Performance Characterization of the Low Complexity Communication Codec

***Bluetooth*® White Paper**

- **Version:** v1
- **Version Date:** 2023-06-13
- **Prepared By:** Generic Audio Working Group
- **Feedback Email:** ga-main@bluetooth.org

Abstract:

This white paper characterizes the Low Complexity Communication Codec (LC3) specification regarding typical codec aspects such as complexity, algorithmic/transmission delay, and subjective audio quality. It provides additional information beyond the technical specification of LC3 which is useful for product designers, codec implementers and profile developers.

Furthermore, this white paper summarizes the results of listening test evaluations, which were conducted during the standardization of the codec. The listening test evaluations compared LC3 to the previous Bluetooth codecs—Sub Band Codec (SBC) and modified Sub Band Codec (mSBC)—as well as other well-established coding formats used for voice transmission, music streaming, and hearing aids.

Version History

Version Number	Date (yyyy-mm-dd)	Comments
v1	2023-06-13	Approved by the Bluetooth SIG Board of Directors.

Acknowledgments

Name	Company
Alexander Tschekalinskij	Fraunhofer IIS
Markus Schnell	Fraunhofer IIS
Maximilian Schlegel	Fraunhofer IIS
Jan Büthe	Fraunhofer IIS
Nick Hunn	GN Hearing A/S
Oren Haggai	Intel Corporation
Georg Dickmann	Sonova AG
Akio Tanaka	Sony Corporation
Andrew Estrada	Sony Corporation
Riccardo Cavallari	WS Audiology Denmark A/S
Jon Tanner	Qualcomm Technologies International, Ltd. (QTIL)
Jeff Solum	Starkey Hearing Technologies
Bjarne Klemmensen	Demant A/S
Sam Geeraerts	NXP Semiconductors N.V.

Contents

1 Introduction

A detailed technical specification of the Low Complexity Communication Codec (LC3) algorithm is provided in the LC3 specification [1]. The codec's operation modes range from voice quality modes for optimal transmission of voice services to high-quality modes for distortion-free music streaming, with two frame durations of 7.5 ms and 10 ms. The codec has been designed with focus on low latency, low computational complexity, and a low memory footprint.

1.1 Scope

This document explains the use cases the LC3 is designed for and characterizes the codec's attributes such as complexity, delay, and audio quality. To characterize the codec's attributes, this document summarizes listening test data and other studies which have been conducted during the standardization phase of the LC3. Laboratory tests were commissioned by the Bluetooth SIG and performed by SenseLab, the test laboratory division of FORCE Technology. The Fraunhofer listening laboratory provided further results.

This document is important to codec implementers and especially to product designers and profile developers who will set up design constraints for actual products. It gives guidance for optimizing an LC3 implementation or choosing the best codec settings for a particular product when multiple settings are allowed (i.e., in TMAP [3]).

1.2 Bluetooth use cases

The LC3 codec was designed for the following use cases: voice, music streaming, and hearing aids.

1.2.1 Voice

Voice means the bi-directional connectivity of a Bluetooth audio accessory device for conversational use cases such as telephony or conferencing applications. For BR/EDR, this scenario is typically supported by the Hands-Free Profile (HFP) [2], and for Bluetooth Low Energy (LE) Audio, this scenario is supported by the Telephony and Media Profile (TMAP) [3].

Typically, voice calls are categorized by the audio bandwidth available in the call. The following list describes the audio bandwidth used in telephone systems:

- Narrow Band (NB), 50 Hz to 4,000 Hz at an 8 kHz sampling rate
- Wide Band (WB), 50 Hz to 8,000 Hz at a 16 kHz sampling rate
- Super Wide Band (SWB), 50 Hz to 16,000 Hz at a 32 kHz sampling rate

1.2.2 Music streaming

For Classic Bluetooth, the music streaming is provided through the Advanced Audio Distribution Profile (A2DP) [4]. For LE Audio, the music streaming is provided through TMAP [3]. In general, music streams with a high sampling rate (e.g., 44.1 kHz or 48 kHz) are forwarded from a Bluetooth Audio Source (e.g., a mobile phone) to a Bluetooth Audio Sink (e.g., headphones).

Voice and music use cases can be used at the same time, using different codec configurations for different streams (for example, performing voice recognition while streaming music).

1.2.3 Hearing aid

The Hearing Access Profile (HAP) [5] was introduced to support the adoption of LE Audio by hearing accessibility devices. Hearing aid devices have resource constraints for complexity and power consumption. Unlike earbuds, hearing aids are typically worn for periods of ten hours or more a day, without opportunities for recharging. Therefore, HAP uses WB quality for bi-directional voice applications, using a sampling rate of 16 kHz, and a semi-super wide band (SSWB) configuration for unidirectional media streaming, using a sampling rate of 24 kHz. These setups are defined in the Basic Audio Profile (BAP) [7]. These options are generally sufficient for users with hearing impairment and allow products to take advantage of reduced complexity and bitrates.

1.2.4 Broadcast Audio

The use-cases described in sections 1.2.1 to 1.2.3 are also enabled for the broadcast functionality with the Public Broadcast Profile (PBP) [18]. PBP defines Standard Quality Public Broadcast Audio at sample rates of 16 kHz and 24 kHz, and High-Quality Public Broadcast Audio at a sample rate of 48 kHz.

1.3 Codec attributes

1.3.1 Bitrates

Because LC3 [1] supports voice call and high-quality music scenarios, the codec supports bitrates from 20 bytes per frame to 400 bytes per frame per audio channel, which corresponds to bitrates from 16 kbps to 320 kbps for a 10-ms frame duration.

In the Basic Audio Profile (BAP) [7], Tables 3.11 and 3.12 list bitrate, frame duration, and sampling rate configurations for LC3 to provide a set of interoperable codec configurations. For some sampling rates, e.g., 48 kHz, more than one dedicated bitrate is provided to give implementers the choice either to optimize towards spectral capacity or towards audio quality. Figure 2.14, for example, shows that for some specific signals the audio quality is further improved by using higher bitrates for LC3.

1.3.2 Complexity

The computational complexity of the LC3 encoder and decoder has been characterized in comparison to the codecs SBC [4] and Opus/CELT (Constrained Energy Lapped Transform) [9][1]. Typical codec configurations representing the uses cases in Section 1.2 were considered, as outlined in Table 1.1.

Use Case	Codec Configuration
WB/HA voice (16 kHz sampling rate, mono)	• SBC: 66 kbps • Opus: 32 kbps, 10 ms frame duration, complexity level 0 • LC3: 32 kbps, 10 ms and 7.5 ms frame duration
SWB voice (32 kHz sampling rate, mono)	• SBC: 132.8 kbps • LC3: 64 kbps, 10 ms and 7.5 ms frame duration
Music streaming (48 kHz sampling rate, mono)	• SBC: 381 kbps • Opus: 64 kbps, 10ms frame duration, complexity level 10 • LC3: 96 kbps, 10 ms and 7.5 ms frame duration
HA music (24 kHz sampling rate, mono)	• SBC: 96 kbps • Opus: 48 kbps ,10 ms frame duration, complexity level 10 • LC3: 48 kbps, 10 ms and 7.5 ms frame duration

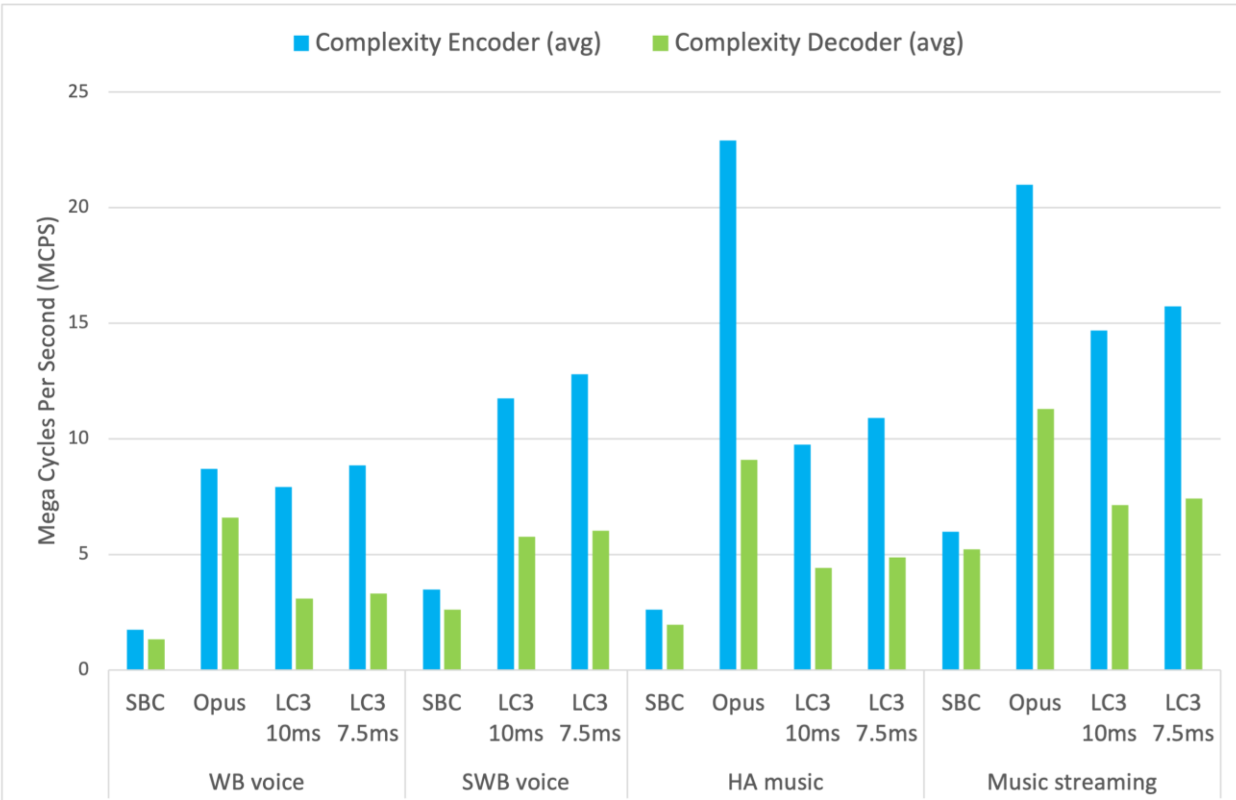
Table 1.1: Codec configurations according to use cases

[1] Opus/CELT is used as an additional reference in the complexity and audio quality evaluation.

The information presented in this section can provide guidance to product designers on the expected computational complexity and the computational delay of the LC3.

The relative complexity of the LC3 compared to SBC and Opus/CELT as well as the complexity ratio of the LC3 encoder and decoder can provide guidance toward the optimization level of the LC3 implementation. This information is of particular interest to codec implementers.

The complexity numbers for all codecs were measured in average and peak Mega Cycles Per Second (MCPS) on the same Digital Signal Processor (DSP) platform. Figure 1.1 shows the MCPS for each encoder and decoder per use case.



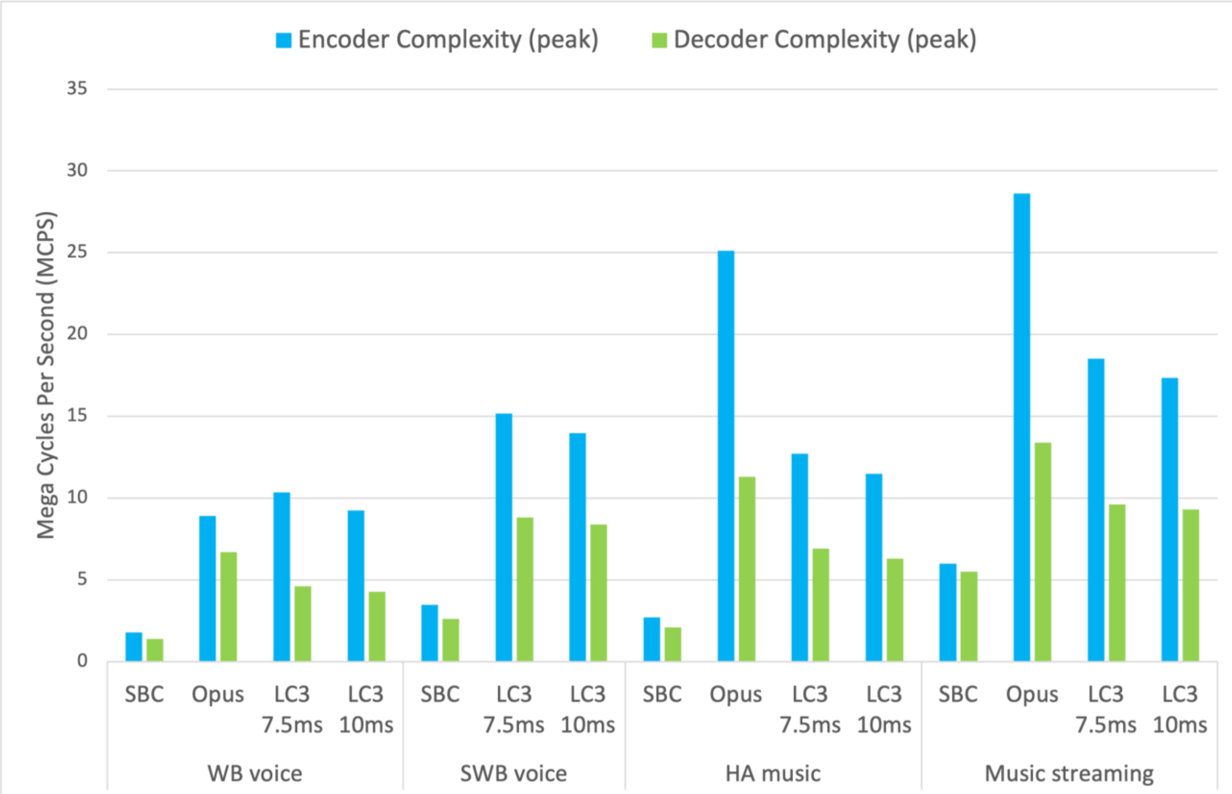


Figure 1.1: Average and peak computational complexity overview per use case

For all use cases, the total complexity of LC3 can be separated into roughly 2/3 for the encoder and 1/3 for the decoder. The 7.5 ms frame duration mode comes with a very minor computational increase compared to the 10 ms frame duration.

The encoder-to-decoder complexity ratio is useful information when optimizing the LC3 implementation for a particular DSP. The encoder complexity should be roughly twice as large as the decoder complexity. This ratio being far off could be an indicator for a missing optimization in a particular module.

Compared to SBC, the LC3 decoder is two to three times more complex and the LC3 encoder is three to four times more complex.

Note that several aspects of the codec influence the total power consumption of a transmission system. For example, because the codec bitrate determines the amount of data to be sent, the computational complexity and the communication bandwidth directly affect the power consumption of both the radio sender and receiver.

