ETSI TR 103 633 V1.1.1 (2022-12)

TECHNICAL REPORT

**Digital Enhanced Cordless Telecommunications (DECT);
Low Complexity Communication Codec Plus (LC3plus);
Performance characterization**

Reference

DTR/DECT-00339

Keywords

audio, codec, DECT, Full-Band, LC3plus, quality, superwideband, voice, VoIP

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - APE 7112B
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° w061004871

Contents

Intellectual Property Rights

Essential patents

IPRs essential or potentially essential to normative deliverables may have been declared to ETSI. The declarations pertaining to these essential IPRs, if any, are publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (https://ipr.etsi.org/).

Pursuant to the ETSI Directives including the ETSI IPR Policy, no investigation regarding the essentiality of IPRs, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Trademarks

The present document may include trademarks and/or tradenames which are asserted and/or registered by their owners. ETSI claims no ownership of these except for any which are indicated as being the property of ETSI, and conveys no right to use or reproduce any trademark and/or tradename. Mention of those trademarks in the present document does not constitute an endorsement by ETSI of products, services or organizations associated with those trademarks.

DECT™, **PLUGTESTS™**, **UMTS™** and the ETSI logo are trademarks of ETSI registered for the benefit of its Members. **3GPP™** and **LTE™** are trademarks of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners. **oneM2M™** logo is a trademark of ETSI registered for the benefit of its Members and of the oneM2M Partners. **GSM**® and the GSM logo are trademarks registered and owned by the GSM Association.

Foreword

This Technical Report (TR) has been produced by ETSI Technical Committee Digital Enhanced Cordless Telecommunications (DECT).

Modal verbs terminology

In the present document "**should**", "**should not**", "**may**", "**need not**", "**will**", "**will not**", "**can**" and "**cannot**" are to be interpreted as described in clause 3.2 of the ETSI Drafting Rules (Verbal forms for the expression of provisions).

"**must**" and "**must not**" are **NOT** allowed in ETSI deliverables except when used in direct citation.

Introduction

The Low Complexity Communication Codec Plus (LC3plus) is standardized in ETSI TS 103 634 [i.4] and integrated in DECT specifications as voice and audio codec. The ETSI Technical Committee "Speech and multimedia Transmission Quality (STQ)" group derived quality metrics and reference conditions to ensure the codec's performance for the application DECT and VoIP. TC STQ designed a test plan [i.2] which was executed within the Testing Test Force (TTF) 005 and the results were statistically analysed.

As conclusion, TC STQ recommends LC3plus as the ETSI codec for global deployment in DECT and VoIP applications. The results from TTF 005 and further inputs are the basis for the present document which aims to characterize LC3plus regarding different codec aspects.

1 Scope

The present document characterizes the Low Complexity Communication Codec Plus (LC3plus) codec [i.4] by using subjective and objective test methodologies as presented in ETSI TS 103 624 [i.2]. The resulting measurements are presented in detail in order to point out the performance of LC3plus in certain use cases such as voice services over DECT and VoIP or music streaming. Other aspects of the codecs such as complexity and memory requirements are discussed as well.

2 References

2.1 Normative references

Normative references are not applicable in the present document.

2.2 Informative references

References are either specific (identified by date of publication and/or edition number or version number) or non-specific. For specific references, only the cited version applies. For non-specific references, the latest version of the referenced document (including any amendments) applies.

NOTE: While any hyperlinks included in this clause were valid at the time of publication, ETSI cannot guarantee their long term validity.

The following referenced documents are not necessary for the application of the present document but they assist the user with regard to a particular subject area.

[i.1] Recommendation ITU-T P.800 (08/1996): "Methods for subjective determination of transmission quality".

[i.2] ETSI TS 103 624 (V1.2.1): "Characterization Methodology and Requirement Specifications for the ETSI LC3plus speech codec".

[i.3] Recommendation ITU-T G.726 (12/1990): "40, 32, 24, 16 kbit/s Adaptive Differential Pulse Code Modulation (ADPCM)".

[i.4] ETSI TS 103 634: "Digital Enhanced Cordless Telecommunications (DECT); Low Complexity Communication Codec plus (LC3plus)".

[i.5] ETSI TS 126 173: "Digital cellular telecommunications system (Phase 2+) (GSM); Universal Mobile Telecommunications System (UMTS); LTE; ANSI-C code for the Adaptive Multi-Rate - Wideband (AMR-WB) speech codec (3GPP TS 26.173)".

[i.6] IETF RFC 6716: "Definition of the Opus Audio Codec".

