kb/s)	L-band
Eureka-147 (192 kb/s)	L-band
VOA/JPL (satellite system)	S-band
AT&T/Lucent IBAC	VHF (FM)
AT&T/Lucent/Amati IBOC (double sideband)	VHF (FM)
AT&T/Lucent/Amati IBOC (lower sideband)	VHF (FM)
USADR FM-1 IBOC	VHF (FM)
USADR FM-2 IBOC	VHF (FM)
USADR AM IBOC	MF (AM)

As the spectral allocation issue became more prominent, the NAB in the mid-1990s started to focus on IBOC systems. In the mean time, advances in perceptual audio coding and orthogonal frequency division multiplexing (OFDM) or digital multitone technology for digital transmission (such as used in Eureka-147) had inspired new hope for the IBOC system. In 1996, audio coders like PAC [38] and MPEG-2 AAC [35] were shown to be able to code stereo music at 96 kb/s without causing audible degradation from original CD materials [39]. These advances inspired a collaboration between two of the original proponents of the IBOC system, USA Digital Radio (USADR) and Lucent Technologies, to join forces to develop a working IBOC system in 1997. In late 1997, a new company, Digital Radio Express (DRE), contacted the NRSC with the claim of possessing viable designs of FM and AM IBOC systems, and the NRSC on IBOC was thus reactivated in February of 1998.

USADR and Lucent subsequently separated in 1999, although development efforts continued in each individual company. In 1999, Lucent Technologies, taking advantage of its research program in audio coding and digital transmission, formed Lucent Digital Radio (LDR) to signify its commitment to this particular technology area. LDR moved rapidly into a new system, with key advances such as multistream audio coding, which can be considered a new generation system. Key components of these systems will be addressed in the following sections.

During 1998 and 1999, the NRSC established Test and Evaluation Guideline documents to assist the technology proponents in self-testing programs so as to identify information that would be needed by the NRSC to validate the viability of the new system. In August 2000, a formal Request for Proposal (RFP) on IBOC was issued to solicit submission of system designs for consideration as a standard for the U.S. During this time, while technology development continued, a number of business mergers took place; DRE was merged into USADR in 1999, and, in 2000, the two remaining proponents, USADR and LDR, with somewhat different system designs, joined together to become a sole company called iBiquity Digital Corp. [62]. Attributes of both systems have been combined, and, in August 2001, test results were presented to the NRSC. Based on the evaluation of these results, NRSC made a recommendation for approval of the FM system to the FCC on November 2001 and the AM system in April 2002 [63]. Deployment of both AM and FM hybrid IBOC is scheduled for the 2002/2003 time frame. Several radio equipment transmitters have IBOC-compliant offers, and several receiver manufacturers have announced IBOC-ready radios.

D. Other Terrestrial Systems

The largest deployed terrestrial digital audio radio services system is Eureka-147 [6], [64]–[66]. This system was the outcome of a large European consortium activity in the early 1980s. The project was done in the context of the Eureka series of research projects, and project 147 began in 1986 to develop a digital audio broadcasting system. The system specification was finalized in 1994 and was adopted as a worldwide ITU-R standard in 1994 and as an ETSI standard in 1997. The system is operational is many Western European countries and Canada, and deployment is scheduled in several Asian countries and Australia. Receivers are widely available and prices are on the order of $200–$300. Eureka-147 is different from IBOC in many ways. Rather than using existing AM and FM bands, it assumes newly allocated bands. To obtain efficient frequency use, several programs are multiplexed and transmitted on a single carrier. Such an ensemble has a transmission bandwidth of 1.536 MHz. Using OFDM modulation (using differential quadrature phase-shift keying (QPSK) for each carrier), the gross capacity of this ensemble is about 2.3 Mb/s. Varying levels of error protection can be selected resulting in net bit rates of 0.6–1.8 Mb/s. Error protection levels can be set for individual programs within an ensemble. Its audio compression scheme relies on MPEG 1, 2 Layer II, which requires 128–192 kb/s for stereo audio broadcasts. It supports both 48- and 24-kHz sampling frequencies and bit rates from 8 to 384 kb/s in mono, stereo, and dual-channel mode. Its basic frame size is 24 ms. Besides audio, Eureka-147 supports program associated data and generic data. The latter is organized in 24-ms logical frames with a data rate of n times 8 kb/s. The system has been designed for mobile reception over a wide range of frequencies (30 MHz and 3 GHz). This has been accomplished by providing four transmission modes, each using a different number of carriers, frame duration, and symbol duration. Transmission modes I and II are the most suitable for terrestrial broadcasting, while mode III can be used for cable and satellite broadcasts. Various frequencies have been allocated at WARC-92, and most countries either transmit in the VHF band or the *L*-band. Due to its robust design against multifading, it is possible to operate in a so-called *single frequency network* (SFN) mode, where several (geographically separated) transmitters all broadcast the same ensemble at the same frequency. This allows robust coverage of a large area. Another advantage of an SFN is that it provides a very power-efficient network compared to (analog) FM for the same coverage efficiency. The need for multiplexing requires careful coordination between content providers and collective responsibility for the transmitter infrastructure. This approach has been found quite feasible in Europe, where broadcasting in general has been organized at a national level and is intended for national coverage. This in contrast to the U.S. where local programming is preferred. Despite a tremendous effort from various governments and its wide availability, its success has so far been limited. Although established for sound delivery, its most successful applications rely on it as a robust high-speed wireless data delivery service. Recent proposals of combining Eureka-147 with GPRS have indicated that this is a viable commercial option.

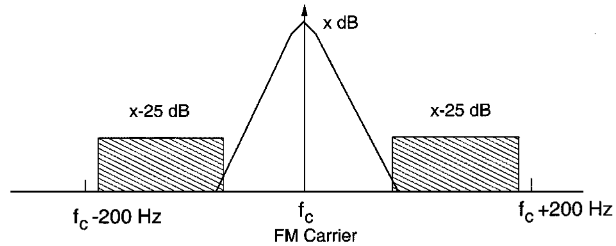

Fig. 18. Basic FM power spectrum.

The Digital Radio Mondiale consortium [8] has developed a system for digital broadcasting at frequencies below 30 MHz. This system [9] has already been recognized by the ITU in a draft recommendation [7]. The U.S. hybrid and all-digital IBOC AM systems are also part of this recommendation. The DRM system has been developed based on the following key requirements.

1) The audio quality must be improved over that achieved by analog AM.
2) The DRM signal must fit within the present channel arrangements in the AM bands.
3) The DRM signal should support operation of an SFN.
4) The DRM signal should support operation of an SFN.

The capacity available for audio within a single 9- or 10-kHz (U.S.) AM channel is limited. Audio coding rates from as low as 10 kb/s up to mid-20 kb/s have been proposed. For the lower rates speech coders can be used, while for the higher rates MPEG-2 AAC with spectral band replication (SBR) is used [9]. Data and audio bitstreams are multiplexed. The transmission system is based on OFDM, which avoids the need for adaptive equalization. Constellation sizes varying from 16 QAM (4 b/s/Hz) to 64 QAM (6 b/s/Hz) have been proposed. The channel coding used in the system is multilevel coding. A variety of OFDM configurations system bandwidths and data rates have been included in the standard. For further details on the DRM system, see [8]. Deployment of DRM services is scheduled for 2003.

V. IBOC FM Systems

Digital broadcasting in the FM band inside the FCC emission mask can take place in a so-called hybrid IBOC system where the digital information is transmitted at a lower power level (typically 25 dB lower) than the analog host FM signal. This digital transmission is achieved in subbands on both sides of the analog host signal. The composite signal is typically 400 kHz wide with the FM carrier in the middle. The digital sidebands are typically about 70 kHz wide at the upper and lower edges of the composite signal (see Fig. 18).

One current design proposal for hybrid IBOC FM systems uses a single 96-kb/s PAC [38], [40], [67] source stream duplicated for transmission over two sidebands using OFDM modulation. A uniform OFDM power profile is used. The channel coding on each sideband is rate 4/5 with memory 6. The total combined rate is 2/5, in a complementary punctured pair convolutional (CPPC) channel coding configuration [68]–[70]. Details of these convolutional codes will be provided below.

To ensure graceful degradation in the presence of severe one-sided first adjacent channel interference, an alternative system uses multistream transmission [71], [72] on the two sidebands combined with multidescriptive audio coding [72]. Further robustness to this type of interference is obtained by introducing a bit error sensitivity classifier in the audio coding algorithm and by transmitting bits in separate classes with different channel codes and different frequency bands [55]. More powerful channel codes [73]–[75], [53], [76], [77] and sideband time diversity give further improvements, especially for slow fading [78].

In Sections V-A and -B, we will give a detailed description of both single-stream and multistream hybrid IBOC-FM digital audio broadcasting systems.

A. Single Source Stream Systems

Hybrid IBOC broadcasting systems for digital audio radio have the capability of simultaneously transmitting analog FM and digital audio of CD-like quality. Due to fading and interference in the already crowded FM band, the signal design for the hybrid IBOC system is very challenging. It has been proposed to use a method of double sideband transmission where the digital information is transmitted by means of OFDM on both sides of the analog host FM and where the digital information can be recovered even when one sideband is partially or totally lost. This leads to an interesting channel coding problem of searching for optimal pairs of high-rate codes that form good combined low-rate codes which are better than classic code-combining techniques. Optimum in this context means channel codes with the best (longest) distance between codewords.

Hybrid IBOC systems have been under consideration for some time, and a number of prototypes have been designed [67], built, and evaluated [79]–[82], [70]. (These systems were earlier referred to as IBOC systems. The term IBOC now refers to all-digital systems, which have no analog host signals in the FM or AM bands.)

Traditional channel coding methods that have been developed for either white noise channels or channels with a known, fixed interference power spectrum are not well suited for the interference environment in the normally crowded FM band. Fig. 19 shows a configuration where the digital audio information is transmitted on two sidebands, one on each side of the analog host. In the interference environment of the FM band, the so-called *first adjacent* channel analog interference facing some receivers may be so severe that the signal-to-interference ratio in one sideband falls well below the operating range (i.e., erasing one sideband) while other receivers may lose the other sideband (depending on geographic location).

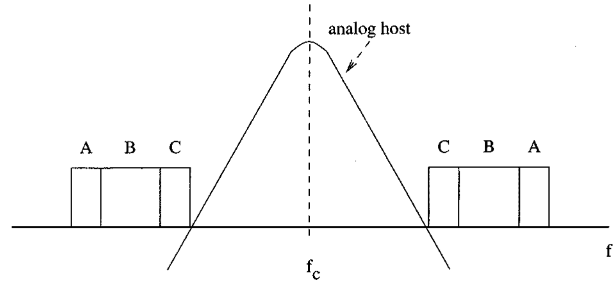

Fig. 19. Basic hybrid IBOC concept. Bands A–C of the OFDM carriers have different interference susceptibilities, and this impacts the channel code design. Band A is more sensitive to interference than band B. Band C is used optionally.

(A first adjacent interferer in FM is 200 kHz from the carrier and a second adjacent interferer is 400 kHz from the carrier.) Thus, one would like to be able to recover all the digital audio information even when either sideband is erased. On the other hand, if neither sideband is erased, one would like to exploit this advantage, for example, to receive the digital audio signal farther away from the transmitter, thus extending the coverage area. In other words, the interference environment is typically location-dependent, and the hybrid IBOC system should be able to adapt to the different scenarios.

Furthermore, even when the sidebands are not completely lost, the carriers within a sideband are exposed to differing levels of interference. For example, the carriers in bands B of the hybrid IBOC OFDM power spectrum in Fig. 19 are deemed to be more robust to interference. Bands A are always used but are deemed to be subject to more adjacent channel interference. Bands C are optionally used in the so-called extended bandwidth mode, yielding a potential increase in channel coding capability. Potentially, transmission in bands C can take place with precancellation techniques [83], [84] by which the self-interference from the analog host FM signal is canceled. The problem is to find the "best" pair of codes, called complementary codes, which together form another "good" code [68]. One code is transmitted on one sideband, and its complementary code is transmitted on the other sideband.