The power savings from lower communication bandwidth are far higher than the extra power usage caused by the higher computational complexity, resulting in a significant reduction of the total system power consumption.

1.3.3 System delay

An example calculation of the total system delay for LC3 can be found in the Gaming Audio Profile (GMAP) [19]. Section A.4 of GMAP [19] analyzes the overall system latency, including the codec algorithm and processing delay for a gaming bi-directional configuration with a FB stereo (gaming) audio stream and a SWB voice (back-)stream. The calculation shows that LC3 helps achieve a total system delay close to 30 ms or less for this use-case. Such low system delay is necessary for low latency applications such as gaming, amplification of live audio, and call forwarding.

1.3.4 Multi-channel coding

LC3 only codes single audio channels. Stereo or multi-channel streams are assembled by aggregating single audio channel instances. Because the audio channels are kept independent, it is possible to send the individual channels to dedicated left and right audio devices, such as untethered earbuds or separate speakers.

For untethered devices, independent channel coding minimizes the overhead of receiving data and decoding complexity compared to a joint coding approach. Joint coding implies that all channels need to be decoded at one device, which results in higher computational complexity. Additionally, a jointly coded payload is larger than a single-channel payload, which means more data needs to be received by the device, leading to higher power consumption and an increased vulnerability to interference.

This information can be considered when comparing the performance of LC3 against other codecs that implement a joint-stereo mode.

1.3.5 Packet loss concealment

Some form of Packet Loss Concealment (PLC), implemented on the receiving ends of audio connections, makes it practical to deliver satisfactory audio quality under challenging channel conditions. For example, connections that are optimized for latency are generally more error-prone and more packet losses are expected in such connections.

The purpose of PLC is to conceal the effect of unavailable or corrupted frame data for decoding. The example PLC algorithm provided in Appendix B of LC3 [1] provides a satisfactory audio quality under typical packet loss conditions. Implementers can choose to modify or implement an alternate PLC scheme to improve the performance under packet loss conditions.

In case implementers choose to implement an alternate PLC scheme, Section 2.1 can be followed for the performance evaluation of the alternate PLC implementation.

The concealment of an LC3 decoder is activated when the decoder receives an externally determined Bad Frame Indicator (BFI) flag that signals a lost frame or the presence of any detected bit error in the received channel payload to the decoder. A receiver generates such a BFI flag for each frame that indicates its integrity.

1.4 Audio quality characterization

1.4.1 Voice

For WB, the audio quality of LC3 was compared to the Bluetooth legacy voice codec mSBC [2], the legacy VoIP codec G.722 (operating in mode 1) [8], and the Opus codec.

For NB, the audio quality of LC3 was verified in comparison to the Bluetooth legacy voice codec Continuously Variable Slope Delta (CVSD) (see Volume 2, Part B, Section 9 of [6]).

For SWB, LC3 was characterized in comparison to SBC codec and Opus.

Additionally, the transcoding performance with the common mobile codecs, (i.e., Adaptive Multi-Rate (AMR) [12], Adaptive Multi-Rate Wideband (AMR-WB) [13], and Enhanced Voice Service (EVS) [14]) was characterized in comparison to the reference codecs mSBC, CVSD, SBC, and Opus.

1.4.2 Music streaming

For music streaming, there were two performance goals targeted: lower bitrate and higher quality compared to A2DP [4], which uses SBC. Therefore, the audio quality of LC3 was characterized in comparison to SBC using the standard settings for medium and high-quality for A2DP. Additionally, LC3 was compared to Opus.

1.4.3 Hearing aids

The WB voice criteria are identical to the criteria listed in Section 1.4.1. For HA music streaming, LC3 was characterized in comparison to G.722 (operating in mode 1 and at 24 kHz), SBC, and Opus (operating at the highest complexity).

2 Listening tests

Sections 2.2 through 2.4 describe the subjective listening tests, which have been conducted to verify the performance requirements described in Section 1.4 and to characterize the audio quality of LC3.

These sections can give guidance to product designers because they present the audio quality performance of the evaluated codecs.

In particular, Sections 2.3.4 to 2.3.7 present not only the audio quality performance of LC3 compared to SBC for music streaming, but also the relative performance of different (optional) LC3 bitrate operation points that are standardized in TMAP [3]. This information can give guidance about whether it may be useful to spend more bits than the minimum bitrate or at which bitrate most users cannot hear any difference from the original signal.

Furthermore, the results of these sections explain the expected bitrate savings with LC3 compared to legacy Bluetooth codecs.

These listening tests only evaluated channel conditions without transmission error (e.g., packet losses). Possibly, transmission errors distort the audio signal and such distortions might have a significant impact on the user experience. When designing Bluetooth products, it is therefore important to implement strategies that counter the effect of transmission errors (e.g., packet loss concealment and retransmission, see also Section 1.3.5).

Three test methods - P.800 Absolute Category Rating (ACR) [11], MUlti Stimulus test with Hidden Reference and Anchor (MUSHRA) [10], and the ITU-R BS.1116-3 test procedure [16] - have been used, depending on the use case. Because different methods were used, any comparison between different listening tests is not valid because using different reference conditions results in a different absolute scale.

While expert or trained listeners rated the conditions for MUSHRA and BS.1116 tests, non-expert listeners rated the conditions in ACR tests. Non-expert (or naive) listeners are considered as typical equipment users without any specific experience in the field of audio processing or listening tests. For ACR tests, un-distorted audio conditions (direct) were graded from zero up to the maximum score of 5.0. A set of Modulated Noise Reference Unit (MNRU) with certain levels of distortion level Q was used to encourage the listener to use the complete scale.

For all MUSHRA tests, statistical groups (A, B, C, …) have been identified by the test lab to find statistically significant differences between several conditions. The groups are based on a Tukey Honestly Significant Difference (HSD) test, which is a type of statistical t-test. In the following, the test was conducted using a confidence level of 95%. Note that whether two conditions differ significantly requires statistical calculation with an HSD test and cannot be directly determined from the visual representation.

2.1 Packet Loss Concealment

This section gives guidance on the performance evaluation of a possible alternate PLC algorithm implementation.

2.1.1 Setup

In order to evaluate the performance of an alternate PLC implementation, a P.800 Absolute Category Rating (ACR) test can be used. This test approach is more reliable than using objective audio quality tools such as PEAQ or P.OLQA for testing the impact of packet losses.

2.1.2 Example PLC performance evaluation

This section presents an example evaluation of an alternate PLC scheme for LC3, where “BT Std PLC” represents the example PLC algorithm provided in Appendix B of LC3, and “FhG Adv PLC” represents the alternate PLC scheme.

The following Bluetooth codecs were evaluated in this WB P.800 ACR test conducted by Fraunhofer with 25 subjects:

- LC3: 32 kbps, 10 ms frame duration, BT Std PLC
- LC3: 32 kbps, 7.5 ms frame duration, BT Std PLC
- LC3: 32 kbps, 10 ms frame duration, FhG Adv PLC
- LC3: 32 kbps, 7.5 ms frame duration, FhG Adv PLC
- mSBC: 60.4 kbps
- OPUS: 32 kbps, 10 ms frame duration, complexity level 0

at the following packet loss rates:

- 0 % PLR
- 1.5 % PLR
- 3 % PLR
- 6 % PLR
- 12 % PLR

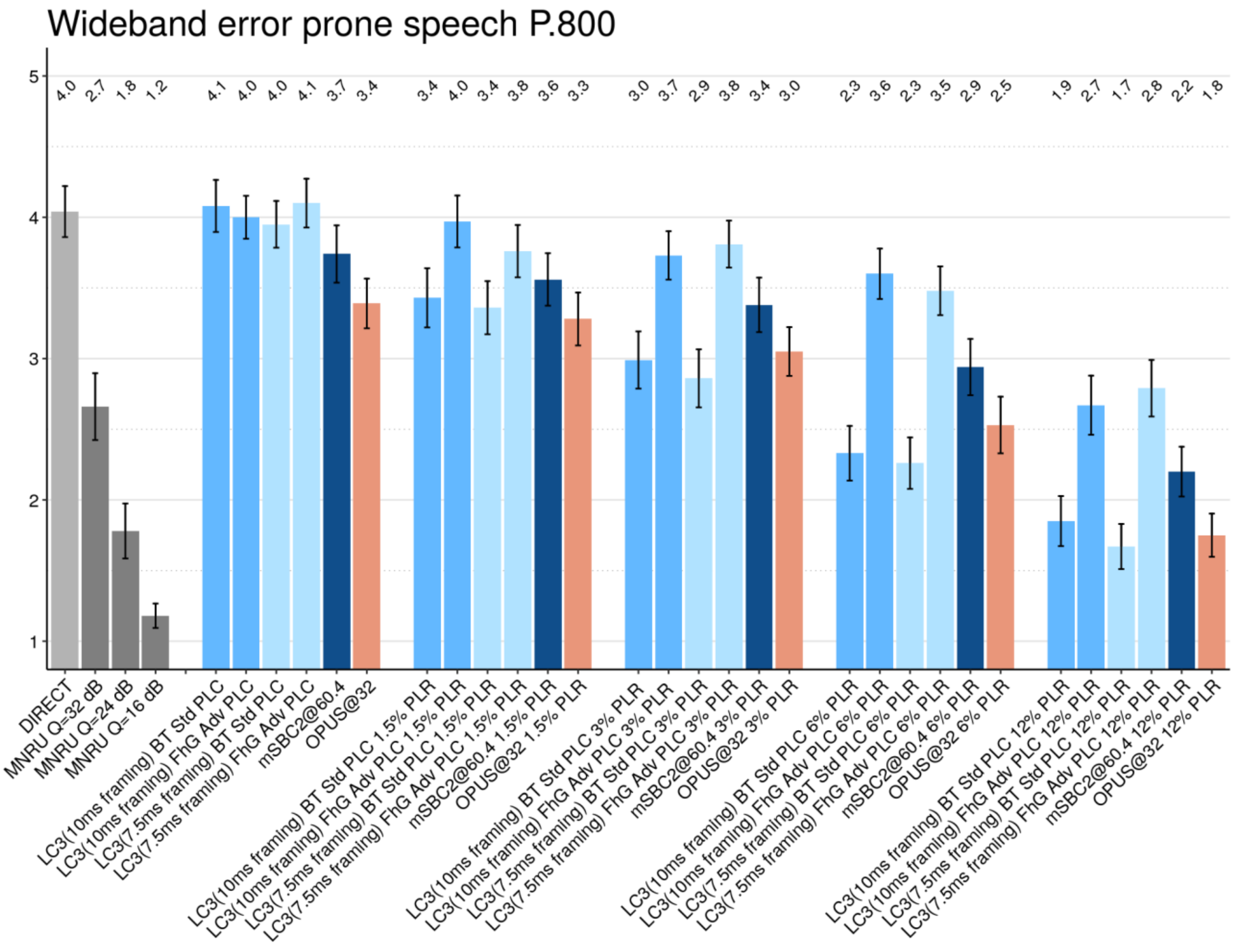


Figure 2.1: P.800 ACR mean scores and 95 percent confidence intervals for WB error-prone voice listening test

From these results, we conclude the following:

- At all packet loss rates above 0%, the alternate PLC scheme “FhG Adv PLC” performs significantly better than the LC3 example PLC algorithm “BT Std PLC”.

2.2 Voice

2.2.1 Setup

The WB and SWB cases were evaluated in lab testing using the MUSHRA test methodology. The SWB configuration of LC3 for HFP was also evaluated using the MUSHRA test methodology.

In addition, all voice use cases were evaluated in lab testing using the ACR test methodology. These tests evaluated the intrinsic speech quality and the transcoding performance of codecs deployed in mobile networks. Figure 2.2 outlines the cases included in the transcoding tests:

- Case A: Receiver is using a Bluetooth headset
- Case B: Sender is using a Bluetooth headset
- Case C: Both sender and receiver are using Bluetooth headsets

Note that the test material has been generated by a simulation using several encoding and decoding processes rather than by real transmissions over several devices.

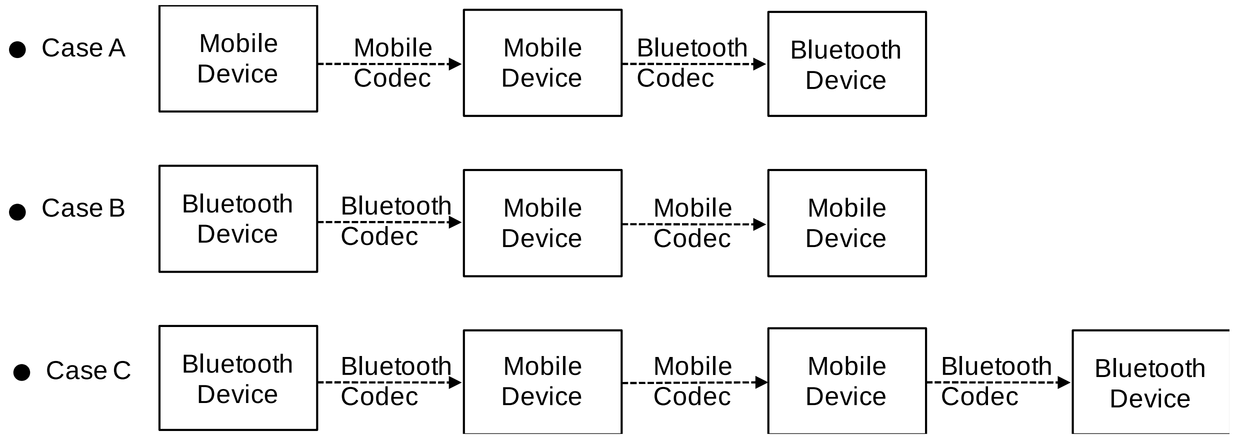


Figure 2.2: Transcoding cases with mobile networks

Table 2.1 lists all four listening tests for the voice use case. Bluetooth devices used the codecs CVSD, mSBC, Opus (as an additional reference), or LC3; mobile devices used the 3GPP codecs AMR [12], AMR-WB [13], or EVS [14].