[i.7] Recommendation ITU-T P.800.1 (07/2006): "Methods for objective and subjective assessment of quality".

[i.8] ETSI EN 300 175-8 (V2.9.1): "Digital Enhanced Cordless Telecommunications (DECT); Common Interface (CI); Part 8: Speech and audio coding and transmission".

[i.9] R. Geiger et al: "Enhanced Mpeg-4 Low Delay AAC - Low Bitrate High Quality Communication", 122nd AES Convention, Paper 6998" (2007 May).

[i.10] Recommendation ITU-T G.711 Appendix I (09/1999): "A high quality low-complexity algorithm for packet loss concealment with G.711".

[i.11] Recommendation ITU-T G.722 Appendix IV (11/2006): "A low-complexity algorithm for packet loss concealment with G.722".

[i.12] ETSI TS 126 441: "Universal Mobile Telecommunications System (UMTS); LTE; 5G; Codec for Enhanced Voice Services (EVS); General overview (3GPP TS 26.441)".

[i.13] ETSI TS 126 073: "Digital cellular telecommunications system (Phase 2+) (GSM); Universal Mobile Telecommunications System (UMTS); LTE; 5G; ANSI-C code for the Adaptive Multi Rate (AMR) speech codec (3GPP TS 26.073)".

[i.14] Recommendation ITU-R BS.1387: "Method for objective measurements of perceived audio quality".

[i.15] Recommendation ITU-R BS.1116: "Methods for the subjective assessment of small impairments in audio systems".

[i.16] Recommendation ITU-T G.191: "Software tools for speech and audio coding standardization".

3 Definition of terms, symbols and abbreviations

3.1 Terms

For the purposes of the present document, the following terms apply:

DIRECT: uncoded original audio signal as reference in subjective experiments

gross rate: total bitrate consisting of source coder rate and forward error correction rate

self-tandeming: several consecutive encoder and decoder operations using one specific codec

3.2 Symbols

For the purposes of the present document, the following symbols apply:

X => Y Transcoding from codec X to Y

3.3 Abbreviations

For the purposes of the present document, the following abbreviations apply:

ACR Absolute Category Rating
AMR-NB Adaptive MultiRate speech codec - Narrow Band
AMR-WB Adaptive MultiRate speech codec - Wide Band
CuT Codec under Test
DCR Degradation Category Rating
DECT Digital Enhanced Cordless Telecommunications
DMOS Degradation Mean Opinion Score
DP DECT Profile

NOTE: E.g. DP0, DP1, etc.

DSP Digital Signal Processor
EPFsize Error Protection File with dedicated loss duration
EVS codec for Enhanced Voice Services
EVS-WB EVS - WideBand
FB FullBand
FER Frame Error Rate
LC3plus Low Complexity Communication Codec Plus
MOS Mean Opinion Score
MOS-LQS Mean Opinion Score - Listening Quality (Subjective)
MPEG Motion Picture Experts Group
NB NarrowBand

ODG	Objective Difference Grade
PEAQ	Perceptual Evaluation of Audio Quality
PLC	Packet Loss Concealment
PLP	Packet Loss Profile

NOTE: E.g. PLP0, PLP1, etc.

PLR	Packet Loss Rate
SDG	Subjective Difference Grade
SNR	Signal to Noise Ratio
STL	Software Tools Library
SWB	Super WideBand
THD+N	Total Harmonic Distortion and Noise
TTF	Testing Task Force
VoIP	Voice over IP
WB	WideBand
WMOPS	Weighted Million Operations Per Second

4 General

The Low Complexity Communication Codec Plus (LC3plus) is standardized in ETSI TS 103 634 [i 4] with the goal to bring the audio quality for wireless audio connections to the next level. This includes the quality of voice calls as well as music streaming applications.

For voice call applications, e.g. DECT and VoIP, LC3plus is designed to achieve the following objectives:

- Introduction of Super-Wideband (SWB) quality in voice services
- Increased capacity of DECT systems when compared to legacy DECT codecs
- Improved robustness for packet loss and bit errors
- Ensure suitable performance in case of transcoding or self-tandeming conditions

For music streaming applications, LC3plus is designed to achieve the following objectives:

- Scaling up to excellent or transparent music quality
- Transmission of High-Resolution audio content

Besides the audio quality aspects, LC3plus is designed to operate with:

- Low latency
- Low computational complexity
- Low memory footprint

5 Terms of Reference

5.1 Voice services

For Voice services, the reference conditions were chosen to allow direct comparison of the LC3plus to the legacy DECT and VoIP codecs depending on the used audio bandwidth.