A well-known technique for generating good high-rate convolutional codes from low-rate codes is *puncturing* [73], [74], [85]. Fig. 20 shows an example of puncturing. Normally only one best high-rate code is sought. A low-rate mother code is first punctured to the full rate (full bandwidth) code used for both sidebands. This code is then in turn punctured to twice its original rate, forming the first code of the complementary pair. The punctured bits form the second code of the pair. The CPPC codes in this section are from [68]. These represent better schemes than classic code combining [73] and do not need to meet the conditions for the so-called equivalent code [86], which is a special subclass of complementary codes [70], [74], [75], [53]. The use of UEP codes [55] further improves the capability of the channel codes, yielding extended coverage areas. This leads to further code optimizations. Throughout this section ideal coherent QPSK and binary-PSK modulation and an additive white Gaussian noise (AWGN) channel have been assumed.

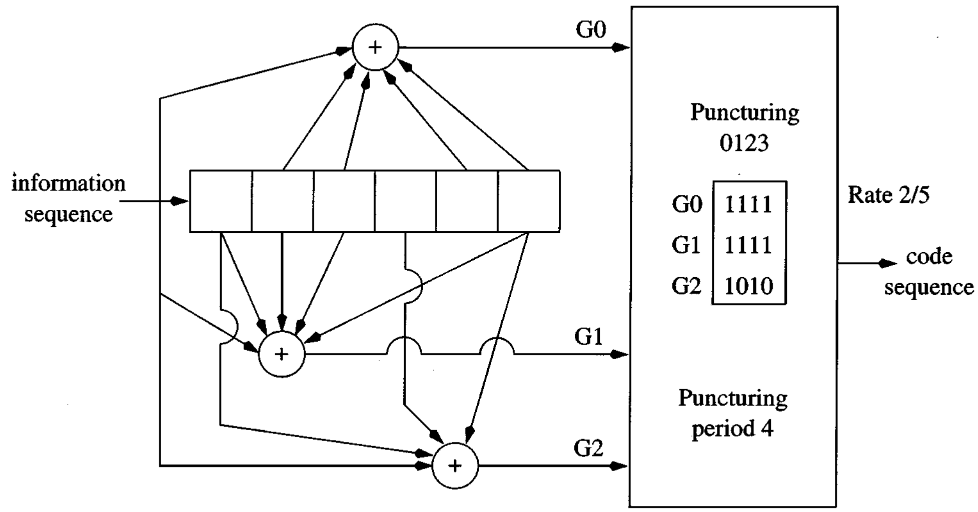

Fig. 20. Convolutional code notations. Rate-1/3, memory $M = 6$ mother code with puncturing period $P = 4$. The output code rate is 2/5.

This optimization then leads to "good" codes for realistic digital audio broadcasting fading channels using, e.g., differentially coherent four-phase modulation (DQPSK) and channel interleaving.

Puncturing a mother code is a well-known technique for obtaining good high-rate convolutional codes with easy decoding by means of the same basic Viterbi algorithm that is used for the mother code. Increased puncturing leads to higher code rates. Puncturing is often performed in a periodic manner with a pattern that is repeated with a period of P bits.

Complementary Punctured Pair Convolutional (CPPC) codes are defined as a pair of punctured codes of rate R_1 that are obtained by puncturing the same mother code with the same puncturing period such that the two codes have no unpunctured bits in common. Hence, the two codes combine to a rate $R_1/2$ code. A special subclass of these codes are so-called equivalent codes described by Kallel [86], which have the property that the puncturing pattern for one code is a cyclically shifted version of the puncturing pattern for the complementary code. It is not, however, necessary to add this constraint when searching for optimal CPPC codes. An "optimal" code is one having the best free (Hamming) distance [73] among those obtained by puncturing its particular mother code. If two codes have the same free distance, then the best code has the smallest information error weight [73], i.e., the smallest average number of bit errors corresponding to free distance error events. These codes are then optimal for transmission over the additive white Gaussian channel with soft decision decoding and BPSK or QPSK modulation at high channel SNRS. They are also optimal for the fully interleaved Rayleigh fading channel with ideal BPSK or QPSK modulation at high average channel SNRs.

The proposed hybrid IBOC system requires rate-4/5 *forward error correction* (FEC) codes for both the upper and lower sideband channels (half-bandwidth codes). These codes combine to form a rate-2/5 error-correction code (full-bandwidth code) [70].

Table 3
Rate-1/3 Mother Code Used for Puncturing

Code	Memory	g_0, g_1, g_2 (octal)	d_f	c_{d_f}/P
Hagenauer	6	133, 171, 145	14	1

Table 4
Properties of the Rate-2/5 Full-Bandwidth Codes

Code	Mother Code	Puncturing Pattern	d_f	c_{d_f}/P
Hagenauer	Hagenauer	(1111, 1111, 1100)	11	1.00
Kroeger	Hagenauer	(1111, 1111, 1010)	11	2.00

The rate 1/3 is the most natural starting rate for puncturing to rate 2/5. Several suitable rate-1/3 mother codes can be found in the literature [74], [85], [76], [77]. In this paper, we only report results obtained with the Hagenauer code. Results for the other codes can be found in [68]. The memory, generators, free distances, and information error weights for the mother code are given in Table 3. The free Hamming distance d_f is the smallest number of code bits that separate two different coded sequences (see [73]). The average information error weight c_{d_f}/P is the average number of information bit errors corresponding to free distance error events. The averaging takes place over all error events starting in any of P positions, where P is the puncturing period.

The full-bandwidth codes constructed are shown in Table 4, along with their free distances and information error weights. The two codes in Table 4, the Hagenauer *rate compatible punctured convolutional* (RCPC) rate-2/5 code [74] and the Kroeger rate-2/5 code [70], are taken from the literature. (The Hagenauer code is optimal in an RCPC sense, and the Kroeger rate-2/5 code gives good noncatastrophic rate-4/5 codes which we are reporting in this section.) These codes are punctured in a complementary fashion to form rate-4/5 CPPC codes. The optimal puncturing patterns are reported below. Other codes can be constructed using the search method from [68].

Table 5
Rate-4/5 Complementary Codes. Noncatastrophic Rate-4/5 Complementary Codes Found by Lucent That Combine to the Kroeger Rate-2/5 Code. $P = 4$

Puncturing Pattern	d_f	c_{d_f}/P	Complementary Pattern	d_f	c_{d_f}/P
(0110, 1001, 0010)	4	2.50	(1001, 0110, 1000)	4	2.50
(0110, 1001, 1000)	4	12.00	(1001, 0110, 0010)	4	12.00

Table 6
Comparisons of Free Distance for CPPC Codes and Classic Code Combining

	Best Rate 4/5	Code Combining R = 2/5	CPPC	
			R = 4/5	R = 2/5
$M = 6$	4	8	4	11
$M = 8$	5	10	5	14

Table 5 lists all noncatastrophic memory–6 complementary codes of puncturing period 4 that have the maximum worst case free distance and combine to the Kroeger rate-2/5 code, respectively. Note that the optimum pair (top line) in Table 5 has puncturing patterns that are cyclically shifted versions of each other and, thus, are "equivalent" in Kallel's sense [86]. These codes have equivalent distance properties. However, in general, optimal complementary codes need not to be equivalent [68].

An alternative approach to CPPC is code combining of two identical codes on the two sidebands. In this case, the high-rate code on one sideband can be optimized without the CPPC constraints. It turns out that, for the cases studied, the CPPC strategy is much better for the case of combining the two sidebands. A slightly lower error coefficient might be obtained for the best one sideband code, but the loss in free distance for the combined two sidebands is significant. Code combining doubles the effective free distance [85], while combining CPPC codes yields a better result. In Table 6, CPPC codes are compared to code combining for the rate-4/5 and 2/5 codes with $M = 6$ and $M = 8$. The asymptotic gain for a Gaussian channel for CPPC over code combining for the rate-2/5 codes is $10 \cdot \log(11/8) = 1.38$ dB for the $M = 6$ case and $10 \cdot \log(14/10) = 1.46$ dB for the $M = 8$ case [68].

The proposed single-stream hybrid IBOC system involves a multicarrier modem with varying interference susceptibility on the subcarriers. In particular, subcarriers farthest away from the analog host signal are most susceptible to interference. Fig. 21 shows a crude model of increasing first adjacent interference. Thus, the mapping of code bits to subcarrier frequencies can affect performance. This mapping of code bits is referred to as bit placement or bit assignment and is given in [68]. Reference [68] also describes some optimal CPPC codes for UEP.

B. Multistream Systems

Different approaches to hybrid IBOC-FM systems for digital audio broadcasting based on multistream transmission methodology and multidescriptive audio coding techniques are introduced in this section (see [71]). These ideas involve a lower per sideband audio coding rate than for single-stream systems and thus allow more powerful channel codes, resulting in robust transmission and graceful degradation in variable interference channels. By also using PFDM techniques combined with UEP and sideband time diversity, new hybrid IBOC-FM schemes are obtained with extended coverage and better peak audio quality than previously proposed.

The FM channel suffers from dispersion in both the time and frequency domains. In the time domain, very severe multipath with delay spread ranging between 3–30 μs has been measured in urban and suburban environments. This broad range of delay spread corresponds to 30–300-kHz channel coherence bandwidth, which is, at the upper limit, comparable to the signal spectrum, thereby introducing flat fades for low delay spread channels such as dense urban environments. In a worst case scenario, no frequency diversity scheme can mitigate the severe flat fading which may extend across the whole spectrum of the radio signal. In the frequency domain, frequency dispersion ranges between 0.2–15 Hz for very low to very high speed vehicles. For static channels, such as the link to a slowly moving vehicle, the channel varies very slowly in time and, therefore, time diversity schemes cannot combat various channel impairments such as selective and flat fading conditions.

Fig. 22 proposes a novel time–frequency distribution of the PAC substreams which is highly robust against various channel impairments and fading conditions. The system tries to achieve maximum diversity across both time and frequency dimensions within the allowable bandwidth and time delay using the multistream PAC format (see Section III-E, [72], and [55]). The elements in the improved systems are the following: multidescriptive (MD) audio coding with 64 kb/s per sideband, allowing for more powerful rate-1/2 channel coding combined with multistream (MS) transmission with two-level UEP and sideband time diversity.

Each of the four substreams corresponds to a nominal average source rate of 32 kb/s with an overall rate of 128 kb/s. To produce these four streams, the audio signal is first encoded using a multidescriptive scheme to produce two streams at 64 kb/s each. Each of the streams is then further subdivided into two substreams of equal sizes using a bitstream classifier, i.e., $\{I, II\}$, and $\{I', II'\}$. The resulting four streams are then transmitted over parts of the FM spectrum by means of OFDM. The most significant bits (streams I and I') are transmitted in the inner bands. At the transmitter side, the substreams I and II are mapped across the upper band, and the complementary substreams I' and II' are assigned to the lower band of the IBOC signal with a 3-s delay. In the four-stream system, there are several built-in digital blending modes which allow for graceful degradation. These modes are summarized in Table 7.

The SNR gains on a Gaussian channel with the rate-1/2 codes are shown in Table 8. Note that an $M = 6$, rate-2/5 (double-sided) code has been added in Table 8 for reference. It can be seen that the *one-sided* 64-kb/s rate-1/2 system with $M = 6$ is comparable to the 96 kb/s, double-sided, rate-2/5, $M = 6$ system. It can also be concluded from Table 8 that the $M \geq 8$ rate-1/2 systems are superior to the $M = 6$, rate-2/5 scheme. It is also interesting to conclude that the rate-1/2, $M = 6$, double-sided system with 128-kb/s audio is identical to the one-sided version in Table 8 and

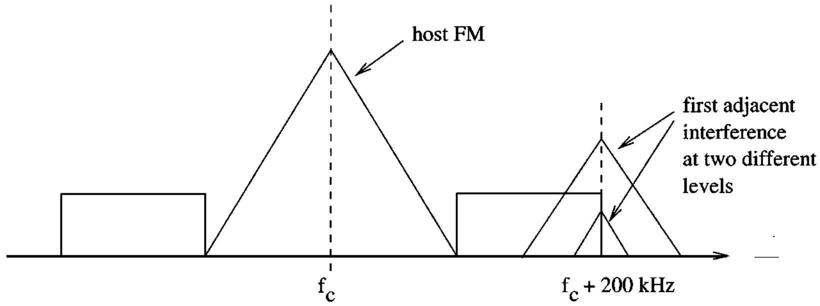

Fig. 21. Impact of a first adjacent FM interference signal at f_c+200 kHz at two different levels. Alternatively, the first adjacent interference may appear at f_c-200 kHz.