Section	Description	Conditions	Number of Subjects
2.2.3	NB P.800 ACR, Fraunhofer	Bluetooth codecs • CVSD: 64 kbps • LC3: 24 kbps, 10 ms frame duration • LC3: 27.8 kbps, 7.5 ms frame duration Mobile codecs • AMR: 10.2 kbps	25
2.2.4	WB P.800 ACR, Fraunhofer	Bluetooth codecs • mSBC: 60.4 kbps • LC3: 32 kbps, 10 ms frame duration • LC3: 32 kbps, 7.5 ms frame duration • Opus: 32 kbps, 10 ms frame duration, complexity level 0 Mobile codecs • AMR-WB: 12.65 kbps and 23.85 kbps • EVS: 13.2 kbps (labeled EVS-WB)	25
2.2.5	SWB P.800 ACR, Fraunhofer	Bluetooth codecs • SBC: 128 kbps • LC3: 64 kbps, 10 ms frame duration • LC3: 64 kbps, 7.5 ms frame duration	25

Section	Description	Conditions	Number of Subjects
		• Opus: 64 kbps, 10 ms frame duration, complexity level 0 Mobile codecs • EVS: 13.2 kbps and 24.4 kbps (labeled EVS-SWB)	
2.2.6	SWB MUSHRA, SenseLab	Bluetooth codecs • SBC: 128 kbps • LC3: 64 kbps, 10 ms frame duration • LC3: 64 kbps, 7.5 ms frame duration • LC3 (HFP configuration): 61.8 kbps, 7.5 ms frame duration Mobile codecs • none	20
2.4.4	WB Voice II (16 kHz), MUSHRA, SenseLab	See Table 2.6 (part of HA listening tests)	See Table 2.6

Table 2.1: Summary of listening tests for voice use case

2.2.2 Test items

All P.800 ACR (2.2.3, 2.2.4, 2.2.5) tests were conducted using the German language with a speech corpus of 24 samples. The corpus was evaluated by non-expert German native listeners.

Table 2.2 lists the items selected for the SWB MUSHRA tests (2.2.6, 2.4.4). The material is from the European Broadcast Union (EBU) test material [15].

Sample	Designation	Material ID	Description	Length (Seconds)
1	Female/male speech	49 50	Female speech, English Male speech, English	12 s
2	Female/male speech and car noise	49 50	Female speech, English, mixed with car background noise at 7 dB Male speech, English, mixed with car background noise at 7 dB	12 s
3	Music	61 66	Soprano Classical music	18 s
4	Music	69	Pop artist	16 s

Table 2.2: Test material for the SWB MUSHRA tests

2.2.3 NB P.800 results

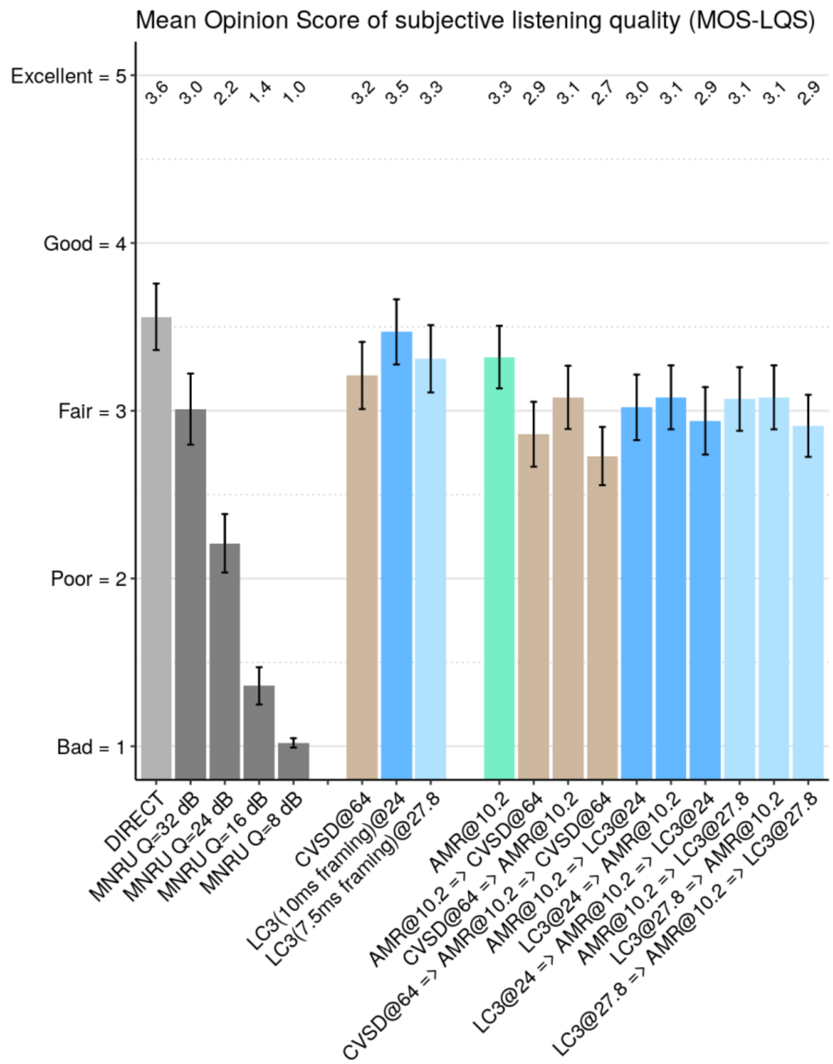


Figure 2.3: P.800 ACR mean scores and 95 percent confidence intervals for NB voice listening test

From these results, we conclude the following:

- During this testing, LC3 was not worse than the former Bluetooth NB codec CVSD, which operates at more than twice the bitrate of LC3.
- During this testing, LC3 was not worse than CVSD in any transcoding case, meaning that transcoding capability with AMR 10.2 kbps is confirmed under test conditions.

2.2.4 WB P.800 results

Because of the large number of WB conditions, the listening test has been separated into parts I and II. Part I contains the direct comparison of the codecs and transcoding with AMR-WB at 12.65 kbps, while part II contains the transcoding with AMR-WB at 23.85 kbps and EVS-WB at 13.2 kbps.

2.2.4.1 Part I

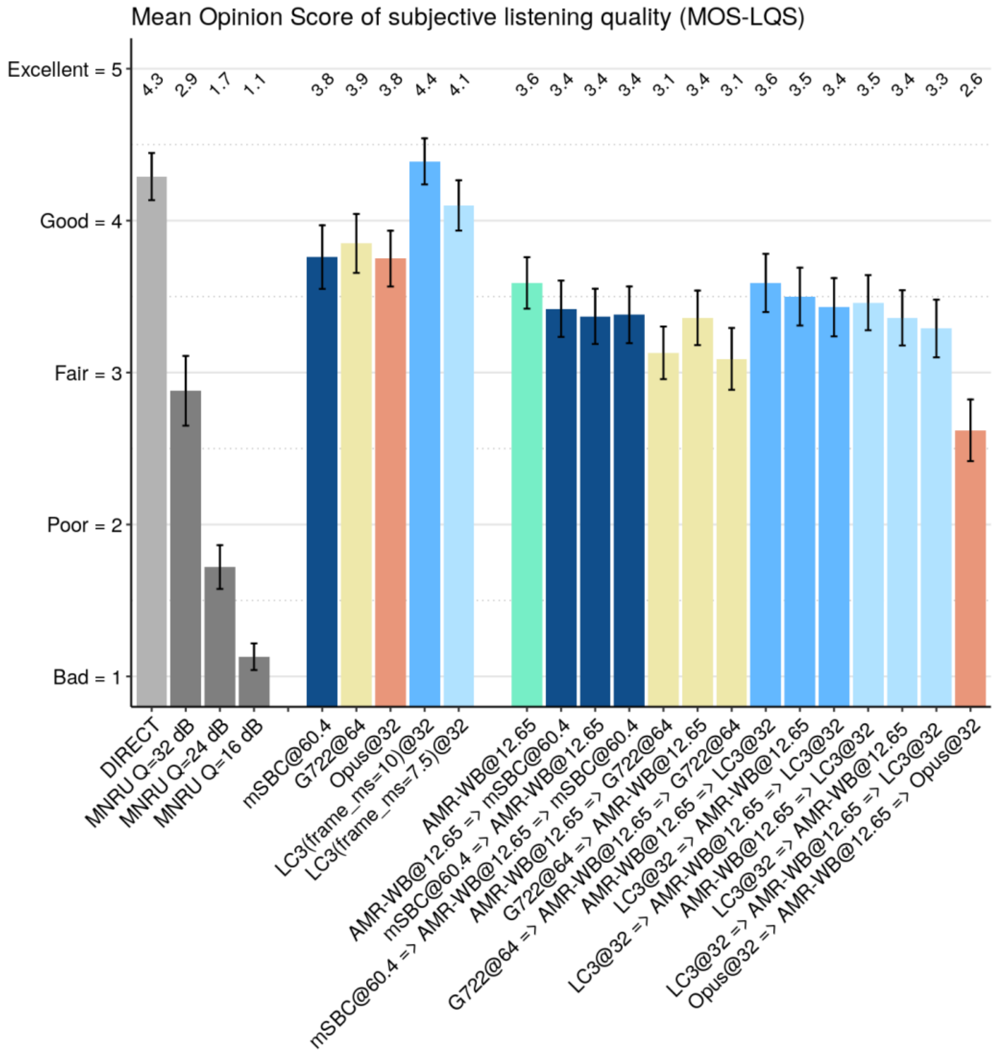


Figure 2.4: P.800 ACR mean scores and 95 percent confidence intervals for WB voice listening test part I

From these results, we conclude the following:

- For intrinsic audio quality, LC3 outperformed G.722 and mSBC during this testing, even though G.722 and mSBC operate at approximately twice the bitrate of LC3.
- For transcoding scenarios with the AMR-WB at 12.65 kbps, LC3 was not worse than G.722 and mSBC, which operate at approximately twice the bitrate of LC3. LC3's transcoding capability with AMR-WB at 12.65 kbps is confirmed under test conditions.

2.2.4.2 Part II

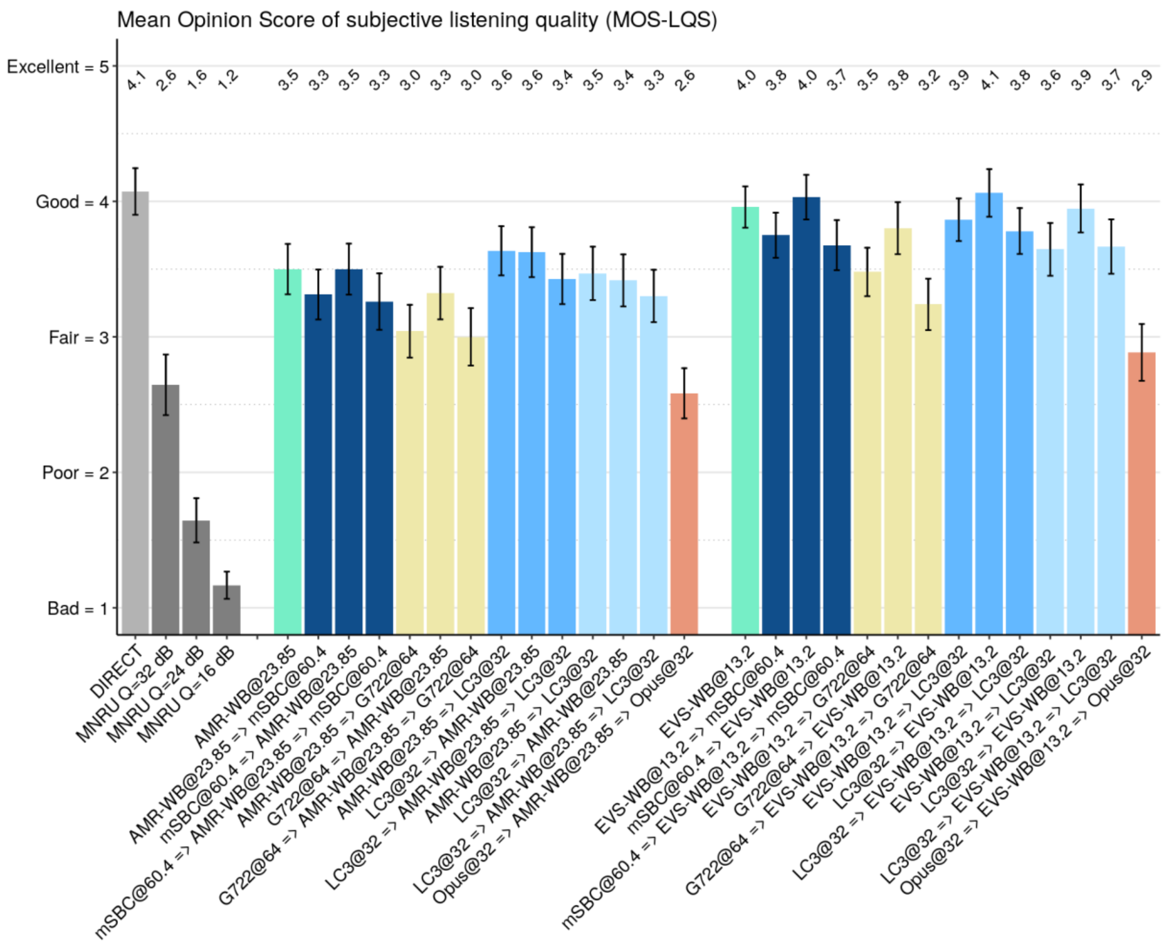


Figure 2.5: P.800 ACR mean scores and 95 percent confidence intervals for WB voice listening test part II

From these results, we conclude the following:

- For transcoding scenarios during this testing, LC3 was not worse than G.722 and mSBC, which operate at approximately twice the bitrate of LC3.
- The transcoding capability of LC3 with AMR-WB 23.85 kbps and EVS-WB 13.2 kbps was confirmed under test conditions without any severe quality degradation.

2.2.5 SWB P.800 results

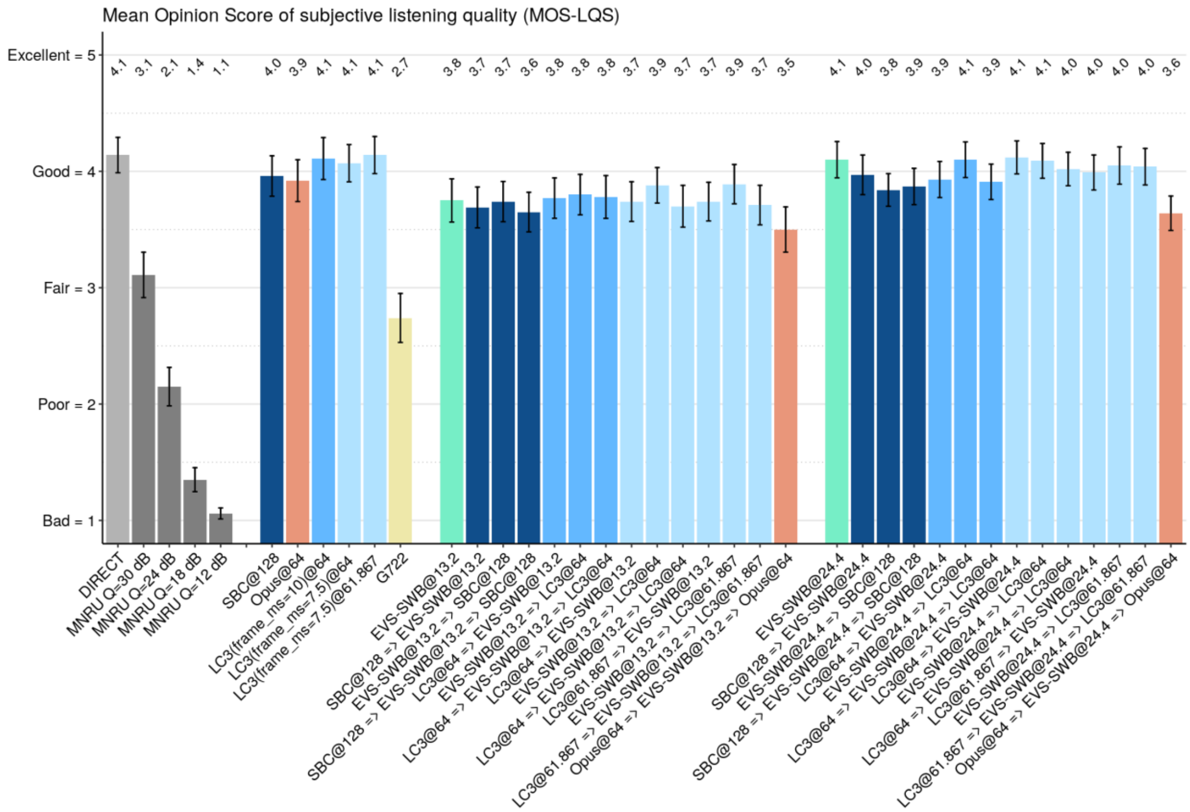


Figure 2.6: P.800 ACR mean score and 95 percent confidence interval for SWB voice listening test

From these results, we conclude the following:

- For intrinsic quality, LC3 was not worse than SBC at twice the bitrate during this testing.
- For transcoding scenarios, LC3 was not worse than SBC at twice the bitrate during this testing.
- The transcoding capability of LC3 with EVS-SWB 13.2 kbps and EVS-SWB 24.4 kbps was confirmed without any severe quality degradation under test conditions.
- The LC3 configuration for HFP showed no significant difference from the other LC3 configurations during this testing.
- During this testing, LC3 operating in SWB improved the perceived quality by more than 1 Mean Opinion Score (MOS) compared to the typical wide-band codec G.722.

