Perceived Audio Quality Analysis in Digital Audio Broadcasting Plus System Based on PEAQ

Karel ULOVEC, Miroslav SMUTNY

Czech Technical University in Prague, Dept. of Radio Engineering, Technicka 2, 166 27, Prague, Czech Republic

{xulovec, smutnmir}@fel.cvut.cz

Submitted November 6, 2017 / Accepted February 1, 2018

Abstract. *Broadcasters need to decide on bitrates of the services in the multiplex transmitted via Digital Audio Broadcasting Plus system. The bitrate should be set as low as possible for maximal number of services, but with high quality, not lower than in conventional analog systems. In this paper, the objective method Perceptual Evaluation of Audio Quality is used to analyze the perceived audio quality for appropriate codecs — MP2 and AAC offering three profiles. The main aim is to determine dependencies on the type of signal — music and speech, the number of channels — stereo and mono, and the bitrate. Results indicate that only MP2 codec and AAC Low Complexity profile reach imperceptible quality loss. The MP2 codec needs higher bitrate than AAC Low Complexity profile for the same quality. For the both versions of AAC High-Efficiency profiles, the limit bitrates are determined above which less complex profiles outperform the more complex ones and higher bitrates above these limits are not worth using. It is shown that stereo music has worse quality than stereo speech generally, whereas for mono, the dependencies vary upon the codec/profile. Furthermore, numbers of services satisfying various quality criteria are presented.*

Keywords

Objective audio quality assessment, Digital Audio Broadcasting Plus (DAB+), Perceptual Evaluation of Audio Quality (PEAQ), MP2, AAC, number of services

1. Introduction

Radio broadcasting is in use for almost a hundred years. Recently, there is a worldwide transition from analog to digital radio systems. One of the systems for digital radio is Digital Audio Broadcasting (DAB) [1]. In many countries, there is experimental broadcasting, which aims to persuade listeners to switch to digital. However, the frequency spectrum is limited and is cluttered with many types of systems, which forces the broadcasters to include more services in DAB multiplexes. This results in a lower bitrate reserved for each service, which is one of the drawbacks in transition from analog to digital, because insufficient bitrate leads to inferior audio quality compared to analog systems [2].

On the other hand, the digital radio systems are evolving. They adopt new more efficient standards for source coding. Hence, the quality is higher or the number of services in one multiplex is increased. New error protection coding techniques are also adopted in modified specifications to further enhance robustness of the systems.

With audio quality in mind, the broadcasters or service providers have to decide on bitrate for each service according to its content. Logically, some services could be broadcasted with lower quality than other ones (anticipated number of listeners could be one of the criteria), but beside that, without proper quality assessment results, the broadcasters can only use their own ears. Detailed quality analysis of source coding methods used in digital radio systems is needed to help broadcasters with this decision, there must be an appropriate balance between the number of services and their audio quality in order to face up to expensive radio spectrum [2].

Although there were many papers examining quality in DAB and other systems of digital radio, such as [2–6], none of them, to authors' best knowledge, compared codecs or profiles of codecs, respectively, in sufficient detail regarding types of signal and number of channels. In this paper, these are investigated in detail. The main aim is to determine dependencies on the type of signal — music and speech, the number of channels — stereo and mono, and the bitrate. Furthermore, according to the results of the quality assessment analysis, the numbers of services in the multiplex are determined for various quality criteria.

The rest of this paper is organized as follows. Section 2 contains background information about DAB and the audio quality assessment and the summary of the related work. Section 3 presents method and results of the audio quality analysis in DAB. Section 4 concludes the paper.

2. Background

In this section, background information about topics discussed in this article is presented.

2.1 Digital Audio Broadcasting

The Digital Audio Broadcasting system is intended for broadcasting radio services. According to [1], the DAB system is one of four types of the current digital radio systems deployed around the world. The DAB system is convenient for fixed, portable, and mobile receivers and is capable of operation at frequency from 30 MHz to 3 GHz. A terrestrial form (T-DAB) is dedicated to portable and mobile reception especially. Audio content distribution can be certainly realized by cellular networks. However, streaming media applies a big load on a telecommunication infrastructure, and also appreciably decreases battery life of mobile devices [3].

Currently, digital audio broadcasting systems are replacing conventional analog radio systems around the world and popularity of these systems is increasing [4]. The audio quality of digital broadcasting systems is an important issue for the successful deployment of the new digital systems. Generally, new digital systems are designed to be robust to errors introduced by transmission thanks to high efficient channel coding (Forward Error Correction, FEC) and modern modulation format.

The DAB system is standardized by ETSI in [7], the first edition of specification was released in 1995. Generally, DAB uses industry standard encoding techniques for audio (and video) to remove redundancy and irrelevancy from source signals and error protection is applied during transmission. Multiplex of several digital audio services is carried via Radio-Frequency (RF) signal simultaneously in the bandwidth of about 1.5 MHz with total raw transport bitrate capacity of over 2.4 Mb/s. The range of the channel coding for error protection can be chosen for each service independently according to requirements of the broadcaster. The standard [7] allows several modes of FEC. The most used mode by the broadcasters is the option allowing protection levels labeled as 1-A, 2-A, 3-A and 4-A. The maximal available useful bitrate for broadcast services altogether (the useful capacity) ranges from 576 kbit/s to 1.728 Mb/s per one multiplex according to error protection.

Services can contain audio, video and data (related or not to audio service). In this paper, only audio services with embedded data in the bit stream are supposed. This data directly linked to the audio program are called Programme Associated Data (PAD). Maximum bitrate of PAD can be 78 kbit/s, but is usually in order of ones to tens of percent of the audio bitrate. Typically, multiplex carries from 10 to 20 audio services with bitrate from 24 kbit/s to 192 kbit/s. The bitrate of one audio service can be set according to [7] from 8 kbit/s to 384 kbit/s in multiples of 8 kbit/s. For audio source coding, DAB allows MPEG Audio Layer II (or MP2 for short) and Advanced Audio Coding (AAC).

MPEG-1 Audio Layer II is a subband audio codec defined in ISO/IEC 11172-3:1993 [8], specified for sampling rates 32 kHz, 44.1 kHz and 48 kHz. Later, it was extended for halved sampling rates and for more than two channels as MPEG-2 Audio Layer II defined in ISO/IEC 13818-3:1998 [9] as backwards compatible method of audio coding. MP2 uses a bank of Quadrature Mirror Filters to split audio to 32 bands. Each band is independently quantized in time domain based on masking threshold determined by the perceptual model. MP2 was originally used as a source audio coding for DAB, the methods of coding and the necessary mandatory requirements for decoding are specified in [10]. In DAB, only subset of available sampling rates is permitted — 48 kHz and 24 kHz. In MP2, according to [8], only several bitrates for 48kHz sampling rate are allowed, as shown in Tab. 1.