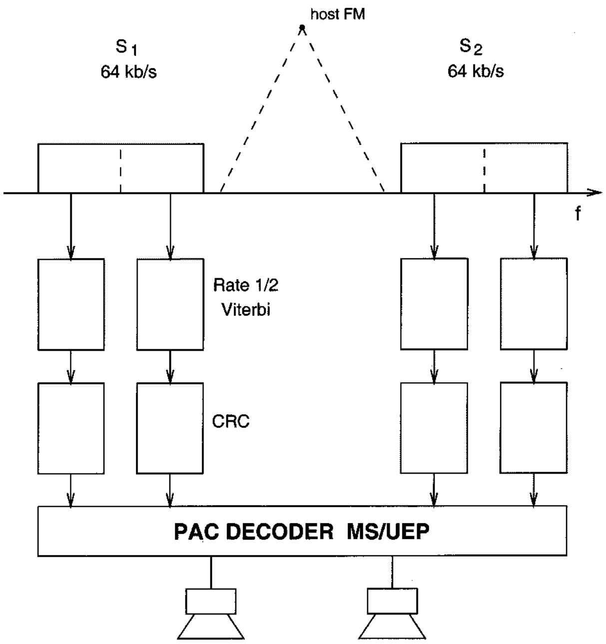

Fig. 22. Simplified block diagram for a proposed system based on 64-kb/s multidescriptive PAC and two-level UEP. The multistream transmission is done for four streams and the interleavers are not shown explicitly.

Table 7
Blend Modes in the Four-Stream Multistream Hybrid IBOC-FM Configurations. See Fig. 22 for Notations

Available Streams	Audio Quality
$I+I'+II+II'$	CD quality
$I+II+I'$ or $I'+II'+I$	CD like quality
$I+II$ or $I'+II'$	Better than Analog FM
$I+I'$	Analog FM like
I or I'	Better than Analog AM

Table 8 Gains With Rate-1/2 Codes on a Gaussian Channel With a Uniform Power Profile With $M = 10$ Codes, an Additional 0.6 dB in Gains, and With $M = 12$ Codes of 1.1 dB

Rate 1/2 Codes M	d_f	Gain in SNR relative to M = 6, R = 4/5*	Gain in SNR relative to M = 6, R = 2/5**
6	10	4.0 dB	−0.4 dB
8	12	4.8 dB	0.4 dB

$* \, d_f = 4$ $\quad ** \, d_f = 11$

thus comparable to the rate-2/5, $M = 6$, 96-kb/s system in asymptotic error rate performance for the Gaussian channel.

Table 9
Frame Throughput (in %) for Different PAC Rates and SNR (E_b/N_0) Values Under Fast Urban Channel Condition (5.2314-Hz Doppler)

PAC rate	SNR (E_b/N_0) 5 dB	6 dB	7 dB	8 dB
I or I' only	99.1	99.9	100	100
$I+II$ or $I'+II'$	98	99.3	100	100
$I+II+II'$ or $I'+II+II'$	79.3	91.7	96.7	98.9
$I+I'+II$ or $I+I'+II'$	80.7	93	97.8	99.4
$I+I'+II+II'$ (full rate PAC)	65	86	94.6	98.4

Table 10 Frame Throughput (in %) for Different PAC Rates and SNR (E_b/N_0) Values Under Slow Urban Channel Condition (0.1744-Hz Doppler)

PAC rate	SNR (E_b/N_0) 6 dB	7 dB	8 dB	10 dB
I or I' only	95.1	97.3	98.9	100
$I+II$ or $I'+II'$	86	91.6	95	100
$I+II+II'$ or $I'+II+II'$	60.3	68	76.3	83.6
$I+I'+II$ or $I+I'+II'$	57.8	67.3	76.3	87.3
$I+I'+II+II'$ (full rate PAC)	37	46.8	59.8	80.3

(There may not be "room" for a rate-1/2 code but rather a rate-8/15 code. Then, the gains in SNR will be somewhat smaller.) These gain numbers will be higher for interleaved fading channels.

End-to-end simulations were performed for the proposed multistream system under urban fast and urban slow fading channel models [71]. In these simulations, 1024 tones over 400 kHz were used with 500-ms interleaving, rate-1/2, $M = 6$ coding, and DQPSK differential modulation in frequency. PAC audio frames of 2000 encoded bits were considered, and the system performance was analyzed in terms of frame error rate versus SNR. A frame error (erasure) is defined when the error-detecting coder [cyclic redundancy check (CRC) in Fig. 22] detects an error. We used in our analysis the 9-ray EIA model with a 5.2314-Hz Doppler for urban fast and a 0.1744-Hz Doppler rate for urban slow [78]. The final results are listed in Tables 9 and 10. Note the robustness and graceful degradation.

A variety of power profiles are presented and discussed in [71] along with BER simulations for some of the coded multistream systems for a number of multipath fading IBOC channels.

C. Advanced Topics and Discussion

There are a number of techniques that may be employed to further improve both the single-stream and multistream systems. Two such ideas that have been mentioned above are the use of UEP [71], [72], [68] channel coding and/or nonuniform power profiles for OFDM. The latter technique may require regulatory approval.

Further improvements are obtainable by introducing the list Viterbi algorithm (LVA) for continuous transmission [87], [88] in the receiver. The LVA has the capability of reducing the frame error mitigation probability, resulting in improved audio quality. This LVA is backward compatible with a system using a standard Viterbi algorithm. In all the systems just mentioned, we assume that the host analog FM signal and the digital OFDM signals are nonoverlapping in frequency. In [83] and [84], principles for simultaneous transmission of digital data and analog FM are described and evaluated. The basic idea is that, since the transmitter knows the analog interference on the much lower (in power) digital signal, an adaptive precancellation algorithm can be employed to modify the digital signal so that the analog FM has no impact on the received digital signal. Near-optimum algorithms are presented in [83] and simpler suboptimum realizable algorithms are given in [84]. Thus, digital streams overlaying the analog FM can in principle be added. These can, for example, be used for further enhancement of the audio or to increase the available rate for the data channel.

There are a number of additional techniques that in principle can be employed to further upgrade IBOC digital audio broadcasting systems in the FM band. Here we will briefly point to a few of these ideas.

Turbo codes [89], [90] are in a class of more advanced channel codes which will give more robust systems at the expense of increased complexity and delay.

Soft combining methods [91], [72] can be used to increase the efficiency of channel decoding, in particular for CPPC codes for single-stream systems. Complexity and robustness versus performance is an issue here.

Screening method for undetected errors is another improvement possible by utilizing the inherent properties of the Huffman code to screen undetected errors (see [92]).

Cancellation or suppression of first adjacent interference from analog FM can in principle be achieved [72], [93] under certain circumstances. Both single-stream and multistream hybrid IBOC systems can be operated without this feature. With a well-functioning canceler, the operating range of the hybrid IBOC system may be extended. Again, there is also here a tradeoff between performance and complexity. An advantage is that a canceler may be optional at the receiver and backward compatible with a system without cancelers.

In summary, we have described the building blocks of both single-stream and multistream hybrid IBOC systems for digital audio broadcasting in the FM band. Both systems have their respective relative advantages and disadvantages. The final choice depends on a complex set of tradeoffs. Which of the advanced features to introduce also depends on tradeoffs between the level of improvement and the required complexity. Some of the techniques can be offered as backward compatible receiver options.

The system chosen to be offered to the U.S. market [62] is based on a single-stream system with CPPC type of channel codes and 96-kb/s PAC audio coding. A separate channel for data services is also provided. The NRSC endorses iBiquity's FM IBOC system and recommends FCC approval.

VI. IBOC AM Systems

This section describes proposed IBOC systems for digital broadcasting in the AM bands (535–1705 kHz). The AM systems differ from the FM systems in many aspects, particularly in terms of the nature of interference due to the modulation scheme. For the FM systems, the digital and analog signals are transmitted without overlapping in frequencies, whereas in the AM systems [72], [94] simultaneous transmission of analog and digital in the same frequency is not only possible but, because of linear analog modulation, it is also necessary because of the severe bandwidth limitations in the AM bands.

The radio channel for broadcasting to mobile receivers in the FM bands (and for cellular mobile radio) is well understood [71]. However, the AM channels are very different and less well understood for digital transmission to mobiles. First of all, daytime and nighttime conditions are very different. During daytime conditions, fairly good stable channels with interference slowly increasing and decreasing in certain bands are obtained when driving in the coverage area. The stable interference is caused by cochannel and adjacent channel interference from other AM or IBOC-AM stations. Impulsive noise should also be taken into account in the signal design. Doppler plays a small role in AM transmission in contrast to the FM case. Changes in the conditions of vehicular reception are caused, for example, by underpasses and power lines, etc. During nighttime, the AM channels can change rapidly due to skywave interference.

The carrier separation in the AM band in the United States is 10 kHz, with stations in the same geographical location separated by at least 20 kHz. That is, only every second adjacent band is assigned in the same city. In AM, the carrier of a first adjacent interfering station is 10 kHz apart and a second adjacent station is 20 kHz apart from the carrier frequency. Digital signals are to be transmitted along with the analog signal in a hybrid IBOC-AM system. To achieve FM like audio quality, an audio coder rate of 32–64 kb/s is required (see Section II). Therefore, bandwidth is extremely limited for the digital audio signal in a hybrid IBOC-AM system.

One proposal to transmit the digital audio signal on top of the analog AM signal consists of using a 30-kHz transmission bandwidth, as shown in Fig. 23, where the digital data are transmitted through bands A-C. In this case, severe second adjacent interference may occur in certain coverage areas and data transmitted in bands A or C can be lost com-

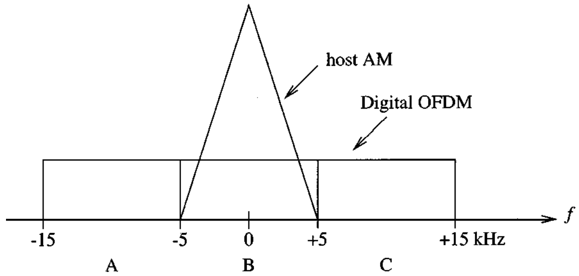

Fig. 23. Conceptual power spectrum of a 30-kHz, hybrid 1130C-AM. The digital data is transmitted in bands A-C.

pletely. For the FM case, as described in the previous section, the digital audio bitstream may be duplicated and transmitted on both sides of the analog host to provide a robust solution to this problem. However, in the AM case, there is not enough bandwidth to transmit a duplicated bitstream. Instead, [95] proposes a more robust strategy built on embedded/multi-descriptive audio coding and separate channel coding/modulation in several frequency bands to provide a remedy to this problem. In this proposed scheme, the audio decoder has the capability of blending down to a lower bit rate, when a certain subband in the hybrid IBOC-AM signal is subjected to severe interference. The design is such that the audio quality of this lower bit rate audio signal is better than that of analog AM. Thus, a graceful degradation is achieved along with a higher degree of robustness to certain channel conditions. The proposed scheme, called the *multistream transmission scheme*, is described in more detail in Section VI-A.

To avoid the second adjacent hybrid IBOC-AM interferer that has the same transmission power in the same geographical area, 20-kHz hybrid IBOC-AM systems are also proposed, and they are described in more detail in Section VI-B.

Similar to FM systems, the modem proposed for a hybrid IBOC-AM system is typically an OFDM modem. The modulation scheme proposed for daytime transmission is quadrature amplitude modulation (QAM) using 16-QAM, 32-QAM, or 64-QAM. For nighttime transmission, since the channel can change very rapidly due to skywave interference, 8-PSK modulation is proposed. The bandwidth and the transmission power are extremely limited in hybrid IBOC-AM systems. To protect the digital audio bitstream, bandwidth-efficient forward error-correction (FEC) schemes and coded modulation schemes have to be designed, and this is addressed in Section VI-C.

A. 30-kHz Hybrid IBOC-AM System

Fig. 23 shows the conceptual power spectrum for a 30-kHz hybrid IBOC-AM system. Depending on the OFDM modem tone allocation, signal set choices and FEC rates, the three frequency bands can carry different fractions of the total data rate. Assuming the same modem constellation (e.g., 32-QAM or 16-QAM) and the same coded modulation scheme and FEC rate used for all tones, bands A and C will carry 40% of the total data while B carries 20%, due to multiplexing with the analog host for the B band [95]. While bands A and C are normally expected to carry the same

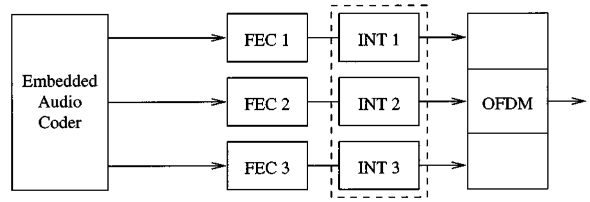

Fig. 24. Conceptual diagram of the multistream transmission system for an embedded audio coder/decoder.

data rate (due to symmetry), the relative data rate in band B (relative to that of bands A and C) is a design parameter [72], [94].