2.2.6 SWB MUSHRA results

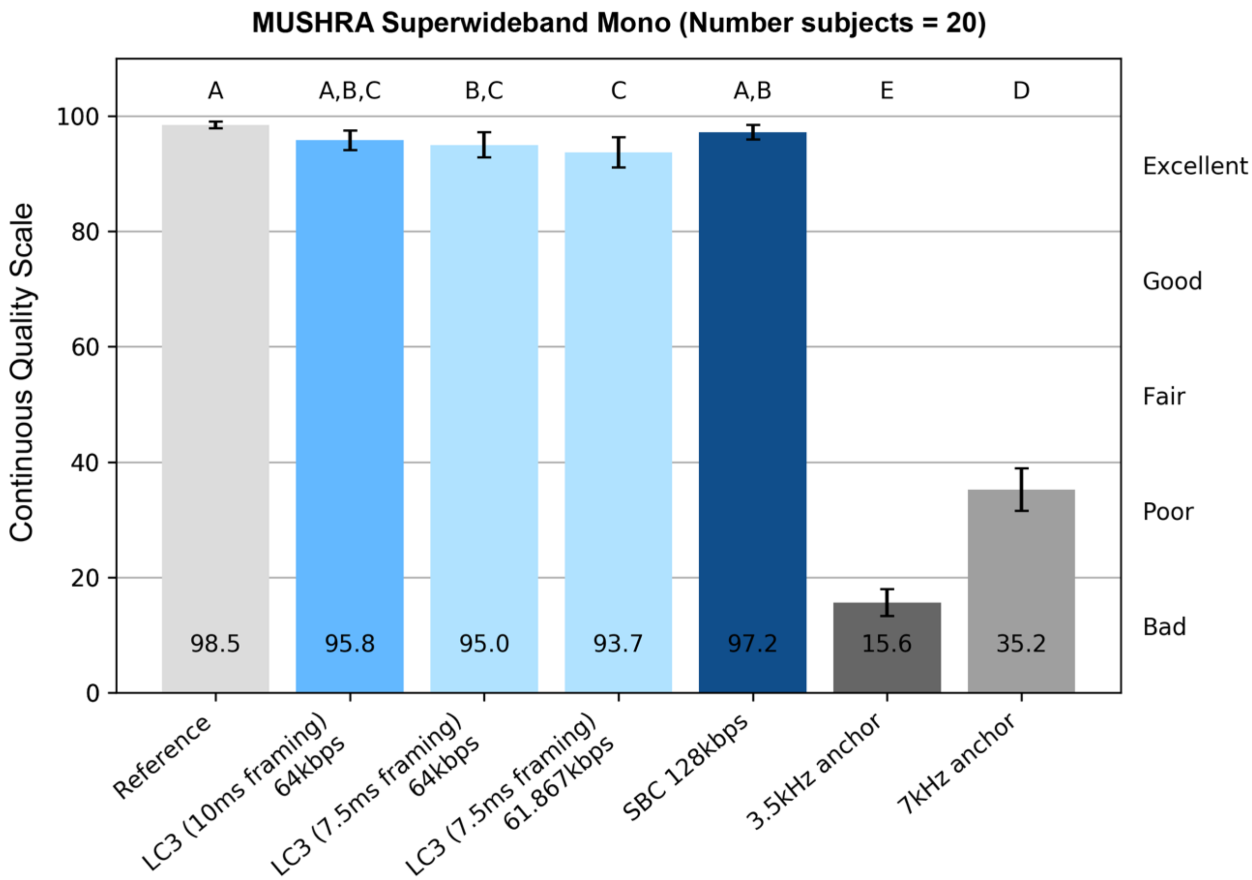


Figure 2.7: Mean scores, 95 percent confidence intervals and statistical grouping (A-E) over all items of the SWB MUSHRA test. Statistical grouping means that conditions with the same letter show no statistically significant difference.

From these results, we conclude the following:

- LC3 at 64 kbps was not worse than SBC at 128 kbps during this testing, and therefore LC3 provided 50 percent higher compression under test conditions.
- All operating points under test were rated in the “Excellent” category.
- There was no significant difference in audio quality between LC3 (10 ms frame duration) at 64 kbps, LC3 (7.5 ms frame duration) at 64 kbps, and LC3 (7.5 ms frame duration) at 61.867 kbps during this testing.

2.3 Music streaming

2.3.1 Setup

The music streaming use case was evaluated in the following two testing phases:

1. Phase I: MUSHRA online tests open for all Associate and Promoter Bluetooth members
2. Phase II: BS.1116 tests conducted in lab testing

In the first testing phase, a MUSHRA test was conducted to compare LC3 to the reference codecs Opus and SBC. An online listening test environment was used, which allowed all Bluetooth members to participate.

For the second testing phase, four listening tests were carried out in lab testing, using the ITU-R BS.1116-3 test procedure [16]. The listening tests were separated by frame duration of 10 ms and 7.5 ms as well as by mono and stereo audio content. The LC3 was compared to the SBC for medium quality and high quality, as used in A2DP [4], at the bitrates 80 kbps, 96 kbps, and 124 kbps, and the corresponding maximum rate of A2DP, i.e., 198 kbps for mono and 2x172 kbps for stereo. The sampling rate of the audio material was 48 kHz for all test cases, except for one LC3 condition in which the sampling rate was 44.1 kHz.

Table 2.3 lists all five listening tests for the music streaming use case. In the phase I listening test, 118 Bluetooth members participated. In phase II for the 10 ms frame duration, 24 post-screened[2] expert listeners participated, while for the 7.5 ms frame duration 20 post-screened listeners participated. For stereo, LC3 operated in dual-mono, meaning that each channel was coded separately at the same bitrate and labeled, for example, as “2x96 kbps”.

Note: The data has been normalized using a statistical z-transformation.

Section	Description	Conditions	Number of Subjects
2.3.3	Phase I, Stereo, 10 ms LC3 frame duration, Bluetooth members	• SBC: 2x96 kbps, 2x198 kbps • LC3: 2x96 kbps, 2x124 kbps • Opus: 2x96 kbps	118
2.3.4	Phase II, Mono, 10 ms LC3 frame duration, SenseLab	• SBC: 132 kbps, 198 kbps • LC3: 80 kbps, 96 kbps, 124 kbps, 198 kbps • LC3 (44.1 kHz): 95.55 kbps	24
2.3.5	Phase II, Mono, 7.5 ms LC3 frame duration, SenseLab	• SBC: 132 kbps, 198 kbps • LC3: 80 kbps, 96 kbps, 124 kbps, 198 kbps • LC3 (44.1 kHz): 95.55 kbps	20
2.3.6	Phase II, Stereo, 10 ms frame duration, SenseLab	• SBC: 237 kbps, 345 kbps • LC3: 2x80 kbps, 2x96 kbps, 2x124 kbps, 2x172 kbps • LC3 (44.1 kHz): 2x95.55 kbps	24
2.3.7	Phase II, Stereo, 7.5 ms frame duration, SenseLab	• SBC: 237, 345 kbps • LC3: 2x80 kbps, 2x96 kbps, 2x124 kbps, 2x172 kbps • LC3 (44.1 kHz): 2x95.55 kbps	20

Table 2.3: Summary of listening tests for music streaming use case

2.3.2 Test items

Table 2.4 lists the items selected for the MUSHRA online tests in phase I. Four samples were selected. The material is from the EBU test material [15].

[2] BS.1116-3 contains rules for post-screening listener scores and excluding less reliable subjects.

Sample	Designation	Material ID	Description	Length (Seconds)
1	Music	57 63	Classical Organ Opera	15 s
2	Music	61 66	Soprano Classical Music	18 s
3	Music	69	Pop artist	14 s
4	Single instruments	27 32 41	Castanets/rhythm Triangles/roll Celesta/melodious phrase	11 s

Table 2.4: Test material for MUSHRA online tests

Table 2.5 lists the items selected for the BS.1116 music streaming tests in phase **II**. Six items were selected based on the objective quality metric Perceptual Evaluation of Audio Quality (PEAQ) [17] in order to find the most challenging items out of the EBU test material [15] for all conditions. For comprehensive material coverage, two items for each of the signal class categories (tonal, transient, and complex) were selected.

Description	Signal Class	EBU[3] SQAM Track	Length (Seconds)
Pop artist	complex	69	12
Bass clarinet	tonal	17	10
Castanet	transient	27	8
Glockenspiel	tonal & transient	35	10
Triangle	tonal & transient	32	8
Classical music	complex	67	10

Table 2.5: Selected items for BS.1116 music streaming tests

[3] Refers to the track on the European Broadcasting Union’s TECH 3253 - Sound Quality Assessment Material recordings for subjective tests CD [15].

2.3.3 MUSHRA test

2.3.3.1 Results over all items

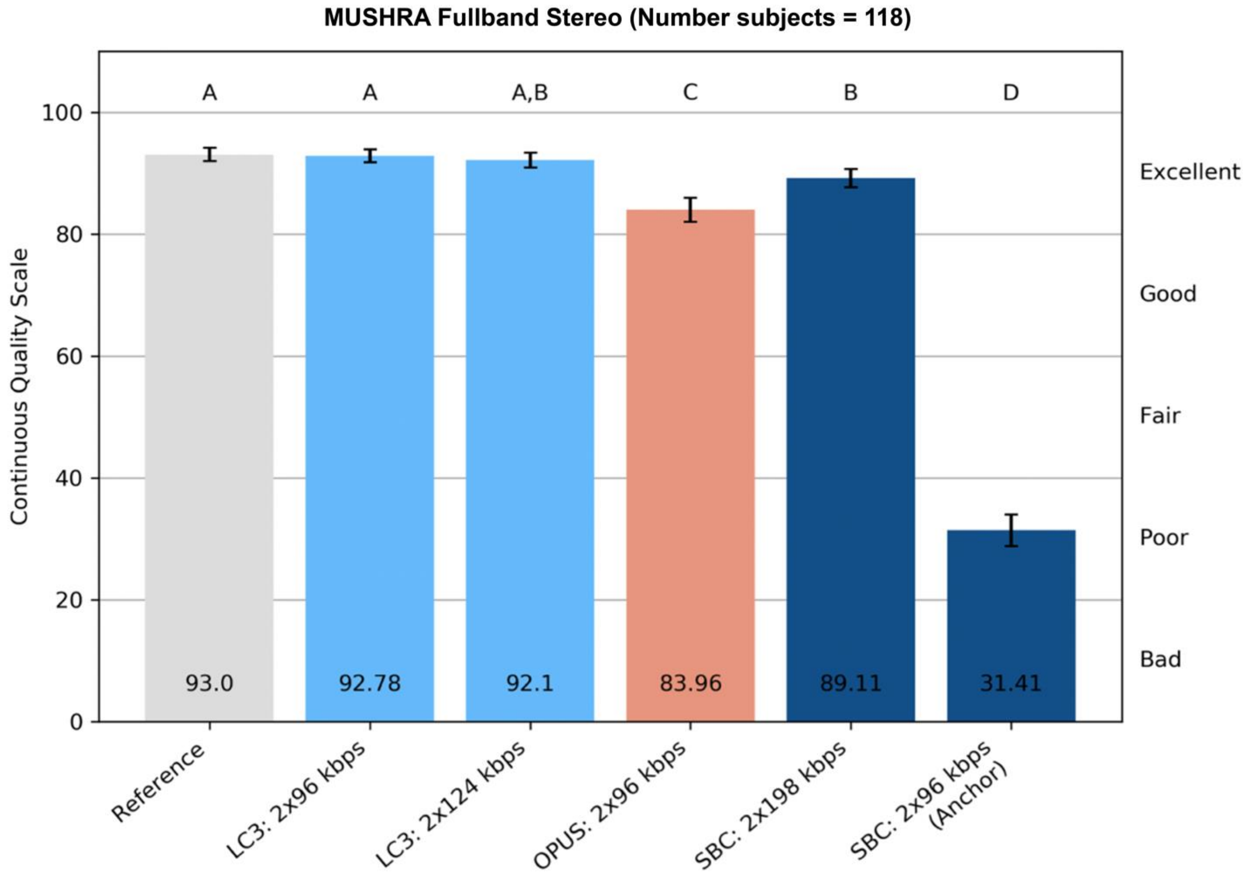


Figure 2.8: Mean scores, 95 percent confidence intervals, and statistical grouping (A-D) over all items in the MUSHRA stereo test for 10 ms frame duration. Statistical grouping means that conditions with the same letter show no statistically significant difference.

From these results, we conclude the following:

- During this testing, LC3 (10 ms frame duration) at 2x96 kbps and at 2x124 kbps showed no statistically significant difference from the uncoded original (Reference) because all conditions were in statistical group A.
- LC3 (10 ms frame duration) at 2x96 kbps was in statistical group A, while SBC at 2x198 kbps was in statistical group B, meaning that LC3 at 2x96 kbps performed significantly better than SBC at 2x198 kbps during this testing.
- During this testing, the overall quality of LC3 was very close to the Reference, which required a more sensitive listening test methodology. Therefore, BS.1116 was chosen for phase II testing.