On the other hand, AAC is a transformational codec. It uses modified discrete cosine transform and quantization is performed in frequency domain. AAC was also defined in MPEG-2 as the Part 7 in ISO/IEC 13818-7:1997 [11] as a non-backwards compatible method with better coding efficiency than its predecessors. AAC-Low Complexity (LC) profile originated in [11]. Later, the AAC codec was updated and more profiles were added in MPEG-4 Part 3, or ISO/IEC 14496-3:1999 [12] and following amendments. The High-Efficiency (HE) AAC version 2 codec, containing three profiles, was added to DAB in 2007 leading to designation of DAB Plus (DAB+). AAC coding in DAB+ system is specified in [13]. Four sampling rate values are allowed — 48 kHz, 32 kHz, 24 kHz and 16 kHz. Three AAC profiles are possible — AAC-LC, HE AAC version 1 (HE-AACv1) and HE AAC version 2 (HE-AACv2). The name of the last profile is the same as of the codec itself, which is quite misleading. In this paper, therefore, the codec is called shortly AAC, whereas the individual profiles are mentioned in the full forms.

Thanks to a hierarchic structure, each codec capable of HE-AACv2 can also encode/decode in the both profiles lower in hierarchy. The first profile, lowest in hierarchy, is AAC-LC. AAC-LC, combined with Spectral Band Replication (SBR) forms more complex profile HE-AACv1 [14]. During encoding of audio signal using HE-AACv1 profile, high frequency content is analyzed and side information about this content is created. Downsampled audio is then passed to AAC-LC. Because the frequency band is narrower, the content at lower frequencies is more precisely described than it would be with original frequency range for the same bitrate. The high frequency content is reconstructed by replication of the lower frequency band with help of side information obtained during coding. The most complex HE-AACv2 profile (the highest in hierarchy) comprises the HE-AACv1 and Parametric Stereo (PS) [14]. The channel correlation is exploited by coding only one channel with the HE-AACv1. Spatial information is regenerated by the PS block using side information created during signal encoding.

Allowed bitrates [kbit/s]						
32	48	56	64	80	96	112
128	160	192	224	256	320	384

Tab. 1. List of allowed bitrates in MPEG-1 Audio Layer II for 48kHz sampling rate.

2.2 Audio Quality Assessment

In many aspects of multimedia applications, there is need to measure the quality of audio signals, i. e. how would an average listener evaluate the given audio. There are two groups of methods for audio quality measurement — subjective and objective.

Subjective methods use subjects, usually trained people, who are instructed to listen to processed (degraded) audio and then compare it with a reference. In order to get the average listener, it is needed to use statistically significant number of listeners, which can be very time consuming.

The Recommendation ITU-T P.800.1 [15] describes the methodology of a speech quality assessment in the context of the telecommunication channel. A Mean Opinion Score (MOS) is determined as the scale of quality by calculating mean of the 5-grade score given by the subjects. However, it is very important to eliminate unreliable subjects and so this method calls for large number of participants.

Next popular method is the ITU-R BS.1116 [16]. The trained subjects are presented with reference, impaired signal and hidden reference. They grade the impaired signal and the hidden reference with five grade scale. Subjective Difference Grade (SDG) is computed as the difference of grades, the value ranges between −4 (very annoying) and 0 (imperceptible). This process eliminates dependence on different quality demands of the subjects.

Another popular subjective method is called MUlti Stimulus test with Hidden Reference and Anchor (MUSHRA), or ITU-R BS.1534 [17]. It is used for intermediate quality impairments.

Although the subjective methods give more reliable results, for convenience, the objective methods are used. They are desirable, as they are less time consuming and cheaper than the subjective ones [18]. Objective methods use computer algorithms to predict perceived quality of the assessed signal. Algorithms can be simple, such as signal to noise ratio, but usually these simple algorithms do not correlate well with average listener assessment [19].

One of widely used objective methods is Perceptual Evaluation of Audio Quality (PEAQ). PEAQ is defined in Recommendation ITU-R BS.1387 [19]. It uses ear model to simulate perception and neural network for pattern recognition and mapping of individual characteristics to its output Objective Difference Grade (ODG). ODG is intended to be as close as possible to SDG from [16]. According to the recommendation ITU-R BS.1387 [19], if two ODG values differ in a tenth, the difference should not be taken in account and the quality of these two impaired signals is to be considered the same. For PEAQ, a basic and an advanced models were created. The advanced model is used for more thorough tests. PEAQ is intended for evaluating higher-quality audio, i. e. music and speech with only small impairment. The company Opticom, according to Holters and Zölzer [20], has created the only implementation of the algorithm fully compliant to [19], although there were numerous attempts by other individuals. According to the OPERA instruction manual [21], the ODG correlates well with subjective measurement down to −3.6. If the output value is lower, then it should not be used.

There are also algorithms, which were originally created for assessing speech quality, but were ported to also support music evaluation. One of them is Perceptual Objective Listening Quality Assessment (POLQA) and its adapted version is called POLQA Music. As proved in [4], the algorithm correlates better with subjective measurement than PEAQ, but it is still in development and is not publicly available.

Another speech evaluation method, which was adapted for assessing music at lower bitrates, is modified version of VIrtual Speech Quality Objective Listener (VISQOL) method — VISQOLAudio. PEAQ, POLQA and VISQOLAudio algorithms were compared with subjective tests (MUSHRA) on low bitrate audio in [22]. It was shown that VISQOLAudio performs quite well. It was also proved that although PEAQ was designed for measuring only small impairments, it also works for low bitrate codecs.

Several criteria for quality evaluation are used. One of them is perceptual “transparency”. It means that an average listener cannot distinguish between the reference and the encoded signal, i. e. the act of coding is perceptually transparent to the listener. The SDG (and also ODG) is 0 (imperceptible). Next, the EBU Tech 3339 [23] defines the quality that is safe to broadcast as having SDG score of −1 (perceptible, but not annoying). In this paper, this quality criterion is called “broadcast quality” for short.

2.3 Related Work

Some work has been carried out to investigate the perceived audio quality related to the DAB system.