One possible interference scenario is second adjacent hybrid IBOC-AM to hybrid IBOC-AM interference where either band A or C has a sufficiently low signal-to-interference ratio that the symbols are effectively "erased" (jammed). When this happens, 40% of the symbols are lost. To require the channel code to recover the audio bitstream in this case is difficult and, at some point, the IBOC-AM is forced to blend to the analog AM signal.

One way to combat this type of interference is to employ embedded or multidescriptive audio coding and match layers of the bitstream to the multistream transmission. A system with such a scheme and multitone modulation is shown conceptually in Figs. 24 and 25. In this system, three parallel coding and modulation schemes are matched to the three bands A–C and the three layers in the bitstream, M, S_1, S_2, of the audio coder. In this case, M is considered to be the underlying essential (nonredundant) *core* of the description (see Section III-E).

Table 11 illustrates an example of the ideal bit-rate parameters for each layer of the audio coder's description. For example, in one system explored, the FEC units in Fig. 24 are concatenated outer RS codes and *inner trellis-coded modulation* (TCM) based on 16-, 32-, or even 64-QAM constellations [72], [96]. Due to the limited bandwidth available for hybrid IBOC-AM systems, alternate FEC schemes have also been explored [97], and the results are summarized in Section VI-C.

In such a scheme, the core M could ideally be transmitted in band B since this provides the most reliable channel. In reality, the division of bit rate among the bands may not match exactly the M, $S1$, and $S2$ bit rate divisions in the source description. If the core rate is larger than the capacity of the B band (e.g., in the case of an 8-kb/s core at a total rate of 32 kb/s), then some of the M core stream is transmitted in the A and C bands. If the core rate is smaller than the capacity of the B band, then some of the $S1$ and $S2$ layers can also be transmitted in the B band. Furthermore, in nonideal systems, there are compatibility problems with leakage from the digital signal to the analog AM host in the $\pm$5-kHz band. This then leads to other configurations, where the relative power levels of the OFDM tones in the $\pm$5-kHz band are reduced, and the core audio bits are transmitted in the S_1 and S_2 streams and the enhancement bits are transmitted in the M stream. The core bits can be identical in the two streams or a multidescriptive coder could be used.

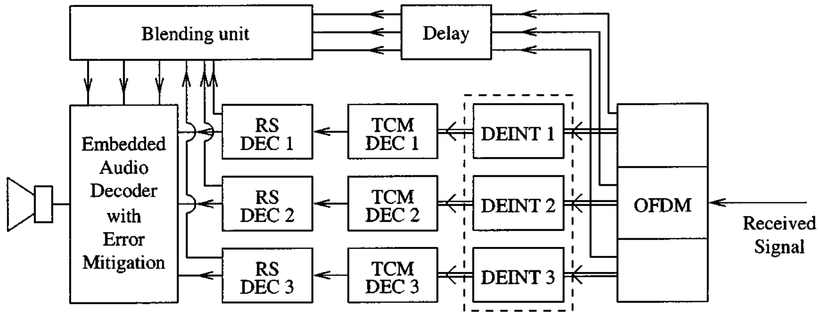

Fig. 25. Receiver for multistream transmission. Conceptual diagram. The OFDM demodulator also contains synchronization, training, equalization, and timing.

Table 11
Rate Allocation (in kb/s) for the Three Frequency Bands A–C. R is the Total Data Rate

	A	B	C
R	$\frac{2}{5}$ R kb/s 40%	$\frac{1}{5}$ R kb/s 20%	$\frac{2}{5}$ R kb/s 40%
48	19.2	9.6	19.2
32	12.8	6.4	12.8
36	14.4	7.2	14.4
40	16	8	16

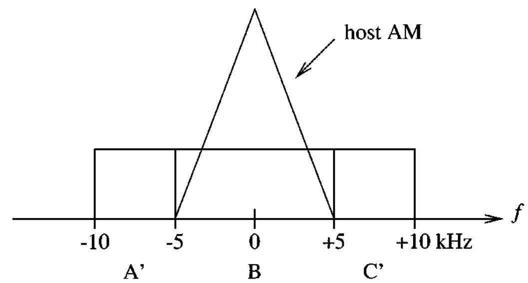

Fig. 26. Conceptual power spectrum for a 20-kHz hybrid IBOC-AM system. The digital data is transmitted in bands A′, B, and C′ with 1/3 of the total rate in each band.

The receiver for the multistream channel coding case produces three parallel error flags for the error mitigation in the embedded audio decoder, as indicated in Fig. 25. This could be done by means of a CRC or an outer RS code, as denoted in Fig. 25. When a high level of interference is detected in any of the three frequency bands, the system stops utilizing the corresponding source bitstream by blending to analog or using a lower decoding bit rate in the embedded audio coder. Note that, in the case where parts of $S1$ and $S2$ are transmitted in band B or parts of M are transmitted in bands A and C, the flagging of a single band may impact more than one layer of the description. It is therefore important to minimize such overlaps between the bands and layers and to have contingencies in cases where part of a layer is lost, e.g., possibly blending in only part of the analog signal or decoding part of a layer. More detailed description of the multistream transmission scheme can be found in [72].

B. 20-kHz Hybrid IBOC-AM System

As an alternative approach to the 30-kHz hybrid IBOC-AM schemes described above, one can make a case for narrowing the bandwidth to 20 kHz, as is sketched in Fig. 26.

There are two main advantages with the 20-kHz system over the 30-kHz system. First, there is no second adjacent hybrid IBOC to hybrid IBOC interference that can potentially be severe, as discussed in the previous section. Furthermore, there is a much greater compatibility with the all-digital IBOC-AM systems [98]–[100]. The potential drawback is a lower data rate due to lower bandwidth available to transmit the digital bitstream. Furthermore, the single-stream 20-kHz schemes lack graceful degradation capability. The blending can only be done directly from digital audio to the analog AM signal. In addition, the single-stream schemes cannot handle severe one-sided first adjacent interference. To combat this problem, the multistream transmission scheme described in the previous section for the 30-kHz system may also be applied [72]. However, since the bandwidth available is very limited for a 20-kHz system, the audio bit-rate allocation for different bands can become very challenging. A dual-stream transmission format for the 20-kHz system is described in [72].

C. FEC and Modulation Schemes

As discussed in the previous section, the power and bandwidth allocated for the transmission of a digital audio bitstream in a hybrid IBOC-AM system is very limited and is not enough to support conventional concatenated coding schemes such as using RS codes as outer codes and TCM [101], [102], [96] as inner codes. Thus, RS codes were first proposed to be used in these systems, and several modulation schemes based on RS codes are constructed in [103] to optimize the performance. The paper showed that, for a code rate of 4/5 and using 32-QAM constellations, an additional coding gain of about 2.5 dB can be achieved at a BER of 10^{-5} ($\sim$4 dB at BER $= 10^{-7}$) by using a multilevel RS coded QAM scheme instead of using the straightforward RS to QAM mapping scheme [103].

To further improve the performance, [103] also proposed using a concatenated coding scheme using an RS code as an outer code and a convolutional code (CC) as an inner code to the lowest levels of the multilevel RS coded QAM scheme, where errors are most likely to occur. The paper showed that, by applying concatenated codes to only the lowest level of the multilevel scheme, an additional 0.8 dB of coding gain can be obtained, compared to the multilevel RS coded QAM scheme [103].

The work in [104] explored the use of TCM instead of using the multilevel RS coded QAM schemes and showed that TCM outperforms the multilevel concatenated RS/CC by 0.8 dB at a BER of 10^{-4}–10^{-5}, but that the probability of error does not decrease as rapidly as that of the multilevel codes. For example, at a BER of less than 10^{-7}, the multilevel code performs better.

One of the advantages of using TCM instead of an RS code is that the underlying code of a TCM is a convolutional code. Since the convolutional code is proposed for the IBOC-FM systems, the IBOC-AM and IBOC-FM systems can potentially share the same hardware to implement the two FEC schemes. However, the underlying convolutional code for a TCM should be carefully chosen so that the TCM performance is optimized [104].

The potential drawback of using a TCM scheme is that a reliable error flag is not readily available, as is the case for RS codes. However, since symbol-by-symbol estimation is possible in TCM decoders by using the forward–backward decoding algorithm [105], a scheme using the resulting symbol-by-symbol soft decision to derive an error flag for error mitigation in audio decoders is proposed for digital audio radio services in [104]. Even though the error flag derived from TCM may not be as reliable as that of RS codes, it has the advantage of being able to match the flags with that of the audio frames as current state-of-the-art error mitigation algorithms perform error mitigation on an audio frame-by-frame basis. Since the audio coded frames are of variable lengths (e.g., approximately 500 to 3000 b), the error flags derived from the RS decoders are based on the RS frames and are often mismatched to the audio frame sizes. Thus, one RS frame error can result in flagging errors for multiple audio frames. On the other hand, flags derived from the TCM decoder described in [104] can be matched exactly to each audio frame since symbol-by-symbol soft decision can be derived at the TCM decoder. Thus, this enables more sophisticated error mitigation algorithms to be designed for audio decoders in order to jointly optimize the overall decoded audio quality.

Finally, to explore the power of using a turbo TCM scheme, the work in [104] also implemented a serially concatenated TCM scheme (SCTCM) and showed that one can approach the Shannon limit within 1–2 dB by using SCTCM schemes. However, the block length of the code needs to be large to achieve the goal. This is not critical for hybrid IBOC-AM applications since decoding delay can be on the order of a second. Therefore, turbo codes such as SCTCM can be a viable choice for these systems as they outperform both TCM and the multilevel schemes. Another possibility is using so-called *pragmatic coded modulation* schemes [106], which belong to a relatively simple family of bandwidth-efficient coded modulation schemes based on standard binary convolutional codes and,for example, multilevel QAM constellations. Finally, so-called bit-interleaved coded modulation [107], [108] is also possible where the rate of binary convolution code has been decoupled from the size of the QAM signal set, giving additional freedom to the system designer.

D. Discussion

We have discussed a number of options for the design of hybrid IBOC-AM systems. The final selection is going to be based on a complex set of tradeoffs including complexity versus performance. The hybrid IBOC-AM system selected for the U.S. market [62] is a multistream system with a bandwidth of 30 kHz and with two core audio streams in the outer frequency bands -15 to -10 kHz and $+10$ to $+15$ kHz, respectively. The enhancement stream is transmitted in the remaining tones at a lower power level. The audio coder is an embedded PAC coder with a 20-kb/s core and a 16-kb/s enhancement, i.e., a total rate of 36 kb/s. The all-digital IBOC-AM system will support PAC rates up to 60 kb/s [62].

It is also important to note that there are in-band, adjacent channel (IBAC) AM nonhybrid solutions proposed for frequencies below 30 MHz. One such system is developed by DRM. This system has its own set of requirements and objectives [8], [9]. The hybrid IBOC-AM and the all-digital IBOC-AM systems proposed for use in the U.S. are also part of the ITU world standard for digital broadcasting below 30 MHz.

VII. SDARS Services

The use of satellite systems for audio broadcasting seems to be a natural match. A satellite provides a large coverage to many receivers, and transmission delay (which is often a problem for communication applications) is not an issue in broadcasting. Nevertheless, most broadcast use of satellites has been limited to television services. In addition, these services mainly provide signals to stationary rather than mobile receivers [109].