2.3.4 LC3-Mono with 10 ms frame duration

2.3.4.1 Results over all items

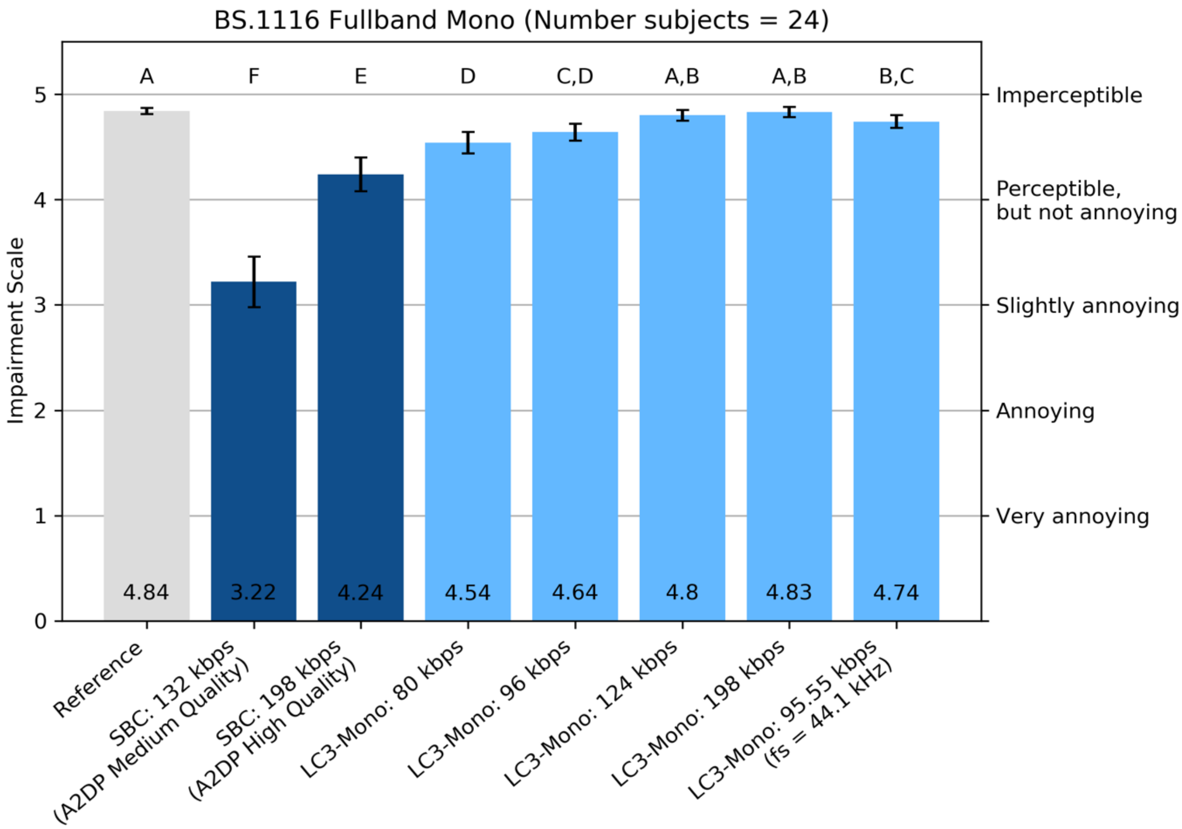


Figure 2.9: Mean scores, 95 percent confidence intervals, and statistical grouping (A-E) over all items in the BS.1116 mono test for 10 ms frame duration. Statistical grouping means that conditions with the same letter show no statistically significant difference.

From these results, we conclude the following:

- During this testing, LC3-Mono (10 ms frame duration) at 124 kbps and at 198 kbps showed no statistically significant difference from the uncoded original (Reference).
- During this testing, LC3-Mono (10 ms frame duration) at 80 kbps performed significantly better than SBC at 198 kbps.
- During this testing, LC3-Mono (10 ms frame duration) at 95.55 kbps with a sample rate of 44.1 kHz performed equivalently to LC3-Mono (10 ms frame duration) at 96 kbps and a sample rate of 48 kHz, which verified the performance of LC3 for 44.1 kHz signals under test conditions.

2.3.4.2 Results for individual items

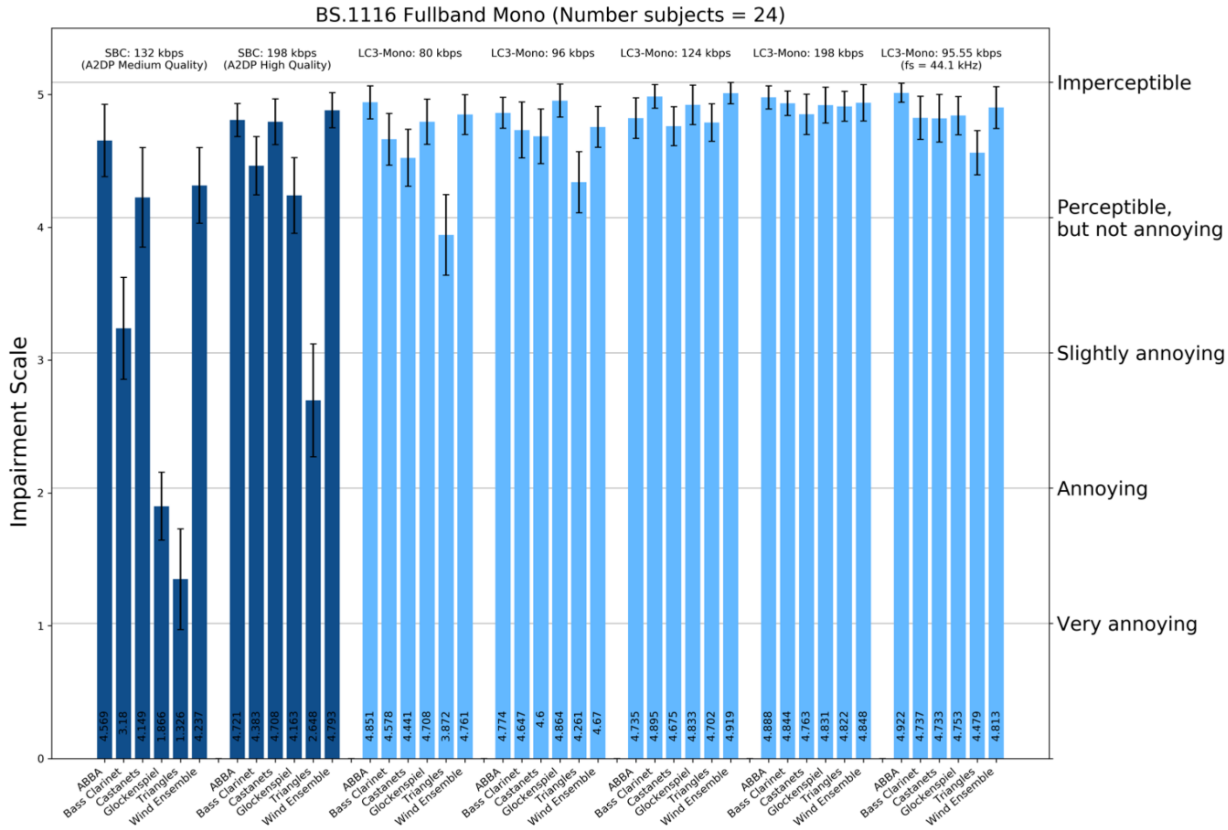


Figure 2.10: Mean ratings and 95 percent confidence intervals for the mono test for 10 ms frame duration per sample

From these results, we conclude the following:

- Starting from 96 kbps per channel, all items coded with LC3-Mono (10 ms frame duration) showed a rating over 4.0, which means there were no annoying distortions during this testing.
- For SBC as specified in A2DP [4], at least one item always showed an annoying quality level during this testing.

2.3.5 LC3-mono with 7.5 ms frame duration

2.3.5.1 Results over all items

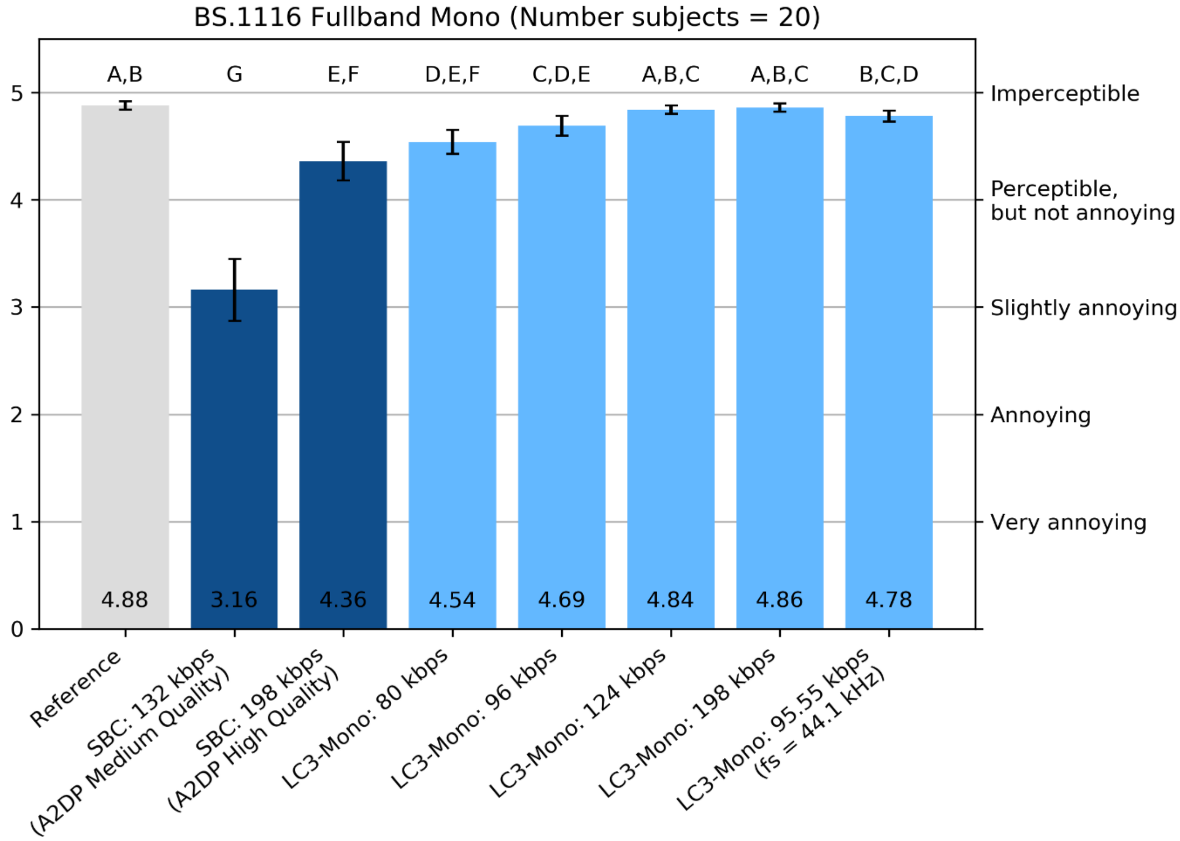


Figure 2.11: Mean scores, 95 percent confidence intervals, and statistical grouping (A-G) over all items in the BS.1116 mono test for 7.5 ms frame duration. Statistical grouping means that conditions with the same letter are not significantly different from at least one of the other conditions in the group.

From these results, we conclude the following:

- During this testing, LC3-Mono (7.5 ms frame duration) at 124 kbps and 198 kbps showed no statistically significant difference from the uncoded original (Reference).
- During this testing, LC3-Mono (7.5 ms frame duration) at 80 kbps has a higher mean quality score than SBC at 198 kbps.
- During this testing, LC3-Mono (7.5 ms frame duration) at 95.55 kbps with a sample rate of 44.1 kHz performed equivalently to LC3-Mono (7.5 ms frame duration) at 96 kbps and a sample rate of 48 kHz, which verified the performance of LC3 for 44.1 kHz signals under test conditions.

2.3.5.2 Results for individual items

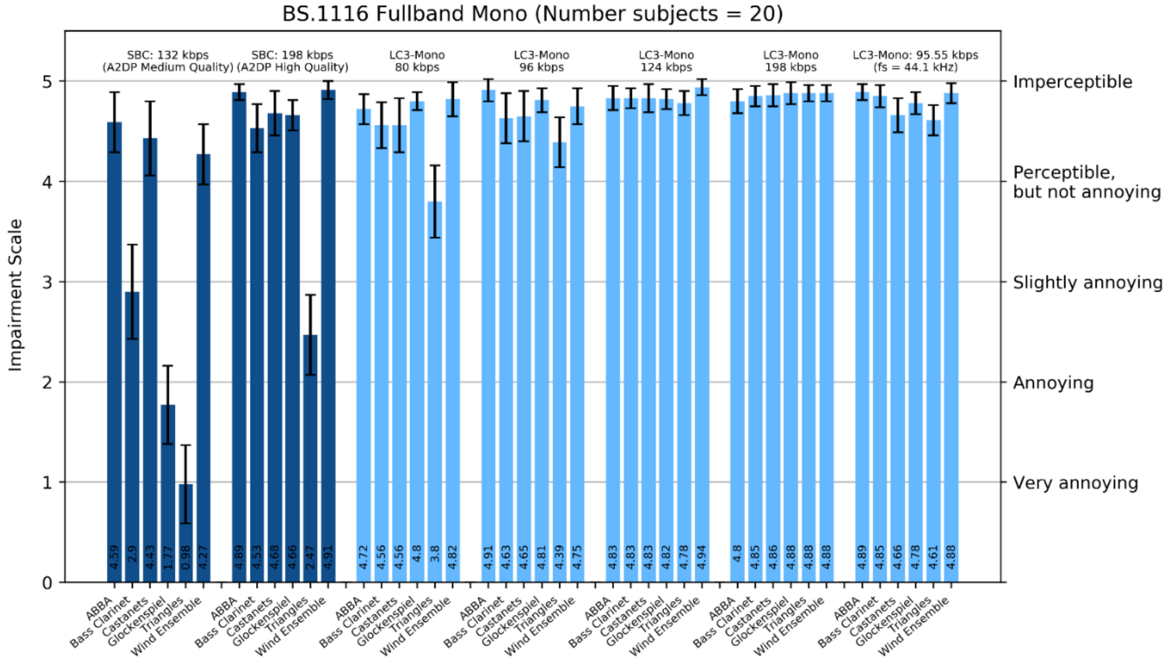


Figure 2.12: Mean ratings and 95 percent confidence intervals for the mono test for 7.5 ms frame duration per sample

From these results, we conclude the following:

- Starting from 96 kbps per channel, all items coded with LC3-Mono (7.5 ms frame duration) showed a rating over 4.0, which means there were no annoying distortions during this testing.
- For SBC as specified in A2DP [4], at least one item always showed an annoying quality level during this testing.