Overview of the procedures to achieve maximum quality of service (QoS) in digital audio broadcasting systems was presented in [5]. The work was aimed to design a new radio platform capable of transmitting via several possible technologies (FM, AM, MW, Internet, DAB). From the perspective of QoS, a DAB network can be split to three key elements: the broadcasting station, the communication channel and the receiving sub-station. As concerned with QoS requirements, the latter two elements are fully specified. A minimum level of QoS can be guaranteed using an appropriate signal management system. [5]

The subjective audio quality of various digital radio systems was investigated in [6]. Radio equipment in the field was utilized for testing. The quality was evaluated as a function of the bitrate. Moreover, various systems were compared between each other. Various audio types of testing signals were employed (speech / music). Four values of bitrate in DAB and four different values of bitrate in DAB+

were tested, the broadcast quality was reached for 64 kbit/s and 96 kbit/s in DAB+ and for 192 kbit/s in DAB.

Related to perceived audio quality of FM and DAB+ radio systems, the results of two tests were presented in [2]. Firstly four values of audio bitrates in DAB+ were tested. Secondly two configurations of FM radio system and six values of audio bitrates in DAB+ were investigated. Profiles of the codec were chosen according to tested bitrate. The sampling frequency was 48 kHz and all signals were encoded in stereo. Four and five different signals were tested in the first and second experiment respectively. Subjective methods (ITU-R BS.1116 and MUSHRA) were applied. The broadcast quality was obtained for audio bitrate of 172 kbit/s in DAB (AAC-LC). One lower grade was assigned to DAB with audio bitrates of 113 kbit/s and 142.4 kbit/s (AAC-LC profile), and to DAB+ with audio bitrate of 86 kbit/s (HE-AACv1 profile). The bitrate for the transparency was estimated higher than 200 kbit/s for AAC codec.

In [4], the impact of the different audio codecs deployed in DAB and in the web-casting applications (among others MP2 and AAC) on the quality assessed by the end user was evaluated using both subjective and objective (PEAQ and POLQA) methods. Six critical signals were selected from public European Broadcast Union (EBU) database to perform the evaluation. Authors verified the quality for two sets of typical values of bitrate for six various codecs — for lower bitrates and for higher bitrates. It was proved that the codec seriously impacts perceived audio quality in case of lower bitrates. Negligible differences between tested codecs were concluded in case of higher bitrates. But, the values of bitrate for different codecs were not the same in both sets, so that the codecs cannot be easily compared between each other. In the paper, relation between SDG and ODG was evaluated. For higher bitrates, PEAQ gave worse results (ODG $\approx -1.5 \cong$ SDG ≈ -1) and there were negligible differences between results for various codecs (in the order of tenths). For lower bitrates, in average, the difference from SDG was greater, (ODG $\approx -4 \cong$ SDG ≈ -2.4). However, the dependence on codecs was significant; SDG values were 0.2 and 0.6 lower than average for AAC-LC and MP2, respectively, whereas the values of SDG were one grade higher for the HE-AACv2 profile. It must be noted that the PEAQ model was not trained for impairments introduced by the HE-AAC codec profiles resulting to worse quality evaluation for HE-AAC.

DAB+ and analog FM radio systems were compared in [3]. Authors analyzed whether the quality of DAB+ surpasses the quality of FM radio. A subjective quality assessment (using MOS scale) and objective quality assessment (ViSQOLAudio metric) of the AAC codec were presented. Seventeen audio signals of various types were prepared for testing. Signals were sampled using 48kHz frequency and encoded with four different bitrates. Authors noted a need for a higher bitrate for the signal with a clear stereo separation between the channels. DAB+ was concluded as an efficient replacement for a traditional FM broadcasting system.

3. Audio Quality Analysis in DAB

3.1 Method

Because subjective tests demand statistically significant number of participants and it was needed to perform large number of repeated tests (in a wide range of bitrates with fine step resolution), the objective measurement was performed using PEAQ. It was desired to compare mono and stereo versions of the signals and also determine the dependence on the type of signal, which multiplied the number of needed tests. Other papers have already proved good correlation of PEAQ's ODG to SDG [4]. Moreover, objective assessment leads to results that are comparable and repeatable.

A set of audio samples was prepared (see Tab. 2). It is a mixture of six speech and music samples from the original test set, which comes with the PEAQ recommendation [19] (the ones whose filename starts with "ref") and newly created

Filename	Content	Description	Signal type
boulevard.wav	Green Day - Boulevard of Broken Dreams	punk music	music
capriccio.wav	P. I. Tchaikovsky - Capriccio Italien	classical music	music
dubstep.wav	DJ Chimera	electronic music	music
pennylane.wav	The Beatles - Penny Lane	pop music	music
refglo.wav	glockenspiel melody	transients and high bandwidth	music
refharp.waw	harpsichord	transients and complex sound	music
refsop.wav	W. A. Mozart - Der Hölle Rache	complex sound, classical music	music
rickroll.wav	Rick Astley - Never Gonna Give You Up	pop music	music
walking.wav	Fats Domino - I'm Walking	rock'n'roll music	music
holmes.waw	A. C. Doyle - Vampire in Sussex	male Czech speech	speech
refsfe.wav	"To administer medicine to animals. . . "	female English speech	speech
refsme.wav	W. Irwing - Old Christmas	male English speech	speech
refsmg.wav	"Die Natur hat den Menschen eine Zunge. . . "	male German speech	speech

Tab. 2. Set of audio samples.

speech sample and music samples covering different genres. The samples were chosen to represent typical material that may appear on radio. Of course, none of the used samples were affected by lossy compression, as the generation loss would worsen the quality. Each sample was trimmed to around 20 to 30 seconds so that averaging the ODG by the algorithm does not affect the result too much. Although it would be possible to use shorter samples (for example half a second), it was intended to use samples of sufficient length, which would be used in subjective tests.

Although the DAB system can use various sampling rates, only 48kHz sampling rate was researched in this paper. Logically, lowering the sampling rate leads to lower bitrate needed to compress the signal at given quality. The quality degradation of the signal at lower sampling rate at given bitrate is lower than when using the sampling rate of 48 kHz. Therefore, the results might serve as the worst case scenario. When using lower sampling rate, the quality degradation should be the same or lower than presented in this paper.