The basic satellite broadcasting configuration (see Fig. 27) consists of a central programming and production facility, which transmits (uplinks) the broadcast signals to a satellite. The satellite takes this signal and shifts the up-link frequency to the proper downlink frequency, applies power amplification, and directs the signal to its designed footprint/service area. The signals are received by both stationary and mobile receivers within this area and are processed to retrieve the audio baseband signals (see [110]). To make uninterrupted broadcast reception possible, it is necessary to maintain a line of sight (LOS) with the satellite. Depending on the elevation angle between the service area and the satellite, this might be difficult to guarantee, especially for mobile receivers. To make these systems more reliable for mobile users, one option is to provide additional transmission diversity using terrestrial repeaters (or gap-fillers). Fortunately, the situation where this is needed the most (such as high-density urban areas with many high-rise buildings) coincides with important market areas and is therefore economically feasible. If the use of terrestrial repeaters is not an option or only can be used sparsely, another solution for countries or regions at high latitudes would be the use of elliptical (or geosynchronous) orbits. This option requires more satellites and a

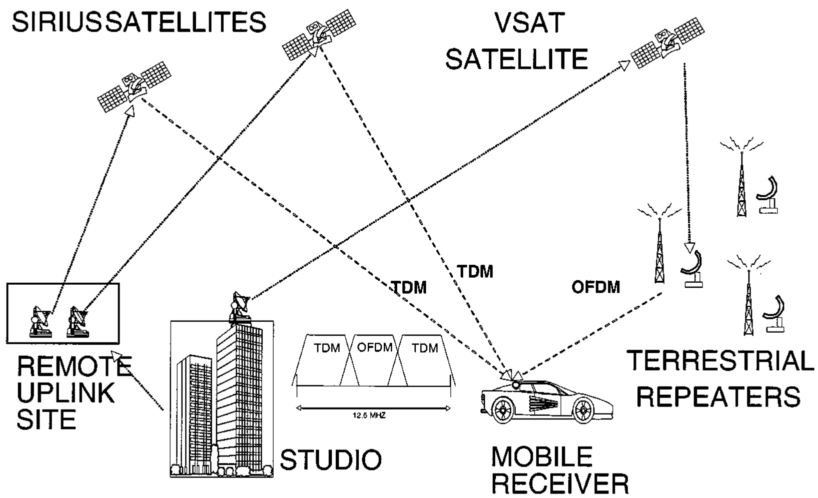

Fig. 27. Basic SDARS system. This configuration matches the Sirius Satellite Radio system.

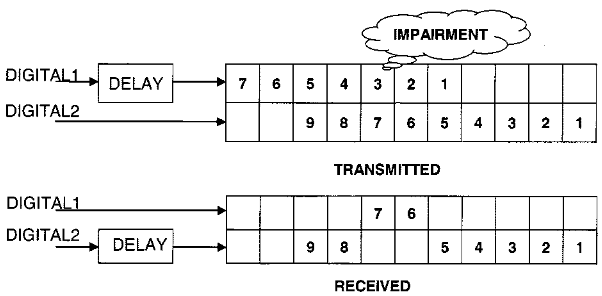

Fig. 28. Concept of time diversity. By delaying stream 1 during transmission, the receiver has two versions available. The contents will be impaired at different time instants.

switching scheme between the satellites to make sure that active transmissions are coming from the satellite having the highest elevation angle with respect to the service area.

To make these systems more robust, it is common to introduce time diversity. To do this, the same program is broadcast from different sources (e.g., two different satellites or one satellite and one terrestrial). One of the channels is delayed with respect to the other, for example by 4 s. Referring to the nondelayed channel as the early channel and the delayed channel as the late channel, at the receiver the early channel is delayed by the same amount to time align the two programs. Listening is done on the late channel and, if the transmission is interrupted by blockage, the listener is switched to the stored early channel. This is illustrated in Fig. 28.

The first SDARS system was Digital Satellite Radio (DSR) [111] and was operational from 1989 to 1999. A more recent system is the Astra Digital Radio (ADR) system [112], which provides digital radio services on its geostationary TV broadcasting satellites. The system covers central Europe and uses stationary receivers. It uses MPEG Layer 11 at 192 kb/s for stereo signals. FEC is based on a punctured convolution code with code rate-3/4 resulting in a 256-kb/s gross bit rate per channel. Transmissions are done in the 11-GHz range. Due to path losses at these frequencies, the antennas need dishes with diameters between 0.5 and 1.2 m.

Another more recent SDARS system is Worldspace [113], which provides digital radio services to developing countries using three geostationary satellites. This is a proprietary standard and limited information is available [114]. The system operates in the *L*-band (1452–1492 MHz) and consists of three geostationary satellites: 1) AfriStar (21°) covering Africa and the near and middle east; 2) AsiaStar (105°) covering China, India, Japan; and 3) AmeriStar (95°) covering central and South America. Each satellite has three spot beams. For each spot, two time division multiplex (TDM) streams are delivered. Each TDM stream contains 96 so-called prime rate channels (PRC), where each PRC transmits at 16 kb/s. For a typical high-quality stereo signal, the signal is transmitted at 128 kb/s using MPEG Layer III, thereby occupying 8 PRCs. At these rates, each spot can deliver $2 \times 12 = 24$ audio channels. The lower path losses for the *L*-band allow reception to be accomplished using low-gain helical antennas that maintain the LOS. Both AfriStar and AsiaStar are operational, and several manufacturers supply receivers [113]. AmeriStar is scheduled for launch beyond 2002. Services can be multimedia (audio, image, and moving image and data) and can be individually encrypted for pay-per-use. Most receivers are used in the stationary mode (home and portables). By using two broadcast channels per program and introducing time diversity, it is possible to make the reception more robust for mobile receivers, although most likely additional terrestrial repeaters are needed. At the writing of this paper, about 150 000 receivers have been deployed.

In the United States, no frequency allocation exists for SDARS in the L-Bband. In 1994, the FCC allocated the 2310–2360 MHz (*S*-band) for SDARS, consistent with the 1992 WARC allocation. In 1996, Congress mandated that the 2310–2320-MHz and 2345–2360-MHz portions of this band should be auctioned for wireless terrestrial communications services (WCS). The remaining bands where auctioned for SDARS services to CD Radio (now Sirius Satellite Radio, Inc.) based in New York, and American Mobile Radio Corporation (now XM Satellite Radio, Inc.) based in Washington, DC. The prices paid for the licenses were

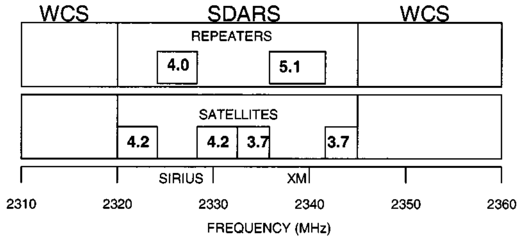

Fig. 29. Frequency–band allocation for SDARS and corresponding bandwidths for repeater and satellite broadcast bands.

$89 million and $93 million, respectively. The frequency allocations are illustrated in Fig. 29.

Both services are similar. For a monthly fee of about $10/month, they provide approximately 50 music channels and 50 talk radio channels. The quality and audio bandwidth of the music channels is somewhere between analog FM and CD quality (stereo), while the talk radio programs approach the quality of mono FM. Most of the programming is commercial-free and can be received anywhere in the continental U.S.. This wide coverage allows for the broadcast of programs that in local markets only have a small target audience but nationwide would reach a much larger audience. This diversity could be one of the attractions of SDARS. The nationwide coverage and commercial-free nature are other appealing factors. Both companies target the automotive market because of the "captive" audience and increasing commuting times for many people. Moreover, in the area between large metropolitan areas the availability of traditional analog broadcasting channels is usually limited and for long-distance commuters, SDARS is an attractive proposal.

Although the service models of both companies are similar, their system design is quite different. Each of them has a license for a 12.5-MHz band, which is roughly divided into three equal-sized bands. The middle band is used for the OFDM repeater signal while the two outer bands are allocated to the satellite signals. The main difference is that XM uses two geostationary satellites, where the average elevation angle will be 45° or less. Due to this low elevation angle, blockage by buildings and other tall obstacles is more likely and the availability of terrestrial repeaters is critical. The current design is based on the use of about 1000 repeaters, which significantly adds to its operation costs. The Sirius Satellite Radio system [10] is designed to limit the number of terrestrial repeaters by using three satellites in elliptical orbit. This geosynchronous orbit effectively makes the satellite follow a "figure-eight" pattern above and below the equator. Each satellite is located north of the equator for about 16 h per day, and at any given point in time two satellites are north of the equator. Since only two satellites can be active at the same time, this requires a hand-over procedure, which makes the overall system more complex. As a result, at any given point in time and at most locations, the minimum elevation angle is about 60°. Note that this angle is time-varying because the satellites are moving relative to the receiver. As a result, for a given stationary reception point, coverage and reception quality can vary as a function of time. The relative high (average) elevation angle makes the need for repeaters less critical, and the current design is based on the deployment of about 150 repeaters. For both systems, the power or equivalent isotropic radiated power (EIRP) of these repeaters can be relatively high (up to 40 kW), and companies using adjacent bands for WCS (see Fig. 29) have filed complaints with the FCC about limiting potential interference.

This different design philosophy about the systems has lead to descriptions of the Sirius system as a true SDARS system with terrestrial gap fillers while the XM system is a terrestrial system with satellite gap fillers. Although the license specifically prevents both operators from providing local programming using their gap fillers, the technical design of the system does not prevent such a service.

The terrestrial repeaters can be fed from the broadcast satellites. However, the interaction between the receiver antenna and the retransmitted signal requires very directive antennas on the base station that are aimed at the satellite. For the XM system, this is the solution of choice. For the Sirius system, this is a more difficult task to accomplish due to the time-varying positions, and their system uses commercial very small aperture terminal (VSAT) satellite services (geostationary) instead.

Reliable signal delivery to a mobile user poses many problems, and the systems have to be designed with a great deal of transmission diversity. Spatial diversity is obtained by transmitting the same information from the two visible satellites. Frequency diversity is provided by having the satellites transmit at frequencies as far as possible within the constraints of the license. By time-delaying the signal between each satellite, additional diversity is provided. The terrestrial signal is also time-delayed by the same amount. Since each stream contains all audio and control signals, only one of the streams needs to be properly received at the receiver. The decoded streams are combined using a maximal ratio combining technique that takes into account level and quality such that the best possible signal is recovered. Additional diversity could be obtained by applying embedded and multidescriptive source coding schemes (see Section III-E) [115].

The antenna built in the car must not only be small but should provide a form factor that can accommodate current car designs. Since its main beam has to see the satellite as the mobile platform turns and moves, the most economic solution is the use of a low-gain antenna with toroidal beam shapes. This shape allows constant gain in the azimuthal plane and directive gain in the elevation plane [110].

The satellite links use TDM transmitted using QPSK. This modulation technique is spectrum-efficient and allows the satellites to be driven at or near saturation. This is in contrast to OFDM, which requires the satellite power amplifier to be operated below saturation, thereby losing power efficiency. OFDM is used for the repeater modulation scheme due to its resistance to multipath fading. This is also the transmission scheme used in Eureka-147, and there is a wealth of experience and data available on its performance and efficient hardware implementations of transmitters and receivers.

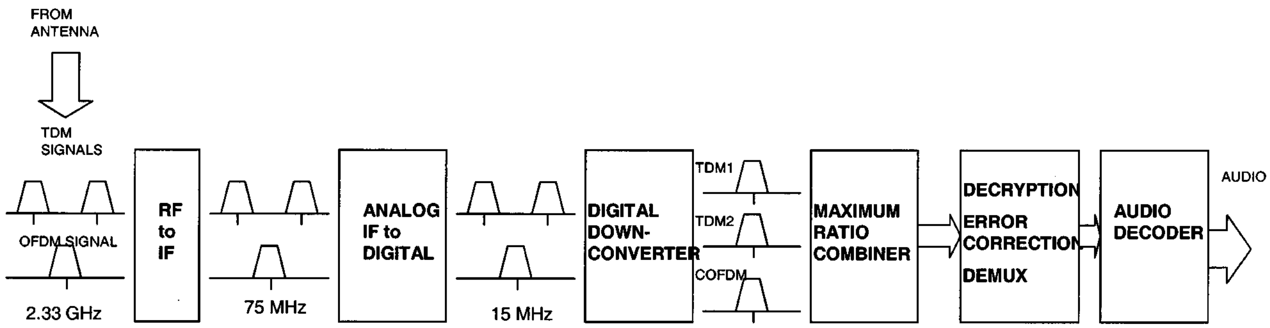

Fig. 30. Decomposition of SDARS receiver chip set.