2.3.6 LC3-Stereo with 10 ms frame duration

2.3.6.1 Results over all items

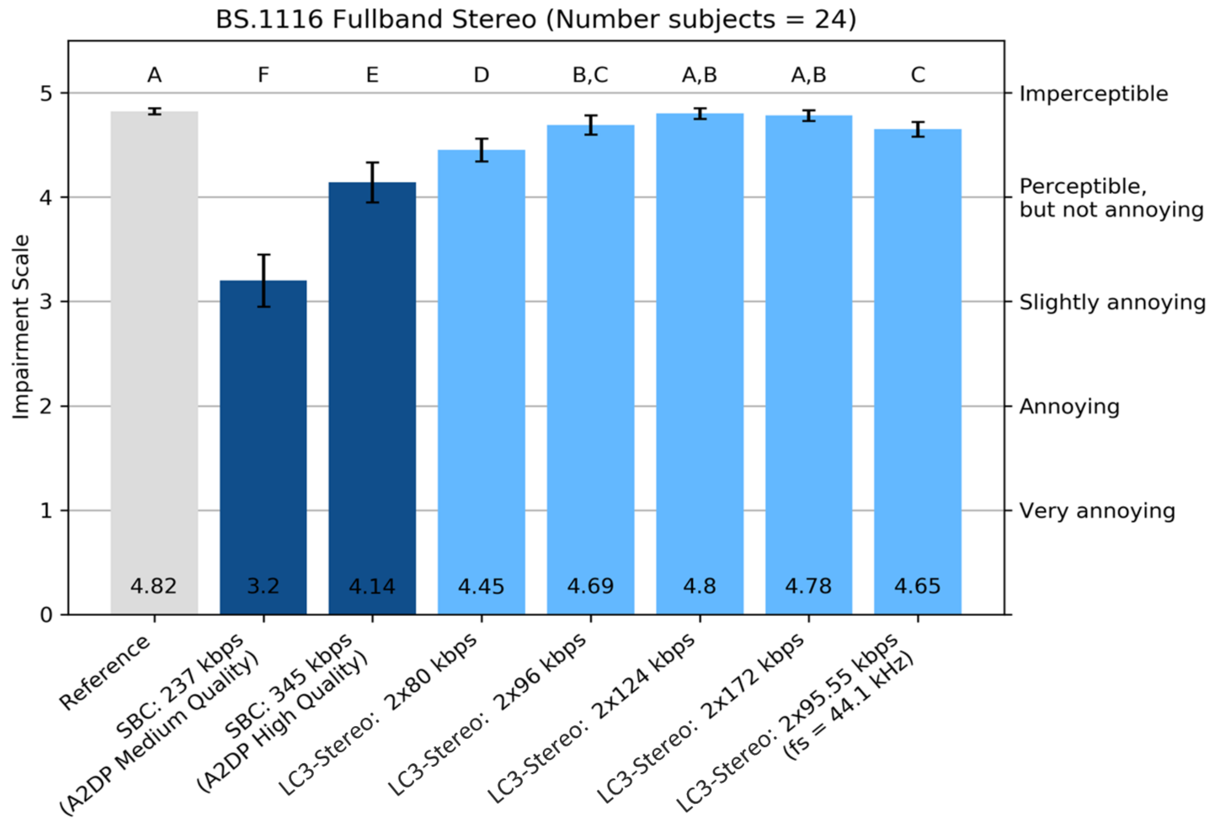


Figure 2.13: Mean scores, 95 percent confidence intervals, and statistical grouping (A-F) over all items in the BS.1116 stereo test for 10 ms frame duration. Statistical grouping means that conditions with the same letter show no statistically significant difference. LC3 operates in dual-mono with two times the bitrates for each channel.

From these results, we conclude the following:

- During this testing, LC3-Stereo (10 ms frame duration) at 2x124 kbps and 2x172 kbps showed no statistically significant difference from the uncoded original (Reference).
- During this testing, LC3-Stereo (10 ms frame duration) at 2x80 kbps performed significantly better than SBC at 345 kbps.
- During this testing, LC3-Stereo (10 ms frame duration) at 2x95.55 kbps with a sample rate of 44.1 kHz performed equivalently to LC3-Stereo (10 ms frame duration) at 2x96 kbps with a sample rate of 48 kHz, which verified the performance of LC3 for 44.1 kHz signals under test conditions.
- Therefore, under test conditions, LC3 was able to provide a higher level of quality compared to SBC when operating at the same bitrate, or was able to reduce the bitrate by roughly 50 percent while providing similar quality.

2.3.6.2 Results for individual items

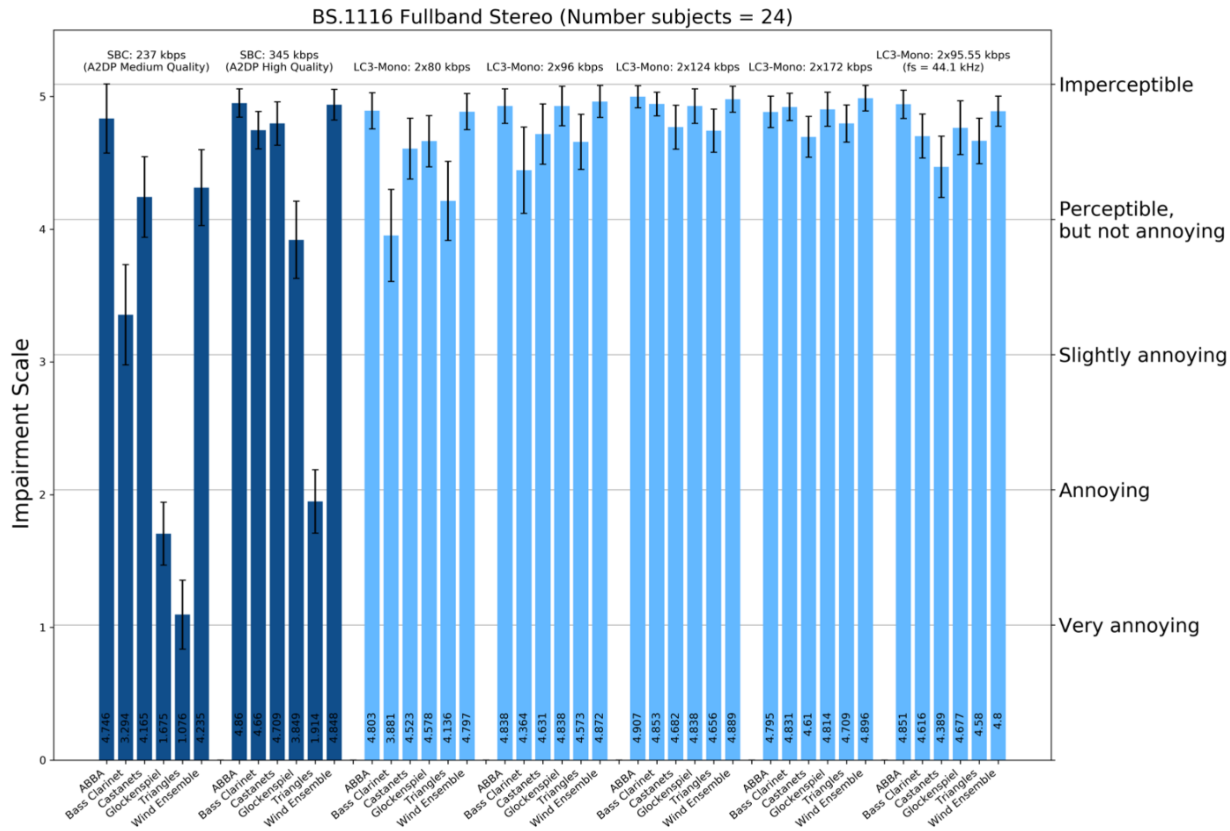


Figure 2.14: Mean ratings and 95 percent confidence intervals for the stereo test for 10 ms frame duration per sample

From these results, we conclude the following:

- Starting from 96 kbps per channel, all items coded with LC3-Stereo (10 ms frame duration) showed a rating above 4.0, which means there were no annoying distortions during this testing.
- For SBC, as specified in A2DP [4], at least two items always received a rating near 2.0, indicating an annoying quality level during testing.

2.3.7 LC3-Stereo with 7.5 ms frame duration

2.3.7.1 Results over all items

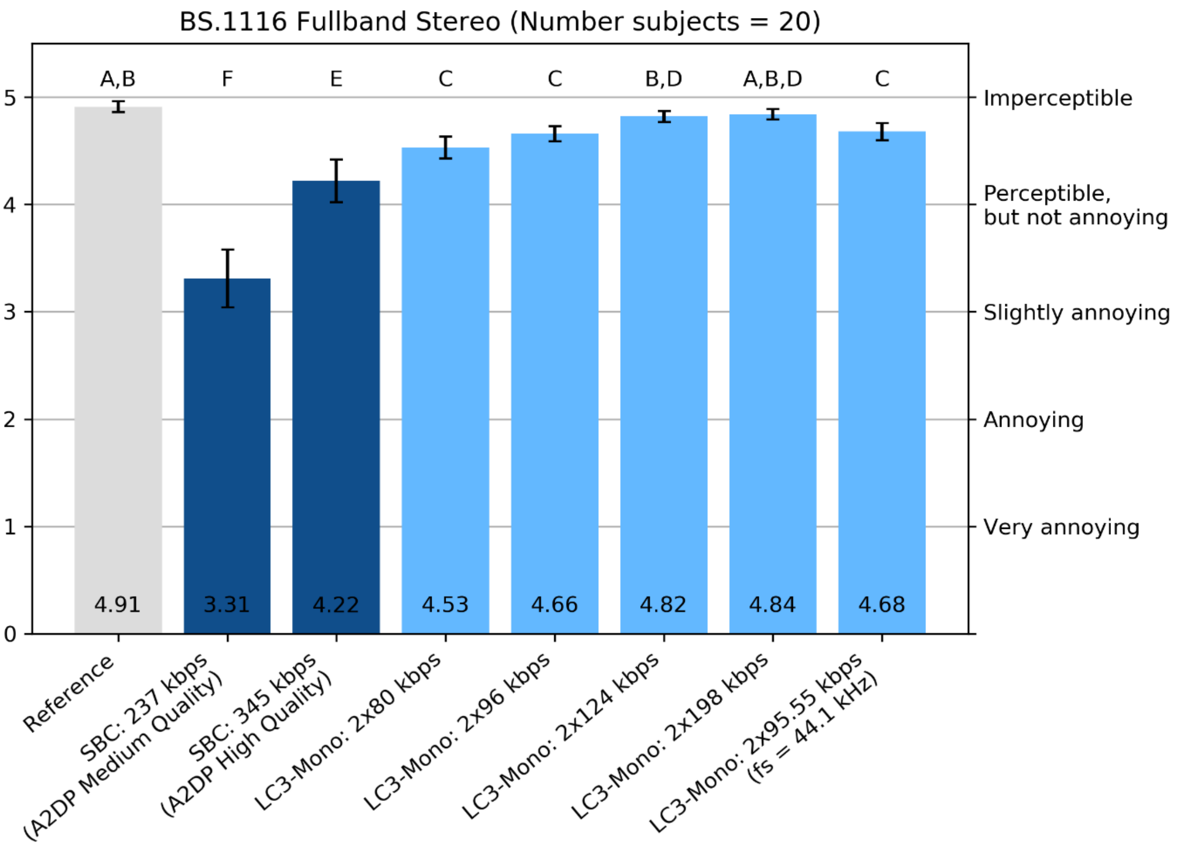


Figure 2.15: Mean scores, 95 percent confidence intervals, and statistical grouping (A-F) over all items in the BS.1116 stereo test for 7.5 ms frame duration. Statistical grouping means that conditions with the same letter are not significantly different from at least one of the other conditions in the group. LC3 operates in dual-mono with two times the bitrates for each channel.

From these results, we conclude the following:

- During this testing, LC3-Stereo (7.5 ms frame duration) at 2x124 kbps and 2x198 kbps and the uncoded original (Reference) performed equivalently.
- During this testing, LC3-Stereo (7.5 ms frame duration) at 2x80 kbps performed significantly better than SBC at 345 kbps.
- During this testing, LC3-Stereo (7.5 ms frame duration) at 2x95.55 kbps with a sample rate of 44.1 kHz performed equivalently to LC3-Stereo (7.5 ms frame duration) at 2x96 kbps with a sample rate of 48 kHz, which verified the performance of LC3 for 44.1 kHz signals under test conditions.
- Therefore, under test conditions, LC3 was able to provide a higher level of quality compared to SBC when operating at the same bitrate, or was able to reduce the bitrate by roughly 50 percent while providing similar quality.

2.3.7.2 Results for individual items

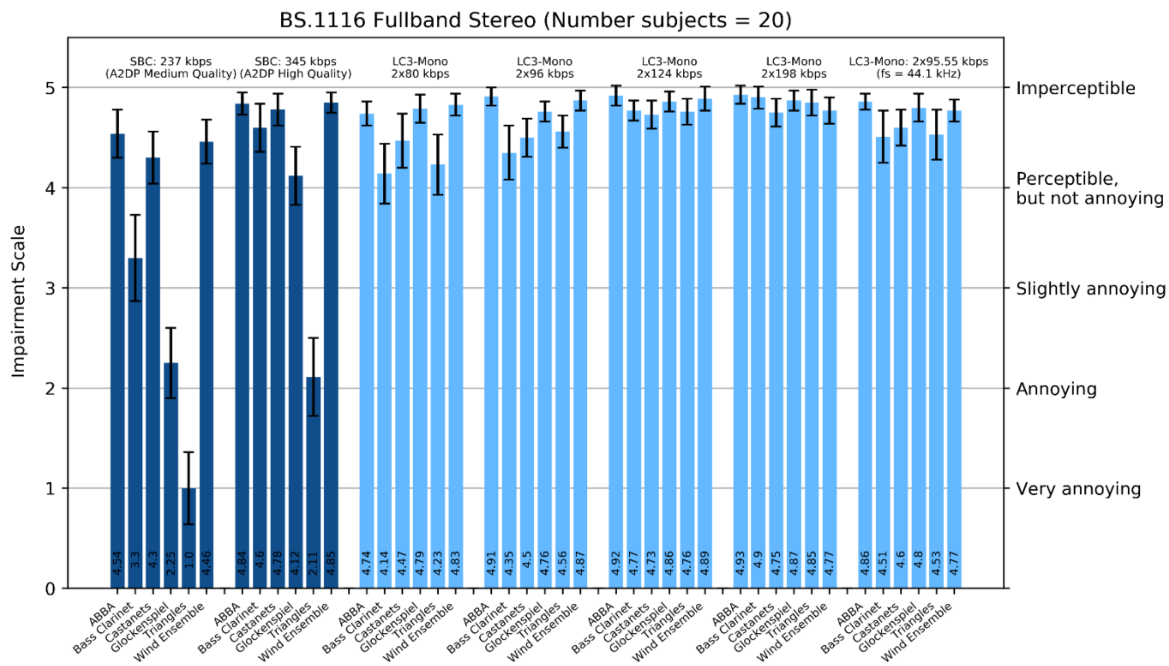


Figure 2.16: Mean ratings and 95 percent confidence intervals for the stereo test for 7.5 ms frame duration per sample

From these results, we conclude the following:

- Starting from 96 kbps per channel, all items coded with LC3-Stereo (7.5 ms frame duration) showed a rating above 4.0, which means there were no annoying distortions during this testing.
- For SBC, as specified in A2DP [4], at least two items always received a rating near 2.0, indicating an annoying quality level during this testing.