Each sample was prepared in two versions, 2-channel (stereo) and 1-channel (mono). The 1-channel version was made by summing the left and right channel and multiplying by factor of 0.5. It was intended to prepare all samples with bitrates in a range between 8 and 384 kbit/s with 8kbit/s step, as possible in DAB. But as Table 1 shows, only several sample rates are possible for MP2. Also, the profiles higher in hierarchy of the AAC codec do not work past defined upper threshold bitrate (as seen in Results), so higher bitrate could not be set. Moreover, as the HE-AACv2 profile is intended for stereo, it naturally does not work for mono samples, so only stereo samples were coded with it. A total of twenty six (thirteen samples times two versions) signals were encoded with MP2 codec and all three profiles (where available) of the AAC codec in a range of relevant bitrates.

All coded signals were then decoded and assessed using the implementation of PEAQ in OPERA software (OPERA V3.5). Because the codecs slightly desynchronize the signal, it was needed to synchronize the degraded and the reference signal with each other before assessing. The advanced model of the PEAQ algorithm was used. Because the AAC codec drastically changes stereo image of the audio, and the penalty of the PEAQ algorithm for this situation is set too high, quality of these was evaluated using mono version of the decoded sample. For result discussion, ODG was averaged in each signal group (speech and music). For extreme low bitrates, some ODG values were lower than −3.6 (AAC-LC under 24 kbit/s, typically), therefore they were not used, because these do not correlate with subjective tests [21].

3.2 Results

The results of the tests and further analysis are divided into three parts. Firstly, the dependencies of ODG on bitrate, the signal type and the channels number are presented. Secondly, the limit bitrates to outperform the quality of the higher AAC profiles in hierarchy by the lower ones are determined. Thirdly, the minimal bitrates to reach several quality criteria and the corresponding numbers of services are evaluated.

ODG dependencies

Figures 1 to 8 present obtained dependencies of ODG on bitrate. In Figs. 1 to 4, the results for stereo and mono music and stereo and mono speech, respectively, can be seen. Figures 5 to 8 show ODG of all signal types for given codec. The first quaternion of figures allow codec performance comparison for each sample type, while the latter are aimed to compare performance of each codec/profile with respect to sample type and number of channels.

Figure 1 shows that until 96 kbit/s, the ODG of the MP2 codec stays around −3.5, whereas at the same bitrate, the AAC-LC and HE-AACv1 reach broadcast quality (ODG over -1). The MP2 codec reaches (almost) imperceptible quality loss at 192 kbit/s. Both High Efficiency profiles perform better than the AAC-LC profile for low bitrates. At 56 kbit/s, the average ODG for stereo music of the AAC-LC profile reaches −3, while HE-AACv1 gets to around −2.5 and HE-AACv2 even to −2.

For mono music (Fig. 2) it may seem that the needed bitrates to reach certain quality are halved bitrates of stereo music. It is almost completely true for MP2 codec. But a closer look at the profiles of the AAC codec (Figures 6 and 7 show it better) reveals that the quality at half of certain bitrate for mono music is slightly worse than the quality of stereo music at that bitrate. It is because of the fact, that in stereo, the codec may exploit channel correlation and use joint stereo mode, for example.

Stereo speech (Fig. 3) gives better results than stereo music for all codecs. The MP2 codec needs much greater bitrate than the AAC codec, it is also almost unusable up to 96 kbit/s. 192 kbit/s seems as an optimum bitrate for broadcast quality. As the figure shows, there is almost no improvement in HE-AACv1 over AAC-LC for stereo speech coding. Over 64 kbit/s, its quality is even lower than the quality of AAC-LC. The HE-AACv2 profile brings greater quality than other profiles for lower bitrates up to 56 kbit/s.

Figure 4, where results for mono speech can be seen, points out again, that the MP2 codec needs much higher bitrate to achieve the same quality as the AAC codec. For MP2, the quality of mono signal is almost the same as the quality of stereo signal at double bitrate (the difference between the ODGs is only in order of a tenth), it can be seen well in Fig. 5. For the AAC family it is not completely true. For AAC-LC, the quality of monaural signal at half of the bitrate is again lower than the quality of stereo signal.

In Fig. 6, mono speech and music gave similar results for AAC-LC profile. Interestingly, for HE-AACv1 profile (Fig. 7), although stereo music gives worse results than stereo speech, mono versions of these signals swapped their position — mono speech gives worse results than mono music.

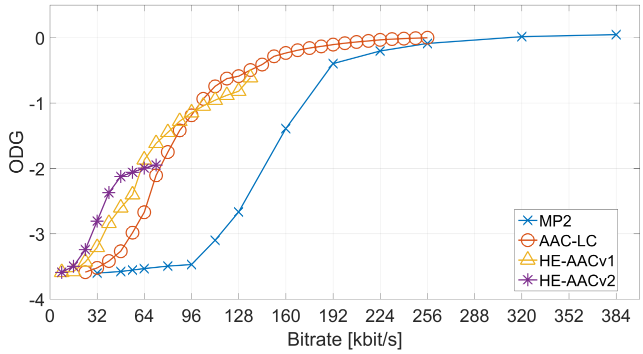

Fig. 1. Results for stereo music, both codecs.

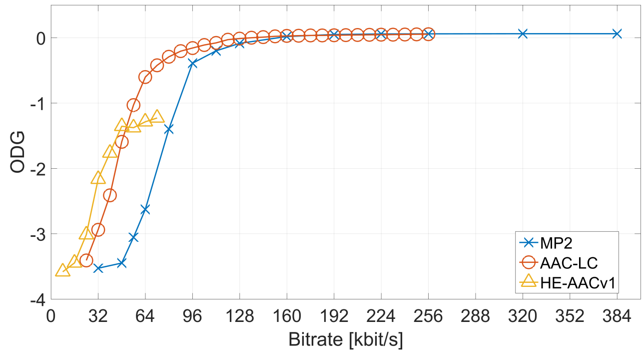

Fig. 2. Results for mono music, both codecs.

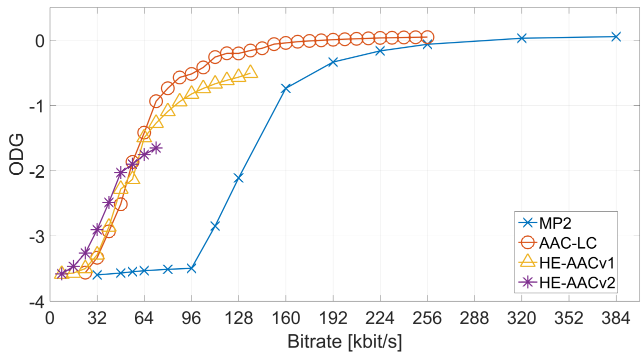

Fig. 3. Results for stereo speech, both codecs.