For both systems, the delivery of a large number of channels with the highest possible quality is a must. The net payload is 4.4 Mb/s in the Sirius system and 4.0 Mb/s in the XM system. Hence, the use of source compression is essential, as mentioned in earlier sections. To accommodate various levels of programming material, both systems allow different allocations of source bit rates for different program materials. A course division is made between music channels, which are typically stereo and of high audio bandwidth (around 15 kHz), and talk radio, which is usually mono and has audio bandwidths as low as 6 kHz. Typical bit rates are between 48–64 kb/s for music services and 24–32 kb/s for talk radio services. Both systems allow adaptation of these allocations in a flexible manner. The XM system follows the approach used in Worldspace (both XM and Worldspace are associated with the same parent company, American Mobile Satellite), maintaining some of the features of the Worldspace system. This approach allocates bit rates in chunks to various programs. The chunk rate is either 8 or 16 kb/s. This requires the source coder to work at a fixed data rate. Since perceptual audio coders are inherently variable bit rate (see Section II-C), this requires special measures such as buffering to make the bit rate constant. If the buffers are large enough, this will have no impact on quality. However, in practical systems, the size of the buffers can negatively impact the quality of the audio. Moreover, a fixed allocation for a certain audio program can lead to either insufficient allocation and poor quality or overallocation leading to a waste of bits. The Sirius system allows for more flexibility by only assuming that the total aggregate rate has to be fixed, but that individual channels can have time-varying bit rates that have individually set averages or averages determined by the source material. This is accomplished by arranging programs in clusters of programs and jointly encoding the various program channels (see Section II-C).

Similar to the IBOC application, the source information has to be protected against channel impairments. Based on the link budgets and the need for error flags to allow error mitigation by the audio coder, a concatenated scheme consisting of RS combined with a convolutional code is used. The Sirius system uses a RS(255, 233, 8) code combined with a rate-2/3 convolutional code, resulting in a 39% coding overhead. A limited amount of interleaving is done but, although delay is in principle not an issue, in practice there is a constraint related to the so-called "tuning" delay (as mentioned in Section II-A),

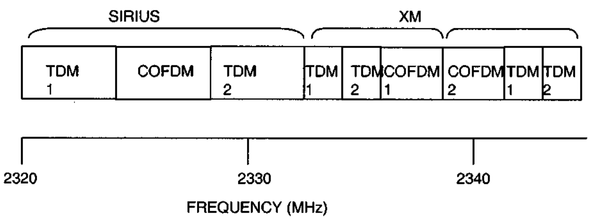

Fig. 31. Allocation of TDM and OFDM bands for XM and Sirius.

Table 12
Comparison Between XM and Sirius SDARS Systems

Parameter	Sirius	XM
No of Satellites	3	2
Satellite names	Sirius 1,2 and 3	Rock and Roll
Satellite type	Loral FS-1300	Boeing Satelite Systems 702
Longitude	nominal 100° W	85° and 115° W
Elevation angles	60°	45°
Uplink frequencies (MHz)	7060-7072.5	7050-7075
Satellite modulation	TDM-QPSK	TDM-QPSK
Repeater feed	VSAT Ku-band	Broadcast Sat (S-band)
Downlink frequencies (MHz)	2320.0 - 2324.0 2328.5 - 2332.5	2332.5 - 2336.5 2341.0 -2345.0
No of repeaters	150	1000
Repeater EIR Power kW	2 - 40	(80%) 2 -10 20% 10-40
Repeater modulation	TDM-OFDM	TDM-OFDM
Source coding scheme	sPAC + iPAC	iPAC + AACPlus
Channel coding	RS + Rate 2/3 Conv	RS + Rate 1/2 Conv
Net transmission rate	4.4 Mb/s	4.0 Mb/s
Number of music programs	60	70
Number of talk/news channels	40	30
Studio location	New York, NY	Washington, DC

which has to be less than 1 s. The signal gets modulated and delivered to the satellite and repeaters.

The receiver takes the signal from either source and demodulates it to the baseband. Fig. 30 shows the stages applied in the chipset developed by Agere Systems for Sirius. Note that a second generation of this set will only require two chips. XM has simplified their receiver architecture by splitting the receive bands into two parts. The diversity is accomplished by putting the TDM signals in two parts as well by means of two spot beams from each satellite. This is illustrated in Fig. 31. This requires an RF stage that only has to deal with half the 12.5 MHz, but tuning to program channels not covered in this band require band switching and will increase tuning delay. Table 12 summarizes the key differences between the XM and Sirius systems.

A. Control and Data Channels and Encryption

To allow for flexible program allocation, it is necessary to have a control channel. This control channel is also used for controlling access. Since the service is a subscription-based service, all information has been encrypted, and mechanisms are in place to control access.

Although the main purpose of SDARS is the delivery of audio content, the systems provide, in essence, a bit pipe to the mobile user. This allows for delivery of data services. Both XM and Sirius are looking into reselling some of their capacity for such services.

B. Current Status and Future Evolution

XM started their (limited-area) services in October 2001, while Sirius started (limited-area) commercial services in February 2002. At this time, both services cover the continental U.S.. The main market is mobile receivers in cars, and many car manufacturers have committed to 2003 models with built-in receivers (and antennas). Earlier models can be equipped with so-called after-market equipment, which can interface with existing radio by rebroadcasting a local FM signal or through an interface that fits in either the CD player or cassette player. It is expected that there will be a strong interest for these services. It is not clear, however, if the subscription-based model will be acceptable. XM has a business model based on a combination of channels with and without commercials, while Sirius has committed to only commercial-free content. Moreover, it remains to be seen how well the coverage is for each of these services. Interrupted services for music distribution are not well received by end users and could quickly reduce the excitement for SDARS. On the other end, the ability to receive a large selection of programs continuously throughout the United States has a very strong value proposition that will change the radio landscape forever.

The expectation is that it will help to accelerate the deployment of IBOC, and, obviously at some point in time, we will have car radios that will be able to receive analog, digital IBOC, and SDARS services.

VIII. Receiver Technology

The success of broadcasting depends on the availability of affordable receivers that can be used in many scenarios. Most digital radios require a significant amount of digital processing and typically require special-purpose VLSI. Economies of scale will push down prices to make radios affordable to anyone. Eureka-147 receivers have been widely available and have come down in price significantly. Satellite receivers have been available in the U.S. for about $300 and are expected to come down in price to about $150 for the low-end models. One potential hurdle for after-market satellite receivers is the need for an external antenna, which requires additional wiring and installation. New cars that are standard equipped with satellite radio most likely will become one of the driving forces for acceptance of this format.

The main distinction of digital radio services compared to analog services is the availability of data transmission capabilities. Although radio data services (RDS) has filled that gap, its bit rates are significantly lower than those possible with the digital services. Besides display of titles and other program information, traffic, weather, and financial information can be easily provided. In the U.S., where most radio services are commercially based, many additional services are to be expected. It should not be ruled out that certain IBOC or satellite channels will be subleased to third parties to provide value-added data services.

Further cost reduction is expected by providing radios that integrate the various broadcasting formats. However, since the standards are quite different in terms of frequency band, modulation schemes, and compression formats, it could be that the only way to provide multiple standards is by use of a programmable platform. From a service perspective, a consumer would likely subscribe to only one of the available satellite services. From a service point of view, integration with IBOC is more likely, since this will eventually replace the current freely available analog FM and AM services.

It is our belief that automotive will be the main application of the digital audio radio services described in this paper. It is expected, however, that a variety of stand-alone receivers will be available for either home use or portable (Walkman-type applications). For the portable applications, power management and antenna issues are the main challenges. For delivery to the home, many other services are already available such as cable TV audio distribution and Internet streaming, and penetration of the new receivers might only happen if they are bundled with other services.

IX. Discussion and Conclusion

The paper has presented some of the technical developments of digital audio broadcasting services within the world. An emphasis has been given to the deployment of digital audio radio services within the United States. These developments include advances in digital audio compression techniques, channel coding techniques, modulation techniques, and receiver technology. Even more importantly, the developments have relied on novel approaches that jointly consider the interactions between multiple elements in the communication chain, i.e., between the source signals, the channels, the receivers, and the human listener.

It is important to note that many of the techniques described may not be deployed in the final systems. Some techniques may see initial deployment only to be phased out later, and new techniques may be introduced as the services develop. The reasons for such dynamics in the design are both technical, e.g., as a result of hardware and performance considerations, as well as nontechnical, e.g., due to mergers and market pressures. The dynamics also reflect the flexibility inherent in the digital nature of the design, one of the many advantages that are driving the deployment of the digital systems. In fact, it is not unreasonable to expect newer nonaudio services to be deployed within the framework of these (primarily) audio systems in the near future.

By describing the range of techniques under consideration, the paper provides a historical perspective on the development process as well as details on the technical advancement. Specifics of final systems, still to be determined, are left to future publications.

ACKNOWLEDGMENT

The authors would like to mention N. Jayant who was one of the early drivers for research and prototyping of IBOC. They also would like to thank J. Johnston, D. Sinha, and F. Baumgarte for their work on the PAC audio coder. The authors acknowledge all other people who have contributed to the systems discussed in this paper but are not mentioned by name.

REFERENCES

[1] Federal Communication Commission. (2002, Aug.), Washington, DC. [Online]. Available: www.fcc.org.

[2] Arbitron., New York. [Online]. Available: www.arbitron.com.

[3] M. R. Aaron, "The digital (r)evolution," *IEEE Commun. Mag.*, pp. 21–22, Jan. 1979.

[4] J. D. Johnston, "Transform coding of audio signals using perceptual noise criteria," *IEEE J. Select. Areas Commun.*, vol. SAC-6, pp. 314–323, Feb. 1987.

[5] K. H. Brandenburg and G. Stoll, "ISO-MPEG-1 audio: A generic standard for coding of high-quality digital audio," *J. Audio Eng. Soc.*, pp. 780–792, Oct. 1994.

[6] W. Hoeg and T. Lauterbach, *Digital Audio Broadcasting Principles and Applications*. New York: Wiley, 2001.

[7] "Draft recommendation ITU-R BS[Doc 6/63] system for digital sound broadcasting bands below 30 MHz," Standard, ITU-R, Sept. 2001.

[8] Digital Radio Mondiale. (2002, June). [Online]. Available: www.drm.org.

[9] J. Stott, *Digital Radio Mondiale (DRM)—Key technical features*: Tech. Rep., Mar. 2001.

[10] Sirius Satellite Radio.. [Online]. Available: www.siriusradio.com.

[11] XM Satellite Radio., Washington, DC. [Online]. Available: www.xmradio.com.

[12] Global Radio., Luxembourg. [Online]. Available: www.global-radio.lu.

[13] WEB-Radio Directory, BRSMEDIA., San Francisco, CA. [Online]. Available: www.web-radio.fm.

[14] D. H. Layer, "Digital radio takes to the road," *IEEE Spectrum*, pp. 40–46, 2001.

[15] N. S. Jayant and P. Noll, *Digital Coding of Waveforms*. Englewood Cliffs, NJ: Prentice-Hall, 1984.

[16] C. E. Shannon, "A mathematical theory of communications I," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423, 1948.

[17] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*. Norwell, MA: Kluwer, 1992.

[18] J. P. Princen and A. B. Bradley, "Analysis and synthesis filter bank design based on time domain aliasing cancellation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 277–284, May 1986.

[19] H. S. Malvar, *Signal Processing With Lapped Transforms*. Norwood, MA: Artech House, 1992.

[20] W. B. Kleijn and K. K. Paliwal, Eds., *Speech Coding and Synthesis*. New York: Elsevier, 1995.

[21] "Recommendation G.722.1—Coding at 24 and 32 kbit/s for hands-free operation in systems with low frame loss," standard, ITU-T., Sept. 1999.

[22] S. Minami and O. Okada, "Stereophonic adpcm voice coding method," in *IEEE Int. Conf. Acoustics, Speech, Signal Processing*, 1990, pp. 1113–1115.

[23] S. A. Ramprashad, "Stereophonic CELP coding using cross channel prediction," in *IEEE Workshop on Speech Coding*, Delevan, WI, Sept. 2000, pp. 136–138.

[24] J. Benesty, D. R. Morgan, and M. M. Sondhi, "A better understanding and an improved solution to the specific problems of stereophonic acoustic echo cancellation," *IEEE Trans. Speech Audio Processing*, vol. 6, pp. 156–165, Mar. 1998.

[25] E. C. Cherry, "Some experiments on the recognition of speech, with one and with two ears," *J. Acoust. Soc. Amer*, vol. 25, pp. 975–979, 1953.

[26] B. Scharf, *Foundations of Modern Auditory Theory*. New York: Academic, 1970.

[27] J. D. Johnston, "Transform coding of audio signals using perceptual noise criteria," *IEEE J. Select. Areas Commun.*, vol. 6, pp. 314–323, Feb. 1988.

[28] J. D. Johnston and K. Brandenburg, "Wideband coding perceptual considerations for speech and music," in *Advances in Speech Signal Processing*, S. Furui and M. M. Sondhi, Eds. New York: Marcel Dekker, 1992.