2.4 Hearing aid

2.4.1 Setup

The hearing aid use case was evaluated in the following two testing phases:

1. Phase I: MUSHRA online tests open for all Bluetooth members
2. Phase II: MUSHRA tests conducted in lab testing

The second testing phase was initiated because the 7.5 ms frame duration of LC3 was introduced later in the standardization process. Table 2.6 lists all four listening tests for the hearing aid use case.

Section	Description	Conditions	Number of Subjects
2.4.3	WB Voice I (16 kHz), Bluetooth member	• SBC: 66.4 kbps • G.722: 64 kbps • Opus: 32 kbps, 10 ms frame duration, complexity level 0 • LC3: 32 kbps, 10 ms frame duration	119

Section	Description	Conditions	Number of Subjects
2.4.4	WB Voice II (16 kHz), SenseLab	• SBC: 66.4 kbps • G.722: 64 kbps • LC3: 32 kbps, 10 ms frame duration • LC3: 32 kbps, 7.5 ms frame duration	20
2.4.5	HA Music I (24 kHz), Bluetooth members	• SBC: 96.4 kbps • G.722: 96 kbps • Opus: 48 kbps, 10 ms frame duration, complexity level 10 • LC3: 48 kbps, 10 ms frame duration	115
2.4.6	HA Music II (24 kHz), SenseLab	• SBC: 96.4 kbps • G.722: 96 kbps • LC3: 48 kbps, 10 ms frame duration • LC3: 48 kbps, 7.5 ms frame duration	20

Table 2.6: Summary of listening tests for the HA use case

2.4.2 Test items

Table 2.7 lists the items the HA working group selected for the WB hearing aid voice listening test. The material is mainly from the EBU test material [15].

Sample	Designation	Material ID	Description	Length (Seconds)
1	Female/male speech and car noise	49 50	Female speech, English, mixed with car background noise at 7 dB Male speech, English, mixed with car background noise at 7 dB	12 s
2	Music	57 other source	Classical Organ Pop artist	15 s
3	Music	61 66	Soprano Classical music	18 s
4	Female speech	49	Female speech, English	14 s

Table 2.7: Test material for phase I and II of WB voice tests

Table 2.8 lists the items the HA group selected for the HA music listening test. The material is mainly from the EBU test material [15].

Sample	Designation	Material ID	Description	Length (Seconds)
1	Music	57 other source	Classical Organ Suzanne Vega: Luca	15 s
2	Music	61 66	Soprano Classical music	18 s
3	Female speech	49	Female speech, English	14 s
4	Single instruments	27 32 41	Castanets / rhythm Triangles / roll Celesta / melodious phrase	11 s

Table 2.8: Test material for phase I and II of HA music tests

2.4.3 WB voice I

The HA working group evaluated the LC3 for generic audio content in comparison to G.722, Opus, and SBC for WB audio material. The audio content contained music and speech items according to Table 2.7. In the phase I MUSHRA test, 119 subjects participated via an online listening environment. Figure 2.17 shows the results of the WB voice I test.

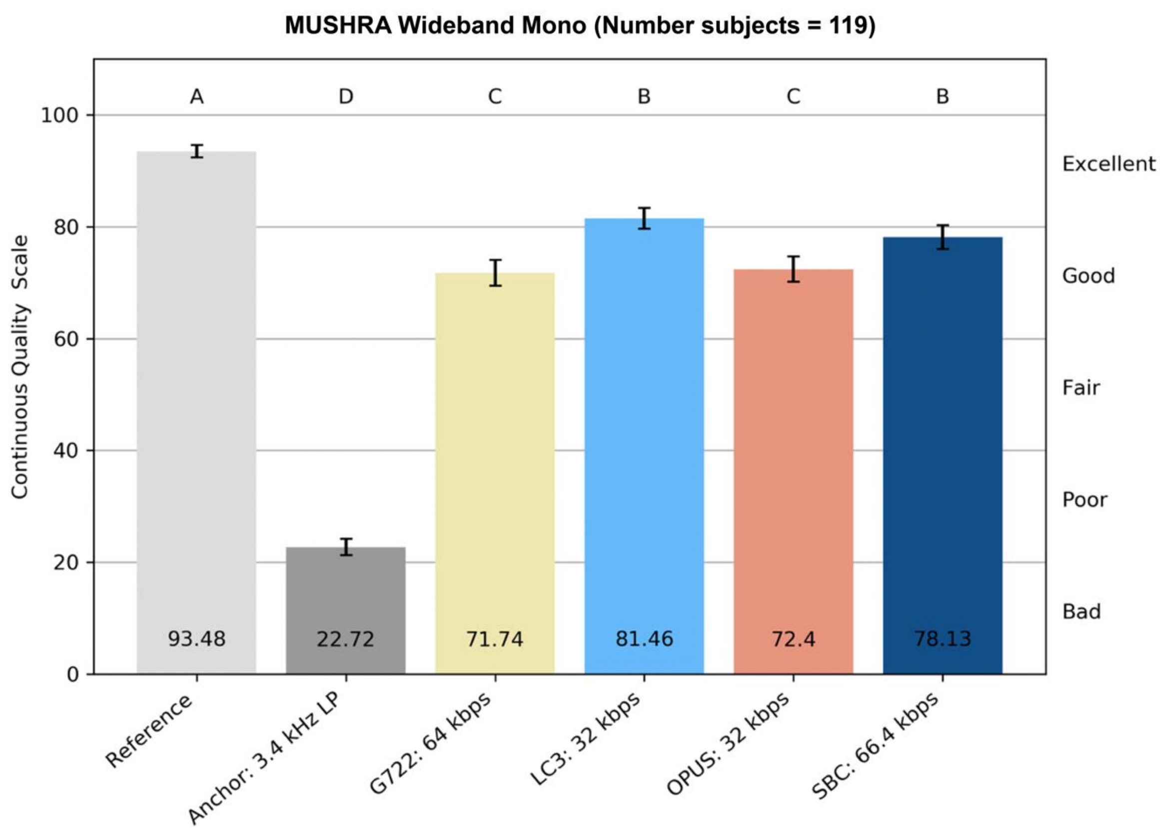


Figure 2.17: Mean scores, 95 percent confidence intervals, and statistical grouping (A-D) over all items of the HA voice I test. Statistical grouping means that conditions with the same letter show no statistically significant difference.

From these results, we conclude the following:

- During this testing, LC3 at 32 kbps offered the same quality as SBC at 66.4 kbps.
- During this testing, LC3 provided significantly better audio quality than G.722 at 64 kbps.

2.4.4 WB voice II

In the phase II MUSHRA test, the audio content from Table 2.7 was used to evaluate LC3 (7.5 ms frame duration) in comparison to LC3 (10 ms frame duration), G.722, and SBC. In the phase II test, 20 expert listeners participated. Figure 2.18 shows the results of the WB voice II test.

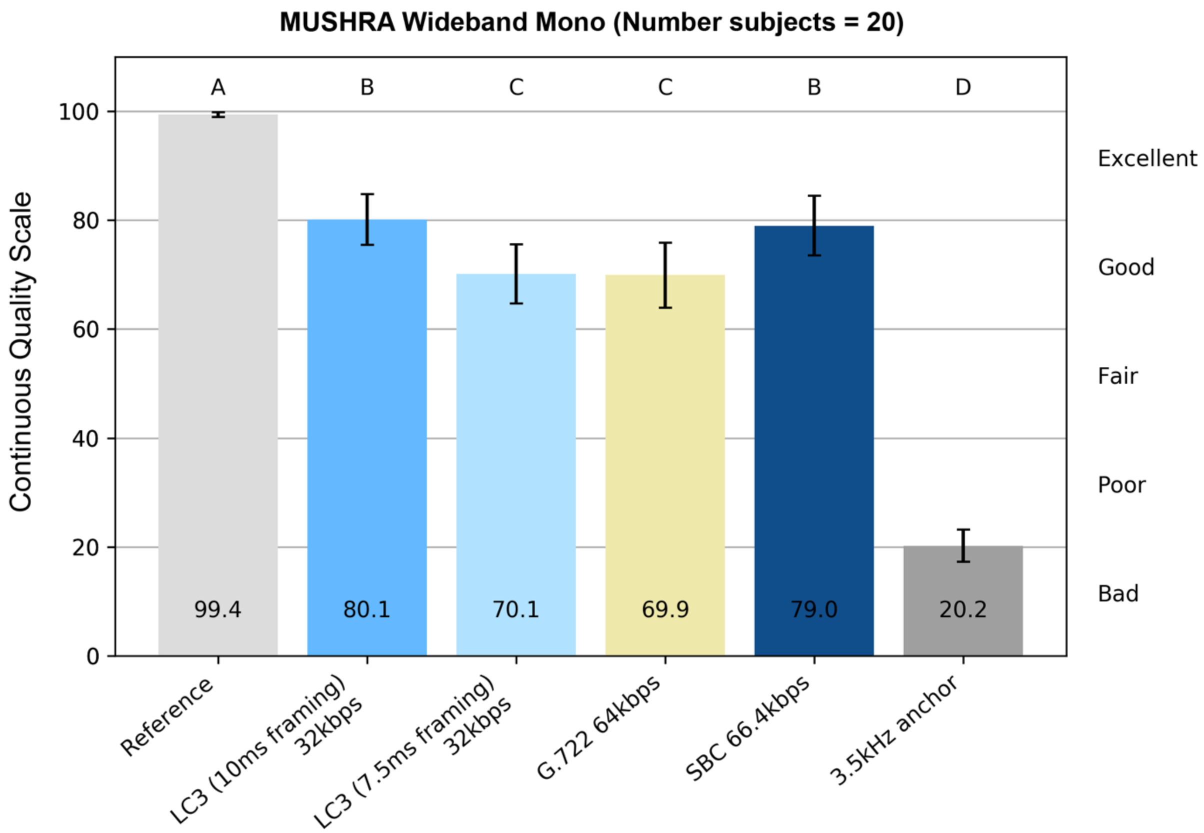


Figure 2.18: Mean scores, 95 percent confidence intervals, and statistical grouping (A-D) over all items of the WB voice II test. Statistical grouping means that conditions with the same letter show no statistically significant difference.

From these results, we conclude the following:

- During this testing, LC3 at 32 kbps and 10 ms frame duration showed no statistical difference from SBC at 66.4 kbps.
- During this testing, LC3 at 32 kbps and 7.5 ms frame duration showed no statistical difference from G.722 at 64 kbps.
- Over all items, the LC3 with 10 ms frame duration performed significantly better than LC3 at 7.5 ms frame duration under test conditions.

2.4.5 HA music I

The HA group evaluated LC3 for generic audio content in comparison to an oversampled G.722, Opus, and SBC for 24 kHz audio material. The audio content contained music and speech items according to

Table 2.8. In the phase I MUSHRA test, 115 subjects participated via an online listening environment. Figure 2.19 shows the results of the HA music I test.

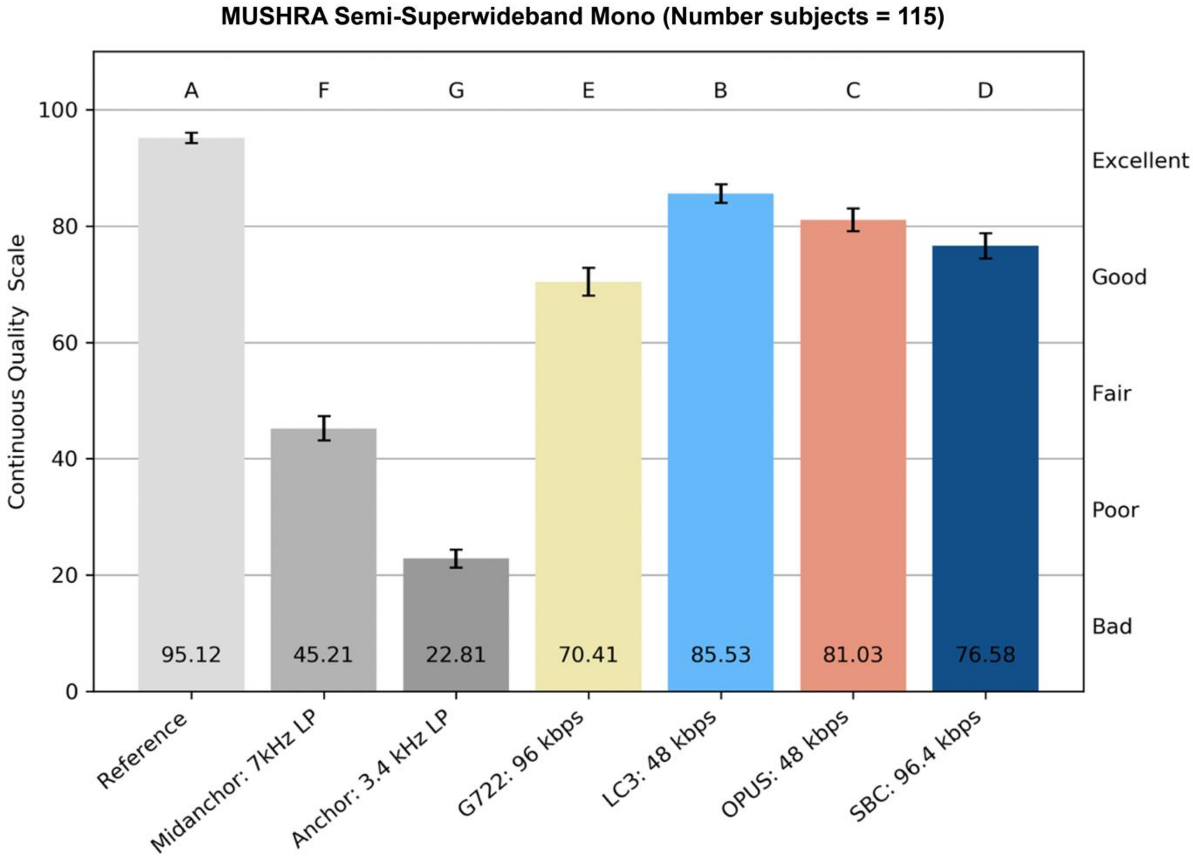


Figure 2.19: Mean scores, 95 percent confidence intervals, and statistical grouping (A-G) over all items of the HA music I test. Statistical grouping means that conditions with the same letter show no statistically significant difference.