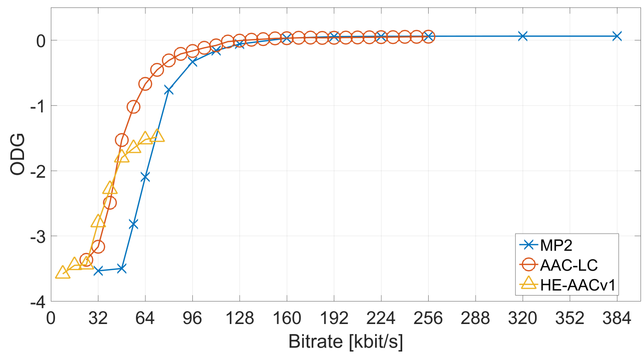

Fig. 4. Results for mono speech, both codecs.

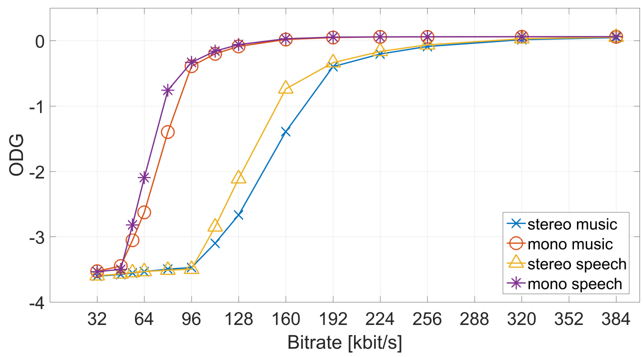

Fig. 5. Results for the MP2 codec.

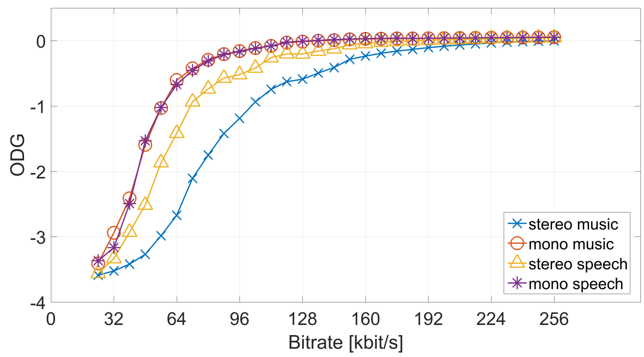

Fig. 6. Results for the AAC-LC profile of the AAC codec.

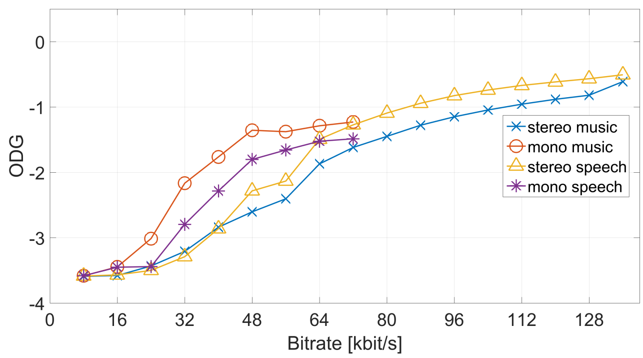

Fig. 7. Results for the HE-AACv1 profile of the AAC codec.

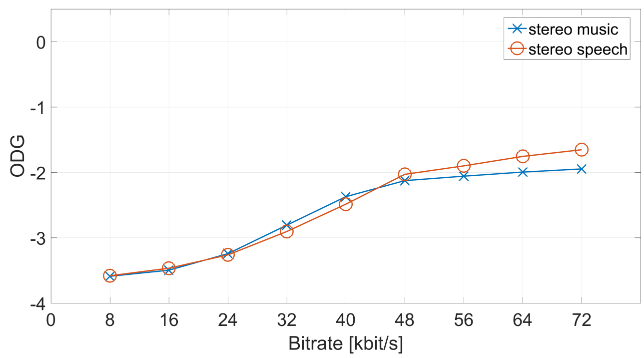

Fig. 8. Results for the HE-AACv2 profile of the AAC codec.

However, as Figure 3 shows, there is almost no improvement in HE-AACv1 over AAC-LC for stereo speech coding, because the difference in quality is only about a tenth or two until the profile is outperformed by AAC-LC after 48 kbit/s. For HE-AACv2 (Fig. 8), there is not clear difference in quality between types of signals up to 24 kbit/s. From 24kbit/s to 48kbit/s, the music gently outperforms the speech and from 48 kbit/s onwards, speech reaches slightly better results again, but now with greater difference.

As seen in the figures, the MP2 codec needs higher (double, at many times) bitrate than the AAC-LC for given quality. Because the MP2 codec does not offer enough possible bitrate settings, the bitrate to reach ODG ≈ -0.5 and broadcast quality is the same. Next, the more complex profiles do not reach transparency (HE-AACv2 does not even reach broadcast quality), because the signal is considerably altered by the spectral band replication and parametric stereo.

It should be noted that although PEAQ's ODG was intended to be as close as possible to SDG, results of subjective and objective methods comparison in [4] show that these values do not exactly match each other, especially for lower quality. Unfortunately, there is also a significant dependence on codecs. Results for AAC-LC profile are more or less correct. However, PEAQ gives better grades to MP2 codec and significantly penalizes HE-AACv2 profile (HE-AACv1 profile was not included in the paper [4]). It should, again, be mentioned that the PEAQ model was not trained for HE-AAC profiles, more detailed information is in Sec. 2.3.

Limit bitrates to outperform quality

AAC profiles higher in hierarchy (more complex ones) shows improvement in quality over profiles lower in hierarchy (less complex ones), but only at lower bitrates. As the bitrate gets higher, the quality saturates and the more complex profile is outperformed by the less complex profile. The codec creators reflected this situation by setting the upper bound of the available bitrate range at around 70 kbit/s for HE-AACv2 and 140 kbit/s for HE-AACv1 in stereo and 70 kbit/s for HE-AACv1 in mono. Higher value can be set, but the codec does not obey the setting. But, the upper bounds of available bitrates do not exactly match the bitrates where the profiles are outperformed. So it is not worth using the profile with higher bitrates, although it is possible. The quality curve of AAC-LC crosses the quality curve of HE-AACv1 at 96 kbit/s for stereo music or even already around 64 kbit/s for stereo speech. HE-AACv2 is outperformed by HE-AACv1 at around 64 kbit/s for both speech and music. This situation also occurs for mono signals, where the HE-AACv1 is outperformed by AAC-LC at around 48 kbit/s. However, there is a slight difference between signals, the crossing for speech signals happens at lower bitrate than for music (when considering interpolation of the results).