[29] J. B. Allen, "Psychoacoustics," in *Wiley Encyclopedia of Electical and Electronics Engineering*, J. G. Webster, Ed. New York: Wiley, 1999.

[30] "Recommendation G723.1—Dual rate speech coder for multimedia communications transmitting at 5.3 and 6.3 kbit/s," standard, ITU-T, Mar. 1996.

[31] "Recommendation G.729—Coding of speech at 8 kbit/s using conjugate-structure algebraic-code-excited linear prediction (CS-ACELP)," Standard, ITU-T, Mar. 1996.

[32] E. Ekudden, R. Hagen, I. Johansson, and J. Svedberg, "The adaptive multi-rate speech coder," in *IEEE Workshop Speech Coding*, Porvoo, Finland, June 1999, pp. 117–119.

[33] T. Painter and A. Spanias, "Perceptual coding of digital audio," *Proc. IEEE*, vol. 88, pp. 451–513, Apr. 2000.

[34] *Final Draft International Standard FDIS 14 496-3: Coding of Audio-visual Objects, Part 3: Audio*, ISO/JTC1 SC29 WG11, Oct. 1998.

[35] M. Bosi, K. Brandenburg, S. Quackenbush, L. Fielder, K. Akagiri, H. Fuchs, M. Dietz, J. Herre, G. Davidson, and Y. Oikawa, "ISO/IEC MPEG-2 advanced audio coding," *J. Audio Eng. Soc.*, vol. 45, no. 10, pp. 789–814, 1997.

[36] J. D. Johnson, S. R. Quackenbush, J. Herre, and B. Grill, "Review of MPEG-4 general audio coding," in *Multimedia Systems, Standards, and Networks*, A. Puri and T. Chen, Eds. New York: Marcel Dekker, 2000, ch. 5.

[37] United States Television Systems Commitee (ATSC). (1995, Dec.) Digital Audio Compression Standard (AC-3). [Online]. Available: www.atsc.org/Standards/A52/

[38] D. Sinha, J. D. Johnston, S. Dorward, and S. Quackenbush, "The perceptual audio coder (PAC)," in *The Digital Signal Processing Handbook*, V. Madisetti and D. B. Williams, Eds. Boca Raton, FL: CRC/IEEE, 1997, ch. 42.

[39] D. Moulton and M. Moulton, "Measurement of small impairments of perceptual audio coders using a 3-facet rasch model," in *AES Conv.*, Preprint, May 1998.

[40] J. D. Johnson, D. Sinha, S. Dorward, and S. R. Quackenbush, "AT&T Perceptual Audio Coding (PAC)," in *Collected Papers on Digital Audio Bit-Rate Reduction*, N. Gilchrist and C. Grewin, Eds: AES, 1996, pp. 73–82.

[41] D. Sinha and C. E. W. Sundberg, "Methods for efficient multi-program audio broadcasting," in *AES Conv. Preprint*, Feb. 2000.

[42] J. H. Hall, "Auditory psychophysics for coding applications," in *The Digital Signal Processing Handbook*, V. Madisetti and D. B. Williams, Eds: CRC/IEEE, 1998, pp. 39-1–39-22.

[43] J. D. Johnston, "Perceptual coding of wideband stereo signals," in *ICASSP-89 Conf. Rec.*, 1989.

[44] K. Brandenburg and J. D. Johnston, "ASPEC: Adaptive spectral entropy coding of high quality music signals," in *AES 90th Conv.*, June 1991.

[45] B. Edler, "Coding of audio signals with lapped transforms and adaptive window functions" (in German), *Frequenz*, vol. 43, no. 9, pp. 252–256, 1989.

[46] Operating Manual Orban, "OPTIMOD-DAB 6200 Digital Audio Processor," Rep..

[47] J. D. Johnston, "Estimation of perceptual entropy using noise masking criteria," in *ICASSP-88 Conf. Rec.*, 1988.

[48] F.-R. Jean, C.-F. Long, T.-H. Hwang, and H.-C. Wang, "Two-stage bit allocation algorithm for stereo audio coder," *Proc. Inst. Elect. Eng.—Vis. Image Signal Process.*, vol. 143, pp. 331–336, Oct. 1996.

[49] C. Faller, "Audio coding using perceptually controlled bitstream buffering," in *Proc. 111th Conv. Aud. Eng. Soc.*, Sept. 2001.

[50] C. Faller, R. Haimi-Cohen, P. Kroon, and J. Rothweiler, "Perceptually based joint program coding for digital audio broadcasting," in *Proc. 113th Conv. Aud. Eng. Soc.*, to be published.

[51] R. V. Cox, W. B. Kleijn, and P. Kroon, "Robust CELP coders for noisy backgrounds and noisy channels," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, 1989, pp. 739–742.

[52] R. Arean, J. Kovačević, and V. K. Goyal, "Multiple description perceptual audio coding with correlated transforms," *IEEE Trans. Speech Audio Processing*, vol. 8, pp. 140–145, Mar. 2000.

[53] R. V. Cox, J. Hagenauer, N. Seshadri, and C.-E. W. Sundberg, "Sub-band speech coding and matched convolutional channel coding for mobile radio channels," *IEEE Trans. Acoust., Speech Signal Processing*, vol. 39, pp. 1717–1731, Aug. 1991.

[54] H. Taddei, S. A. Ramprashad, C.-E. W. Sundberg, and H-L. Lou, "Mode adaptive unequal error protection for transform predictive speech and audio coder," in *Proc. ICASSP*, Orlando, FL, 2002.
[55] D. Sinha and C.-E. W. Sundberg, "Unequal error protection (UEP) for perceptual audio coders," in *Proc. ICASSP*, Phoenix, AZ, Apr. 1999, pp. 2423–2426.
[56] N. Enbom and W. B. Kleijn, "Bandwidth expansion of speech based on vector quantization of the mel frequency cepstral coefficients," in *Proc. IEEE Workshop Speech Coding*, Porvoo, Finland, June 1999, pp. 171–173.
[57] C. Faller and F. Baumgarte, "Efficient representation of spatial audio using perceptual parametrization," in *Proc. IEEE Workshop Applications of Signal Processing to Audio and Acoustics*, Oct. 2001.
[58] ——, "Binaural cue coding: A novel and efficient representation of spatial audio," in *Proc. ICASSP*, Orlando, FL, May 2002.
[59] ——, "Binaural cue coding applied to stereo and multi-channel audio compression," in *Preprint 112th Conv. Aud. Eng. Soc.*, May 2002.
[60] F. Baumgarte and C. Faller, "Why binaural cue coding is better than intensity stereo," in *Preprint 112th Conv. Aud. Eng. Soc.*, May 2002.
[61] S. A. Ramprashad, "A multimode transform predictive coder (MTPC) for speech and audio," in *Proc. IEEE Workshop Speech Coding*, Porvoo, Finland, June 1999, pp. 10–12.
[62] IBIQUITY Digital Corporation., Warren, NJ. [Online]. Available: www.ibiquity.com.
[63] NAB National Radio Systems Committee. (2001), Washington, DC. [Online]. Available: www.nab.org/scitech/nrsc.asp.
[64] EUREKA147. (2002, Apr.) Digital Audio Broadcasting System. [Online]. Available: www.eurekadab.org.
[65] ETSI, "Radio broadcasting systems; Digital Audio Broadcasting (DAB) to mobile, portable and fixed receivers," Rep., ETS 300 401, May 1997.
[66] (2002, Aug.) WorldDab. [Online]. Available: www.worlddab.org.
[67] N. S. Jayant and E. Y. Chen, "Audio compression: Technology and applications," *AT&T Tech. J.*, vol. 74, no. 2, pp. 23–34, Mar./Apr. 1995.
[68] B. Chen and C.-E. W. Sundberg, "Complementary punctured pair convolutional codes for digital audio broadcasting," *IEEE Trans. Commun.*, vol. 48, pp. 1829–1839, Nov. 2000.
[69] B. Kroeger and D. Cammarata, "Complementary punctured convolutional codes with application to IBOC DAB," unpublished, June 1997.
[70] ——, "Robust modem and coding techniques for FM hybrid IBOC DAB," *IEEE Trans. Broadcast.*, vol. 43, pp. 412–420, Dec. 1997.
[71] C.-E. W. Sundberg, D. Sinha, D. Mansour, M. Zarrabizadeh, and J. N. Laneman, "Multistream hybrid in-band-on-channel systems for digital audio broadcasting in the FM band," *IEEE Trans. Broadcast.*, vol. 45, pp. 410–417, Dec. 1999.
[72] C.-E. W. Sundberg, D. Sinha, and H. Lou, "Multistream transmission for hybrid IBOC-AM with embedded/multidescriptive audio coding," *IEEE Trans. Broadcast.*, to be published.
[73] S. Lin and D. J. Costello, Jr, *Error Control Coding: Fundamentals and Applications*. Englewood Cliffs, NJ: Prentice-Hall, 1983.
[74] J. Hagenauer, "Rate compatible punctured convolutional codes (RCPC-codes) and their applications," *IEEE Trans. Commun.*, vol. 36, pp. 389–400, Apr. 1988.
[75] J. Hagenauer, N. Seshadri, and C.-E. W. Sundberg, "The performance of rate compatible punctured convolutional codes for digital mobile radio," *IEEE Trans. Commun.*, vol. 38, pp. 966–980, July 1990.
[76] T. Ottosson, "Coding, modulation and multiuser decoding for DS-CDMA systems," Ph.D. dissertation, Chalmers Univ. of Technol., Göteborg, Sweden, Nov. 1997.
[77] P. Frenger, P. Orten, T. Ottosson, and A. Svensson, "Multirate convolutional codes," Chalmers University of Technology, Dept. of Signals and Systems, Communication Systems Group, Göteborg, Sweden, Tech. Rep. 21, , Apr. 1998.
[78] R. D. Culver *et al.*, "EIA-CEG/DAR subcommittee, WG-B/laboratory testing and SG-1/VHF channel characterization Test. Final Report," EIA-DAR Channel Characterization Task Group, July 1995.
[79] J. A. C. Bingham, "In-band digital audio radio: An update on AT&T Amati PAC/DMT solution," in *Proc. Int. Symp. Digital Audio Broadcasting*, Toronto, ON, Canada, Mar. 1994, pp. 270–277.
[80] N. S. Jayant, E. Y. Chen, J. D. Johnston, S. R. Quackenbush, S. M. Dorward, K. Thompson, R. L. Cupo, J.-D. Wang, C.-E. W. Sundberg, and N. Seshadri, "The AT&T in-band adjacent channel system for digital audio broadcasting," in *Proc. Int. Symp. Digital Audio Broadcasting*, Toronto, ON, Canada, Mar. 1994, pp. 254–267.
[81] C.-E. W. Sundberg, "Digital audio broadcasting. An overview of recent activities in the US," in *Norwegian Signal Processing Symp. NORSIG'97 Conf. Rec.*, Tromsoe, Norway, May 1997, pp. 1–5.
[82] C.-E. W. Sundberg, "Digital audio broadcasting in the FM band," in *ISIE '97 Conf. Rec.*, July 1997, pp. SS37–SS41.
[83] H. C. Papadopoulos and C.-E. W. Sundberg, "Simultaneous broadcasting of analog FM and digital audio signals by means of adaptive precanceling techniques," *IEEE Trans. Commun.*, vol. 46, pp. 1233–1242, Sept. 1998.
[84] B. Chen and C.-E. W. Sundberg, "Digital audio broadcasting in the FM band by means of contiguous band insertion and precancelling techniques," *IEEE Trans. Commun.*, vol. 48, pp. 1634–1637, Oct. 2000.
[85] G. C. Clark, Jr. and J. B. Cain, *Error-Correction Coding for Digital Communications*. New York: Plenum, 1981.
[86] S. Kallel, "Complementary punctured convolutional (CPC) codes and their applications," *IEEE Trans. Commun.*, vol. 43, pp. 2005–2009, June 1995.
[87] B. Chen and C.-E. W. Sundberg, "List Viterbi algorithm for continuous transmission of concatenated cyclic redundancy check and convolutional codes," *IEEE Trans. Commun.*, vol. 49, pp. 784–792, May 2001.
[88] C. Nill and C.-E. W. Sundberg, "List and soft symbol output Viterbi algorithms: Extensions and comparisons," *IEEE Trans. Commun.*, vol. 43, Feb.–Apr. 1995.
[89] R. Maerkle and C.-E. W. Sundberg, "Turbo codes for digital audio broadcasting in the FM band," *IEEE Trans. Commun.*, submitted for publication.
[90] ——, "Channel codes based on hidden puncturing for partial band interference radio channels," *IEEE Trans. Commun.*, submitted for publication.
[91] J. N. Laneman and C.-E. W. Sundberg, "Soft selection combining for terrestrial digital audio broadcasting in the FM band," *IEEE Trans. Broadcast.*, vol. 47, pp. 103–114, June 2001.
[92] J. N. Laneman, C.-E. W. Sundberg, and C. Faller, "Huffman code based error screening and channel code optimization for error concealment in perceptual audio coding (PAC) algorithms," *IEEE Trans. Broadcast.*, to be published.
[93] H. C. Papadopoulos and C.-E. W. Sundberg, "Postcanceling techniques for simultaneous broadcasting of analog FM and digital data," *IEEE Trans. Commun.*, to be published.
[94] D. Goldston, "Modem proposals for IBOC-AM: Channel characterization," unpublished.
[95] C.-E. W. Sundberg, D. Sinha, and H. Lou, "Multi-stream transmission for spectrally and spatially varying interference AM-DAB channels," in *Proc. IEEE Communications Conf.*, Helsinki, Finland, June 2001, pp. 1654–1659.
[96] E. Biglieri, D. Divsalar, P. J. McLane, and M. K. Simon, *Introduction to Trellis-Coded Modulation With Applications*. New York: Macmillan, 1991.
[97] S.-Y. Chung and H. Lou, "Multilevel rs/convolutional concatenated coded QAM for hybrid IBOC-AM broadcasting," *IEEE Trans. Broadcast.*, vol. 46, pp. 49–59, Mar. 2000.
[98] R. Cupo and M. Shariat, "All-digital AM system architecture," unpublished, 1998.
[99] R. L. Cupo, M. Sarraf, M. Shariat, and M. Zarrabizadeh, "An OFDM all-digital in band-on-channel (IBOC) AM & FM radio solution using the PAC encoder," *IEEE Trans. Broadcast.*, vol. 44, pp. 22–27, Mar. 1998.
[100] B. Kroeger, "AM hybrid & all-digital IBOC DAB formats," unpublished.
[101] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Trans. Inform. Theory*, vol. IT-28, pp. 55–67, Jan. 1982.
[102] ——, "Trellis-coded modulation with redundant signal sets: Parts I and II," *IEEE Commun. Mag.*, vol. 25, Feb. 1987.
[103] S.-Y. Chung and H. Lou, "Multilevel RS/convolutional concatenated coded QAM for hybrid IBOC-AM broadcasting," in *Proc. IEEE Global Communications Conf.*, Nov. 1999.
[104] ——, "TCM and turbo TCM for audio broadcasting," in *Proc. IEEE Vehicular Technology Conf.*, May 2000.
[105] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 284–287, Mar. 1974.
[106] A. J. Viterbi, J. K. Wolf, E. Zehavi, and R. Padovani, "A pragmatic approach to trellis-coded modulation," *IEEE Commun. Mag.*, vol. 40, pp. 11–19, July 1989.