During this testing, LC3 at 48 kbps performed significantly better than G.722 at 96 kbps and SBC at 96.4 kbps.

2.4.6 HA music II

In this phase II MUSHRA test, the audio content from Table 2.8 was used to evaluate LC3 (7.5 ms frame duration) in comparison to LC3 (10 ms frame duration), G.722, and SBC. In the phase II test, 20 expert listeners participated. Figure 2.20 shows the results of the HA music II test.

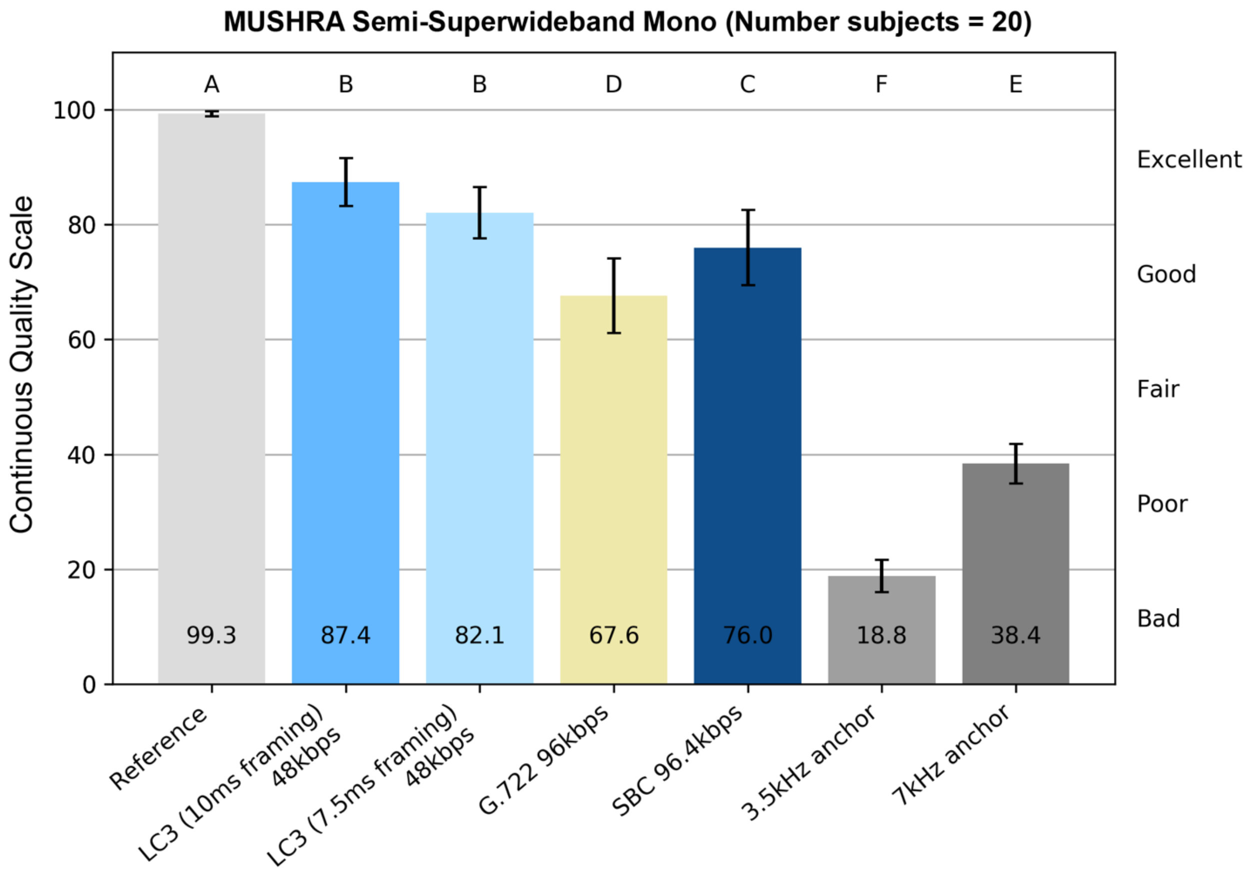


Figure 2.20: Mean scores, 95 percent confidence intervals, and statistical grouping (A-F) over all items of the HA music II test. Statistical grouping means that conditions with the same letter show no statistically significant difference.

From these results, we conclude the following:

- During this testing, LC3 at 48 kbps, operating at 10 ms and 7.5 ms frame duration, performed significantly better than G.722 at 96 kbps and SBC at 96.4 kbps.
- During this testing, LC3 operating at 10 ms frame duration showed no statistical difference from LC3 at 7.5 ms frame duration.

3 Conformance and testing

3.1 Conformance

The Bluetooth Special Interest Group (SIG) provides the LC3 Test Suite (TS) [1] which contains a complete reference script implementation of the conformance procedure and explains the test methodology. The package includes a fixed-point reference executable of the LC3 encoder and decoder. This implementation was used for the listening test evaluation outlined in Section 2.

For conformance tests, implementations under test are compared to the reference executable by metrics of Objective Difference Grade (ODG) and Root Mean Square (RMS). The conformance procedure is designed such that implementations of various platforms and precisions are able to pass the conformance criteria. The thresholds for the conformance criteria are based on a comparison between the fix-point reference executable and a floating-point implementation. The encoder and decoder are tested separately.

3.2 Sinusoidal test sequences

For conducting Total Harmonic Distortion (THD), Signal to Noise Ratio (SNR), or gain measurements on the LC3, some specific behaviors of LC3 need to be considered. For lower bitrates, LC3 uses a Long-Term Post Filter (LTPF) which possibly amplifies the signal, especially sinusoidal signals. Figure 3.1 shows the gain difference of input and decoded output signal of sinusoids at dedicated frequencies for the sampling rates 8 kHz, 16 kHz, 24 kHz, 32 kHz, 44.1 kHz, and 48 kHz.

As observed from Figure 3.1, the LTPF amplifies a full-scale sine signal up 0.5 dB. This effect is reduced by using a test frequency of 6,427 Hz, which usually results in a minimal gain difference of less than 0.03 dB, for testing the sampling rates 16 kHz, 24 kHz, 32 kHz, and 48 kHz. For the 8 kHz sampling rate, the optimal test frequency is 1,951 Hz, because of the up-sampling to the LTPF internal sampling rate of 12.8 kHz. For the 44.1 kHz sampling rate, the optimal test frequency is 5,905 Hz (which equals $6427 \times \frac{44100}{48000}$).

Gain is measured with the ITU-T P.56 RMS tool. In Figure 3.1, LC3 with 10 ms frame duration is shown in blue, and LC3 with 7.5 ms frame duration is shown in red. Note that for normal speech and music files, the average gain difference is typically less than 0.1 dB.

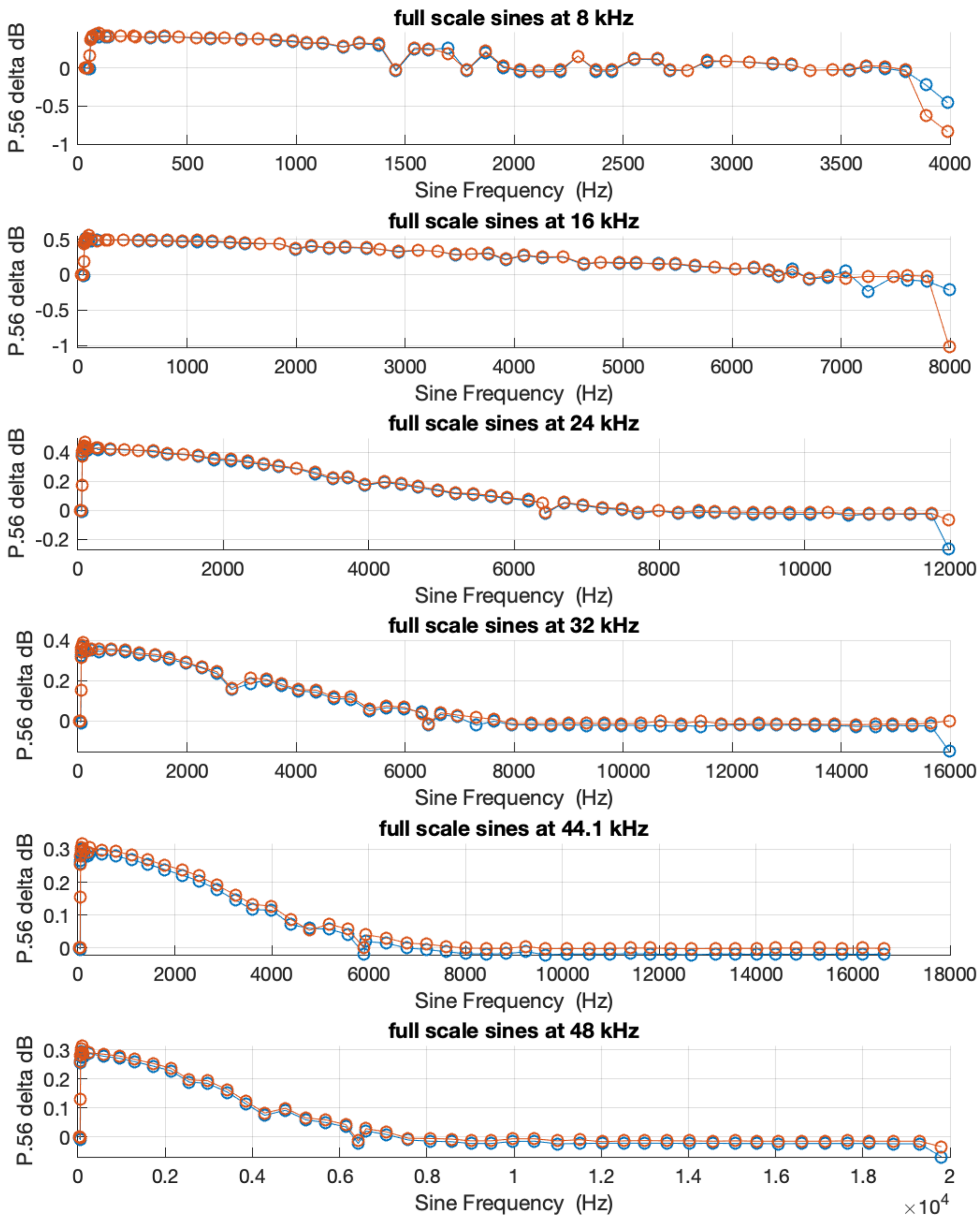


Figure 3.1: Gain difference of input signal and decoded output signal for sinusoids at dedicated frequencies

4 Summary and Conclusions

The extensive tests and studies presented in this document have shown that LC3 fulfills all codec requirements of LE Audio and exceeds most of these requirements.

The computational complexity of LC3 is significantly higher than SBC; however, because of the efficiency gain, LC3 enables a significant power reduction by LE Audio compared to Bluetooth Classic.

An analysis of the Bluetooth transmission using LC3 has shown that it is possible to achieve a total system delay close to 30 ms for audio transmission in a gaming-audio setup.

For voice services (see Section 2.2), LC3 introduces the super-wideband audio bandwidth into Bluetooth. Furthermore, with LC3 it is possible to reduce the bitrate by approximately 50 percent compared to the former mSBC codec, while maintaining the audio quality. Extensive transcoding tests have shown that LC3 interworks perfectly with the 3GPP mobile codecs AMR, AMR-WB, and EVS.

For music streaming (see Section 2.3), LC3 provides superior audio quality over the SBC codec, while operating at less than half the bitrate. Furthermore, starting from 2x96 kbps, any audio content coded by LC3 was rated higher than 4.0 MOS on a BS.1116 [16] rating scale during testing, meaning that not even slightly annoying artefacts were perceived by the user. LC3 enables a new level of audio quality compared to the SBC codec.

LC3 outperforms G.722 and SBC for the HA music streaming case tests (see Section 2.4.4). For the speech quality case tests (see Section 2.4.3), LC3 outperforms G.722 and shows the same level of quality as SBC while operating at approximately half the bitrate.

5 References

[1] Low Complexity Communication Codec Specification, Version 1.0 or later

[2] Hands-Free Profile Specification, Version 1.8 or later

[3] Telephony and Media Audio Profile Specification, Version 1.0 or later

[4] Advanced Audio Distribution Profile Specification, Version 1.4 or later

[5] Hearing Access Profile Specification, Version 1.0 or later

[6] Bluetooth Core Specification, Version 5.3 or later

[7] Basic Audio Profile Specification, Version 1.0.1 or later

[8] International Telecommunication Union (ITU), “Recommendation ITU-T G.722 - 7 kHz audio-coding within 64 kbit/s", September 2012, https://www.itu.int/rec/T-REC-G.722-201209-I/en

[9] Internet Engineering Task Force (IETF), “IETF Request for Comment (RFC) 6716 - Definition of the Opus Audio Codec", September 2012, https://tools.ietf.org/html/rfc6716

[10] ITU, “Recommendation ITU-R BS.1534-3 - Method for the subjective assessment of intermediate quality level of audio systems" aka MUSHRA, October 2015, https://www.itu.int/dms_pubrec/itu-r/rec/bs/R-REC-BS.1534-3-201510-I!!PDF-E.pdf

[11] ITU, “ITU-T P.800 - Methods for subjective determination of transmission quality", August 1996, https://www.itu.int/rec/dologin_pub.asp?lang=e&id=T-REC-P.800-199608-I!!PDF-E&type=items

[12] 3rd Generation Partnership Project (3GPP), “3GPP TS 26.071 - Mandatory speech CODEC speech processing functions; AMR speech Codec; General description”, July 2020, https://www.3gpp.org/ftp//Specs/archive/26_series/26.071/26071-g00.zip

[13] 3GPP, “3GPP TS 26.171 - Speech codec speech processing functions; Adaptive Multi-Rate - Wideband (AMR-WB) speech codec; General description, Release 16", July 2020, https://www.3gpp.org/ftp//Specs/archive/26_series/26.171/26171-g00.zip

[14] 3GPP, “3GPP TS 26.441 - Codec for Enhanced Voice Services (EVS); General overview, Release 16”, July 2020, https://www.3gpp.org/ftp//Specs/archive/26_series/26.441/26441-g00.zip

[15] European Broadcast Union (EBU), “EBU – TECH 3253 - Sound Quality Assessment Material recordings for subjective tests", September 2008, https://tech.ebu.ch/docs/tech/tech3253.pdf

[16] ITU, “Recommendation ITU-R BS.1116-3 - Methods for the subjective assessment of small impairments in audio systems", February 2015, https://www.itu.int/dms_pubrec/itu-r/rec/bs/R-REC-BS.1116-3-201502-I!!PDF-E.pdf

[17] ITU, “Recommendation ITU-R BS.1387-1 - Method for objective measurements of perceived audio quality" aka PEAQ, November 2001, https://www.itu.int/rec/R-REC-BS.1387

[18] Public Broadcast Profile Specification, Version 1.0 or later

[19] Gaming Audio Profile