Stereo music gives worse results than stereo speech, because the channel differences in speech signals are smaller than in music signals. On the other hand, each codec treats mono signals differently — MP2 keeps the order, i. e. mono music has worse quality than mono speech; AAC-LC profile of the AAC codec gives very similar results for music and speech and using the HE-AACv1 profile results in mono speech having worse quality than mono music. It is probably caused by the spectral band replication, which distorts the speech more than the music. Thus, for mono speech, HE-AACv1 does not reach broadcast quality unlike for mono music.

Bitrates to reach quality criteria, the numbers of services

The next part is to determine the numbers of services which can be put in the multiplex. First, it is needed to find bitrates to reach several quality criteria including transparency and broadcast quality. The bitrates are shown in Tab. 3. Because it is not probable, according to the table, that the broadcasters would set such high bitrates needed for transparency, ODG ≈ -0.5 is added, as a reasonably high quality. Also, ODG ≈ -2 (slightly annoying) and ODG ≈ -3 (annoying) are used as criteria. The first one might serve for lower quality broadcasting, the second one as a border for "unusable quality". When deciding which bitrate fulfills the criterion (especially for MP2, due to the sparsity of the bitrates available), the insignificance of the difference of ODG of a tenth of a grade is taken in account. If the difference of the quality from the criterion is greater than 0.2, higher bitrate in the list is taken. If the quality cannot be reached (higher bitrate is not available), the cell is denoted by 'x'.

Secondly, the number of services is estimated. A 5% increase of bitrate in each service is assumed for additional information (PAD). More 10% increase is supposed as a space for an additional FEC in case of AAC codec. The numbers in Tab. 4 for 4-A protection level and Tabs. 6 to 8 for protection levels 1-A to 3-A (in the Appendix) are calculated using the equation

$$N_{\mathrm{n,i,j,k}} = \left\lfloor \frac{C_{\mathrm{n}}}{B_{\mathrm{i,j,k}} \cdot (1 + M_{\mathrm{i}})} \right\rfloor \tag{1}$$

where $N_{\mathrm{n,i,j,k}}$ is the number of services, C_{n} is the multiplex capacity for n-th protection level, $B_{\mathrm{i,j,k}}$ is bitrate of i-th codec/profile for j-th signal type for k number of channels, M_{i} is the margin for PAD and optionally for an aditional FEC for i-th codec/profile and $\lfloor\cdot\rfloor$ denote the floor function. The margin M_{i} equals 0.05 for MP2 and 0.15 for all three profiles of the AAC codec.

The numbers in Tab. 4 are discussed in detail below. It can be seen that for transparency, the broadcasters or service providers can put only 5 stereo services or 10 mono services with MP2. With AAC-LC, the numbers are the same for music, for speech, they are increased to 7 stereo services and 11 mono. Moreover, the services can be more robust thanks to the optional FEC. When broadcasting reasonably high-quality audio (ODG over -0.5), there can be 8 services when using MP2 in stereo, or 17 in mono. For AAC-LC, the number of services for music increases to 11 and 20 or to 15 and 20 services with high quality speech for stereo and mono signals, respectively. For HE-AACv1, the number of chan-

	Quality criterion	Bitrates [kbit/s]						
		MP2		AAC-LC		HE-AACv1		HE-AACv2
		stereo	mono	stereo	mono	stereo	mono	stereo
music	Transparency	320	160	256	144	x	x	x
	ODG ≈ −0.5	192	96	136	72	136	x	x
	Broadcast quality	192	96	104	56	104	72	x
	ODG ≈ −2 (slightly annoying)	160	80	72	48	64	32	64
	ODG ≈ −3 (annoying)	112	56	56	32	40	24	32
speech	Transparency	320	160	192	136	x	x	x
	ODG ≈ −0.5	192	96	96	72	136	x	x
	Broadcast quality	160	80	72	56	88	x	x
	ODG ≈ −2 (slightly annoying)	128	64	56	48	56	48	48
	ODG ≈ −3 (annoying)	112	56	40	32	40	32	32

Tab. 3. Bitrates to reach quality criteria.

	Quality criterion	Number of services						
		MP2		AAC-LC		HE-AACv1		HE-AACv2
		stereo	mono	stereo	mono	stereo	mono	stereo
music	Transparency	5	10	5	10	x	x	x
	ODG ≈ −0.5	8	17	11	20	11	x	x
	Broadcast quality	8	17	14	26	14	20	x
	ODG ≈ −2 (slightly annoying)	10	20	20	31	23	46	23
	ODG ≈ −3 (annoying)	14	29	26	46	37	62	46
speech	Transparency	5	10	7	11	x	x	x
	ODG ≈ −0.5	8	17	15	20	11	x	x
	Broadcast quality	10	17	20	26	17	x	x
	ODG ≈ −2 (slightly annoying)	12	25	26	31	26	31	31
	ODG ≈ −3 (annoying)	14	29	37	46	37	46	46

Tab. 4. Number of services to reach quality criteria for 4-A protection level.

nels is smaller to fulfill this quality criterion in case of speech signal. Therefore, it is better to use AAC-LC instead.

For broadcast quality, the numbers further increase (except for MP2, because the bitrates were not dense enough). There can be 14 stereo music services with AAC-LC or 20 stereo speech services. In mono, numbers of services increase to 26, numbers are same for both music and speech. The HE-AACv1 profile is of no use, because the number of services to reach broadcast quality is the same (stereo music) or lower than for AAC-LC. For mono speech, the HE-AACv1 profile does not even reach broadcast quality.

If the broadcaster makes do with lower quality, still useable ODG −2 yields to 10 stereo music and 12 stereo speech services encoded with MP2, or 20 and 26 services encoded with AAC-LC. The high efficiency profiles start to outperform the AAC-LC at this quality criterion. For stereo music, the numbers of services for both HE-AACv1 and HE-AACv2 are the same, 23, 3 more than for AAC-LC, whereas for speech, AAC-LC and HE-AACv1 offer the same number of services for both stereo and mono, 26 and 31, respectively. HE-AACv2 offers stereo speech with the same number of services as HE-AACv1 for mono, 31.