[107] E. Zehavi, "8-psk trellis codes for a Rayleigh channel," *IEEE Trans. Commun.*, vol. 40, pp. 873–884, May 1992.
[108] X. Li and J. A. Ritcey, "Bit-interleaved coded modulation with iterative decoding," *IEEE Commun. Lett.*, vol. 1, pp. 169–171, Nov. 1997.
[109] W. Pritchard and M. Ogata, "Satellite direct broadcast," *Proc. IEEE*, vol. 78, pp. 1116–1140, July 1990.
[110] R. D. Briskman, "Satellite DAB," *Int. J. Satellite Commun.*, vol. 13, pp. 259–266, 1995.
[111] W. Schambeck, "Digitaler hörfunk über satelliten," in *Fernseh-und Rundfunk Satellitenempfang*: Expert Verlag, 1987, pp. 145–177.
[112] Astra, Betzdorf, Luxembourg. [Online]. Available: www.ses- astra.com
[113] Worldspace, Washington, DC. [Online]. Available: www.worldspace.com
[114] S. J. Campanella, "Seminar on the worldspace satellite direct digital audio broadcast system," in *Proc. IEE Colloq. Communication*, July 19, 1998.
[115] H. Lou, M. J. Fernández-Getino García, and V. Weerackody, "FEC scheme for a TDM-OFDM based satellite radio broadcasting system," *IEEE Trans. Broadcast.*, pp. 60–67, 2000.

Christof Faller received the M.S. (Ing.) degree in electrical engineering from ETH Zurich, Switzerland, in 2000.

During his studies, he worked as an independent consultant for Swiss Federal Labs, applying neural networks to process parameter optimization of sputtering processes and spent one year at the Czech Technical University (CVUT), Prague. In 2000, he became a Consultant for the Speech and Acoustics Research Department, Bell Laboratories, Lucent Technologies. After one and a half year consulting, partially from Europe, he became a Member of Technical Staff, focusing on new techniques for audio compression applied to digital satellite radio broadcasting. Recently, he moved to the newly formed Media Signal Processing Research Department, Agere Systems, Murray Hill, NJ. His research interests include generic signal processing, specifically audio coding, control engineering, and neural networks.

Mr. Faller won the first price in the Swiss national ABB (Asea Brown Boveri) youth science contest organized in honor of the 100-year existence of ABB (formerly BBC) in 1991.

Biing-Hwang Juang (Fellow, IEEE) is Director of Acoustics & Speech Research (DAR) at Bell Labs, Lucent Technologies. He is engaged in a wide range of communication related research activities, from speech coding, speech recognition to multimedia and broadband communications. He recently led the DAR research team which developed the fundamental technologies for the upcoming digital audio broadcasting systems in North America. He has published extensively and holds a number of patents in the area of speech communication and communication services. He is co-author of the book *Fundamentals of Speech Recognition* (Englewood Cliffs, NJ: Prentice-Hall, 1993).

Dr. Juang received the 1993 Best Senior Paper Award, the 1994 Best Senior Paper Award, and the 1994 Best Signal Processing Magazine Paper Award, and was co-author of a paper granted the 1994 Best Junior Paper Award, all from the IEEE Signal Processing Society. In 1997, he won the Bell Labs' President Award for leading the Bell Labs Automatic Speech Recognition (BLASR) team. He also received the prestigious 1998 Signal Processing Society's Technical Achievement Award and was named the Society's 1999 Distinguished Lecturer. In 2000, he was awarded the IEEE Third Millennium Medal for his contributions to the field of speech processing and communications. He was endowed the Bell Labs Fellow Award, the highest honor in technical fields in Bell Laboratories. He is currently Editor-in-Chief of the IEEE TRANSACTIONS ON SPEECH AND AUDIO PROCESSING and member of the Editorial Board of the PROCEEDINGS OF THE IEEE. He also serves on international advisory boards outside the United States.

Peter Kroon (Fellow, IEEE) received the M.S. and Ph.D. degrees in electrical engineering from Delft University of Technology, Delft, The Netherlands. His Ph.D. work focused on time-domain techniques for toll-quality speech coding at rates below 16 kb/s. The regular-pulse excitation coding technique described in his thesis forms the basis of the current GSM cellular system.

In 1986, he joined Bell Laboratories, Murray Hill, NJ, where he has been working on a variety of speech coding applications, including the 4.8-kb/s secure voice standard FS1016, the ITU-T G.729 8 kb/s standard, and the IS-127 enhanced variable rate coder. From 1996 to 2000, he supervised a group developing new algorithms for audio/speech and channel coding applicable to digital cellular systems, voice mail systems, and both terrestrial and digital audio broadcasting. In 2001, he became director of the Media Signal Processing Research Department, Agere Systems, Murray Hill, NJ. He has published more than 35 papers and holds 9 U.S. patents,

Dr. Kroon received the 1989 IEEE Signal Processing Award for authors less than 30 years old for his paper on regular pulse coding. From 1997 to 2000, he served as an Associate Editor of the IEEE TRANSACTIONS ON SPEECH AND AUDIO PROCESSING. He was Chairman of the 1997 IEEE Speech Coding Workshop and a Co-Editor of a special issue on Speech Coding for the *International Journal of Speech Technology*, and, since 2000, he has served as a member-at-large on the IEEE Signal Processing Society Board of Governors.

Hui-Ling Lou (Member, IEEE) received the M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, in 1988 and 1992, respectively.

She focused on digital communications and signal processing algorithm and architecture design during her studies. Prior to her graduate studies, she designed digital circuit boards and computer-aided design tools for the 5ESS electronic switching system at AT&T Bell Laboratories. In 1992, she consulted at Amati Communications Corporation, currently a subsidiary of Texas Instruments, and developed a reconfigurable codec chip for an Asymmetric Digital Subscriber Line (ADSL) system. She was with the Multimedia and Wireless communications research laboratories at AT&T/Lucent Bell Laboratories from 1993 to 2001 where she contributed to the physical and link layer system, algorithm, and architecture design for both wired and wireless communication systems. In particular, she was involved in chip architecture design for the IS-95 CDMA cellular system, system and channel coding design for both the terrestrial and satellite audio broadcasting systems, and joint source and channel coding design for audio and video transmission over second- and third-generation mobile systems. Currently, she is a Senior Systems Architect at Marvell Semiconductor, Inc., Sunnyvale, CA. She has organized several conferences and special sessions for both wireless and multimedia communications. She has 21 patents, granted and pending, and has coauthored 37 technical publications.

Sean A. Ramprashad (Member, IEEE) was born in London, U.K., in 1968. He received the B.S.E. degree (*summa cum laude*) in electrical engineering from Princeton University, Princeton, NJ, in 1991 and M.S. and Ph.D. degrees in electrical engineering from Cornell University, Ithaca, NY, in 1993 and 1996, respectively.

He was an intern with AT&T Bell Laboratories during the years 1993 to 1995 and was a Member of Technical Staff at Lucent Technologies, Bell Laboratories, from 1996 to 2000. He is now a Distinguished Member of Technical Staff in the Communications System Technology Laboratory of Agere Systems, Murray Hill, NJ. His research interests have covered a range of areas including speech and audio coding, detection theory, video coding, and information theoretic aspects of vector quantization and feature extraction. He has been involved in product teams focused on a number of different applications, including video- teleconferencing, satellite audio broadcast, Bluetooth audio, and ITU-T coding standards.

Dr. Ramprashad is a member of the committees for the 1997 and 2002 IEEE Workshop(s) on Speech Coding, is presently serving as an Associate Editor for the IEEE TRANSACTIONS ON SPEECH AND AUDIO PROCESSING, and is a member of various subcommittees of the IEEE Speech Technical Committee. He is also a member of Phi Beta Kappa and Tau Beta Pi.

Carl-Erik W. Sundberg (Fellow, IEEE) received the M.S.E.E. and Dr.Techn. degrees from the University of Lund, Lund, Sweden, in 1966 and 1975, respectively.

He is a currently a Senior Scientist with iBiquity Digital Corporation,Warren, NJ. During 2000–2001 he was a Distinguished Member of Technical Staff at Agere Systems, Murray Hill, NJ, and during 1984–2000 he was a Distinguished Member of Technical Staff at Bell Laboratories, Lucent Technologies, Murray Hill, NJ. During 1976, he was with the European Space Research and Technology Centre (ESTEC), Noordwijk, The Netherlands. From 1977 to 1984, he was a Research Professor (Docent) in the Department of Telecommunication Theory, University of Lund. His consulting company, SundComm, has been involved in studies of error control methods and modulation techniques for many private companies, and international organizations. He has published over 90 journal papers and contributed over 140 conference papers. He has over 65 patents, granted and pending. He is coauthor of the books *Digital Phase Modulation* (New York: Plenum, 1986), *Topics in Coding Theory* (Berlin, Germany: Springer-Verlag, 1989), and *Source-Matched Digital Communications* (New York: IEEE Press, 1996). His research interests include source coding, channel coding, digital modulation methods, digital audio broadcasting systems, fault-tolerant systems, digital mobile radio systems, spread-spectrum systems, digital satellite communications systems, and optical communications.

Dr. Sundberg and his coauthor received the IEEE Vehicular Technology Society's Paper of the Year Award, and, in 1989, he and his coauthors were awarded the Marconi Premium *Proceedings of the Institution of Electrical Engineers* Best Paper Award.