It should be noted that the number of services for ODG ≈ −3 is only a limit and it is advisable that the broadcasters do not reach these numbers of services when using any of the codecs/profiles as the broadcasted signals turn out to be unusable. So, there should be less than 14 stereo or 29 mono services encoded with MP2 codec or less than 26 stereo music or 37 stereo speech or 46 mono services encoded with AAC-LC. With HE-AACv1, the number of services should not exceed 37 for stereo or 62 for mono music and 46 for mono speech. The limit for both types of signals for HE-AACv2 is 46.

To sum up, mono signals are of worse quality than stereo signals at double bitrate except for MP2, where the quality is the same. AAC-LC needs much lower bitrate than MP2, but some of the bandwidth is reserved for optional FEC, so the number of services is not strictly proportional to bitrate ratios. For example, broadcast quality for stereo signals is reached with 192 kbit/s for music and 160 for speech for MP2, while AAC-LC needs only 104 kbit/s for music and even only 72 kbit/s for speech. With this quality, there can be 8 services with music or 10 services with speech. With AAC-LC, the number of services increases to 14 and 20 for music and speech, respectively, together with additional error protection.

The signal type also affects the quality of the encoded signal. With MP2 and AAC-LC, stereo speech is more efficiently encoded than music. However, the type of signal influence the High-Efficiency profiles differently, because the quality is almost the same up to 48 kbit/s. In mono, the dependence on the type of signal is various for each codec (Tab. 5). In MP2, the difference between types of signal is the same as in stereo, just at half the bitrate. AAC-LC gave similar quality for both speech and music and with HE-AACv1, speech has worse quality than music.

Codec/Profile	Mono signals quality
MP2	speech better (same order as in stereo)
AAC-LC	same quality
HE-AACv1	music better (reversed order)

Tab. 5. Behavior of codecs/profiles to types of mono signals.

4. Conclusion

For DAB Plus codecs (MP2 and AAC with three profiles), the dependencies of perceived audio quality on the type of signal, the number of channels, and the bitrate was examined. It was shown that MP2 encoded audio needs greater bitrate than AAC encoded audio to achieve the same quality. Speaking of the AAC codec, only the AAC-LC profile reached transparency and AAC-LC and HE-AACv1 profiles reached broadcast quality. For stereo, encoded speech signals have higher quality than music signals. For mono signals, the dependence on the type of signal is various for each codec.

In the case of AAC, mono signals did not reach the same quality as stereo signals of double bitrates because in stereo signals, the codecs can make use of inter-channel correlation and reduce average bitrate for one channel. Although the HE-AACv2 profile is an extreme case of inter-channel correlation utilization, it did not receive higher quality grade than -2. The HE-AACv1 profile, although the maximum quality for stereo signals crossed -1 (broadcast quality), is almost unusable for high-quality broadcasting, because for ODG ≈ -0.5, the needed bitrate is higher than of AAC-LC profile. Mono signals did not reach transparency with AAC-LC, either. According to the results, the High-Efficiency profiles can be used only for higher-mid or low quality services, where they perform better than the AAC-LC profile.

The High-Efficiency profiles could be set only in limited bitrate range. According to the analysis, bitrates where the profiles are outperformed by the profile lower in hierarchy are lower than the limits set by the codec creators. Also, this bitrate depends on signal type and number of channels. Significant difference between the crossing of the quality curves for different signal types can be found in stereo mode. In mono, the difference is only slight.

Furthermore, bitrates and numbers of services in the DAB+ multiplex to reach given quality criteria for each codec/profile were determined for speech and music in stereo and mono modes. According to the fact, that the High-Efficiency profiles are outperformed by the less complex ones at certain bitrate, it is not worth using these profiles, if the number of services is lower than for the less complex profile.

In the future, similar analysis will be carried out to study other digital radio systems, which are in use today.

Acknowledgments

Research described in the paper was supported by the Grant Agency of the Czech Technical University in Prague, grant No. SGS17/190/OHK3/3T/13.

References

[1] BODSON, D. Digital audio around the world. *IEEE Vehicular Technology Magazine*, 2010, vol. 5, no. 4, p. 24–30. DOI: 10.1109/MVT.2010.939105

[2] BERG, J., BUSTAD, C., JONSSON, L., et al. Perceived audio quality of realistic FM and DAB+ radio broadcasting systems. *Journal of the Audio Engineering Society*, 2013, vol. 61, no. 10, p. 755–777.

[3] GILSKI, P., STEFANSKI, J. Subjective and objective comparative study of DAB plus broadcast system. *Archives of Acoustics*, 2017, vol. 42, no. 1, p. 3–11. DOI: 10.1515/aoa-2017-0001

[4] POCTA, P., BEERENDS, J. G. Subjective and objective assessment of perceived audio quality of current digital audio broadcasting systems and web-casting applications. *IEEE Transactions on Broadcasting*, 2015, vol. 61, no. 3, p. 407–415. DOI: 10.1109/TBC.2015.2424373

[5] MANOUSELIS, N., KARAMPIPERIS, P., VARDIAMBASIS, I. O., et al. Digital audio and internet radio broadcasting systems under a QoS perspective. In *Proceedings of the 4th WSEAS International Conference on Telecommunications and Informatics*. Prague (Czech Republic), 2005, p. 1–6. ISBN: 960-8457-11-4

[6] LEE, S., LEE, Y.-T., SEO, J., et al. An audio quality evaluation of commercial digital radio systems. *IEEE Transactions on Broadcasting*, 2011, vol. 57, no. 3, p. 629–635. DOI: 10.1109/TBC.2011.2152910

[7] ETSI *ETSI Standard EN 300 401 V2.1.1. Radio Broadcasting Systems; Digital Audio Broadcasting (DAB) to Mobile, Portable and Fixed Receivers*. 2017.

[8] INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, GENEVA, SWITZERLAND. *ISO/IEC 11172-3:1993 Information Technology – Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1,5 Mbit/s – Part 3: Audio*.

[9] INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, GENEVA, SWITZERLAND. *ISO/IEC 13818-3:1998 Information Technology – Generic Coding of Moving Pictures and Associated Audio Information – Part 3: Audio*.

[10] ETSI *ETSI Specification TS 103 466 V1.1.1. Digital Audio Broadcasting (DAB); DAB audio coding (MPEG Layer II)*. 2016.

[11] INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, GENEVA, SWITZERLAND. *ISO/IEC 13818-7:1997 Information Technology – Generic Coding of Moving Pictures and Associated Audio Information – Part 7: Advanced Audio Coding (AAC)*.

[12] INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, GENEVA, SWITZERLAND. *ISO/IEC 14496-3:1999 Information Technology – Coding of Audio-visual Objects – Part 3: Audio*.

[13] ETSI *ETSI Specification TS 102 563 V2.1.1. Digital Audio Broadcasting (DAB); DAB+ Audio Coding (MPEG HE-AACv2)*. 2016.

[14] MELTZER, S., MOSER, G. HE-AAC v2 - audio coding for today's digital media world. *EBU Technical Review*, 2006, p. 37–48.

[15] ITU TELECOMMUNICATION STANDARDIZATION SECTOR, GENEVA SWITZERLAND. *Mean Opinion Score (MOS) Terminology*. Recommendation ITU-T P.800.1. 2016. Approved in 2016-07-29.

[16] ITU RADIOCOMMUNICATION SECTOR, GENEVA SWITZERLAND. *Methods for the Subjective Assessment of Small Impairments in Audio Systems*. Recommendation ITU-R BS.1116. 1994–2015. Approved in 2015-02-04.

[17] ITU RADIOCOMMUNICATION SECTOR, GENEVA SWITZERLAND. *Method for the Subjective Assessment of Intermediate Quality Levels of Coding Systems*. Recommendation BS.1534. 2001–2015. Approved in 2015-10.

[18] RUND, F., KHADDOUR, H., SCHIMMEL, J., et al. Objective quality assessment for the acoustic zoom. In *Proceedings of the 38th International Conference on Telecommunications and Signal Processing (TSP 2015)*. Prague (Czech Republic), 2015, p. 392–396. DOI: 10.1109/TSP.2015.7296290

[19] ITU RADIOCOMMUNICATION SECTOR, GENEVA SWITZERLAND. *Method for Objective Measurements of Perceived Audio Quality*. Recommendation ITU-R BS.1387-1. 1998–2001. Approved in 2001-11-24.

[20] HOLTERS, M., ZOLZER, U. GSTPEAQ - An open source implementation of the PEAQ algorithm. In *Proceedings of the 18th International Conference on Digital Audio Effects (DAFx-15)*. Trondheim (Norway), 2015, p. 181–184. ISSN: 2413-6689

[21] OPTICOM. *OPERA User Manual Version 3.5*. 204 pages. [Online] Cited 2017-08-10. Available at: http://www.opticom.de/download/OperaManual3.5.pdf

[22] HINES, A., GILLEN, E., KELLY, D., et al. ViSQOLAudio: An objective audio quality metric for low bitrate codecs. *Journal of the Acoustical Society of America*, 2015, vol. 137, no. 6, p. 449–455. DOI: 10.1121/1.4921674

[23] EUROPEAN BROADCASTING UNION, GENEVA, SWITZERLAND. *EBU Evaluations of Multichannel Audio Codecs, Phase 3*. EBU Tech 3339. March 2010.

About the Authors . . .

Karel ULOVEC was born in Prague, Czech Republic in 1977. He received the M.S. and Ph.D. degrees in electrical engineering from the Faculty of Electrical Engineering, Czech Technical University in Prague, in 2001 and 2008, respectively. Since 2004, he is an assistant professor at the Department of Radio Engineering at the Faculty of Electrical Engineering, Czech Technical University in Prague. His research interests are in areas of broadcasting systems and measurements of radio transmitters, receivers and systems.

Miroslav SMUTNÝ was born in Městec Králové in 1994. In 2013, he finished his secondary education and started studying at the Faculty of Electrical Engineering, CTU in Prague, where he obtained his bachelor's degree in Communication, Multimedia and Electronics.

Appendix 1

The following tables show possible number of services computed for protection levels 1-A, 2-A and 3-A.

	Quality criterion	Number of services						
		MP2		AAC-LC		HE-AACv1		HE-AACv2
		stereo	mono	stereo	mono	stereo	mono	stereo
music	Transparency	1	3	1	3	x	x	x
	ODG ≈ −0.5	2	5	3	6	3	x	x
	Broadcast quality	2	5	4	8	4	6	x
	ODG ≈ −2 (slightly annoying)	3	6	6	10	7	15	7
	ODG ≈ −3 (annoying)	4	9	8	15	12	20	15
speech	Transparency	1	3	2	3	x	x	x
	ODG ≈ −0.5	2	5	5	6	3	x	x
	Broadcast quality	3	6	6	8	5	x	x
	ODG ≈ −2 (slightly annoying)	4	8	8	10	8	10	10
	ODG ≈ −3 (annoying)	4	9	12	15	12	15	15

Tab. 6. Number of services to reach quality criteria for 1-A protection level.

	Quality criterion	Number of services						
		MP2		AAC-LC		HE-AACv1		HE-AACv2
		stereo	mono	stereo	mono	stereo	mono	stereo
music	Transparency	2	5	2	5	x	x	x
	ODG ≈ −0.5	4	8	5	10	5	x	x
	Broadcast quality	4	8	7	13	7	10	x
	ODG ≈ −2 (slightly annoying)	5	10	10	15	11	23	11
	ODG ≈ −3 (annoying)	7	14	13	23	18	31	23
speech	Transparency	2	5	3	5	x	x	x
	ODG ≈ −0.5	4	8	7	10	5	x	x
	Broadcast quality	5	10	10	13	8	x	x
	ODG ≈ −2 (slightly annoying)	6	12	13	15	13	15	15
	ODG ≈ −3 (annoying)	7	14	18	23	18	23	23

Tab. 7. Number of services to reach quality criteria for 2-A protection level.

	Quality criterion	Number of services						
		MP2		AAC-LC		HE-AACv1		HE-AACv2
		stereo	mono	stereo	mono	stereo	mono	stereo
music	Transparency	3	6	3	6	x	x	x
	ODG ≈ −0.5	5	11	7	13	7	x	x
	Broadcast quality	5	11	9	17	9	13	x
	ODG ≈ −2 (slightly annoying)	6	13	13	20	15	31	15
	ODG ≈ −3 (annoying)	9	19	17	31	25	41	31
speech	Transparency	3	6	5	7	x	x	x
	ODG ≈ −0.5	5	11	10	13	7	x	x
	Broadcast quality	6	13	13	17	11	x	x
	ODG ≈ −2 (slightly annoying)	8	17	17	20	17	20	20
	ODG ≈ −3 (annoying)	9	19	25	31	25	31	31

Tab. 8. Number of services to reach quality criteria for 3-A protection level.