Table 1: Reference codecs per application

Bandwidth	Application	
	DECT	VoIP
NB	Recommendation ITU-T G.726 [i.3], IETF Opus [i.6] (see note 1)	Recommendation ITU-T G.711 [i.10], IETF Opus [i.6] (see note 1)
WB	Recommendation ITU-T G.722 [i.11], IETF Opus [i.6] (see note 1)	Recommendation ITU-T G.722 [i.11], IETF Opus [i.6] (see note 1)
SWB	3GPP EVS [i.12], Recommendation ITU-T G.722 [i.11], IETF Opus [i.6] (see notes 1 and 2)	3GPP EVS [i.12], Recommendation ITU-T G.722 [i.11], IETF Opus [i.6] (see notes 1 and 2)
NOTE 1: Opus configuration with 10 ms frame duration, complexity level 0, constant bitrate and restricted low delay mode. NOTE 2: Opus operated on 48 kHz internal sampling rate.		

The chosen bitrates for LC3plus were mainly motivated to fit DECT normal and long slots, i.e. 32 kbps for NB and WB and 64 kbps for SWB.

5.2 Music services

The quality of high-resolution audio is usually determined by metrics such as THD+N or SNR. Therefore, extreme low distortion levels of less than -120 dB THD+N and more than 120 dB SNR at higher bit rates are envisioned.

For medium bitrates, the codec should scale to perceptually transparent audio quality measured in terms of Objective Difference Grades (ODG).

6 Introduction to Testing

6.1 Subjective quality tests

All subjective experiments were conducted using the Recommendation ITU-T P.800 [i.1] procedure using clean speech material. Subjects were naïve listeners and native speakers. MESAQIN.com performed the experiments according to the procedures and test plan specified in ETSI TS 103 624 [i.2]. All subjective quality tests were under the responsibility of the ETSI TC STQ group.

All experiments were conducted in English language. Each bandwidth (NB, WB and SWB) as well as clean and error-prone channels were tested separately. The experiments combined conditions for DECT and VoIP scenarios.

6.2 Objective Measurements

Objective measures have been used to estimate the following aspects:

- The music streaming performance of LC3plus has been estimated with the help of Recommendation ITU-R BS.1387 [i.14] (PEAQ).
- The capability of LC3plus to transmit high-resolution content has been assessed by the distortion metrics Total Harmonic Distortion plus Noise (THD+N) and Signal to Noise Ratio (SNR).

- The complexity of LC3plus in fixed-point arithmetic has been analysed with the help of the Recommendation ITU-T G.191 [i.16] software tools.

7 Subjective quality tests

7.1 Introduction to use cases

The following clauses characterize the performance quality for the following use cases:

- DECT with clean channel conditions.
- DECT with error prone channel conditions.
- VoIP without packet loss conditions.
- VoIP including packet loss conditions.

Information about the preparation of test files, used codec and software versions, speech material, etc. can be found in the characterization test plan [i.2], Annex B.

7.2 Charts

All figures showing listening test results in the present document are shown as bar or line charts and depict the arithmetic mean score together with 95 % confidence intervals per tested condition. For NB, WB and SWB error-prone channel experiments mean opinion score for subjective listening quality (MOS-LQS) [i.7] is used as measure with the opinion scale ranging from 1 (= bad) to 5 (= excellent) plotted on the y axis. For SWB error-free channel experiments Degradation Mean Opinion Score (DMOS) [i.7] is used as measure and rated according to the five-point degradation category scale from 1 (Degradation is very annoying) to 5 (Degradation is inaudible).

The MOS values for each condition are calculated from 96 different data points and rated by 24 different naïve and native listeners. The results of the statistical tests comparing the relevant conditions can be found in Annex A.

The printed scores above the datapoints are rounded to one decimal place, whereby the actual datapoint is more precise.

7.3 DECT scenarios with error-free channels

7.3.1 Overview

This clause demonstrates that LC3plus in DECT provides the same or better voice quality and may provide higher efficiency than the DECT legacy audio codecs. This is also true when DECT interoperates with legacy VoIP networks, where the following transcoding scenarios are evaluated:

- Voice calls from legacy VoIP to DECT.
- Voice calls from DECT to legacy VoIP.
- Voice calls from DECT over legacy VoIP to DECT.

The DECT legacy codecs are G.726 [i.3] (NB) and G.722 [i.11] (WB). Legacy VoIP terminals utilize G.711 [i.10] (NB) and G.722 [i.11] (WB).

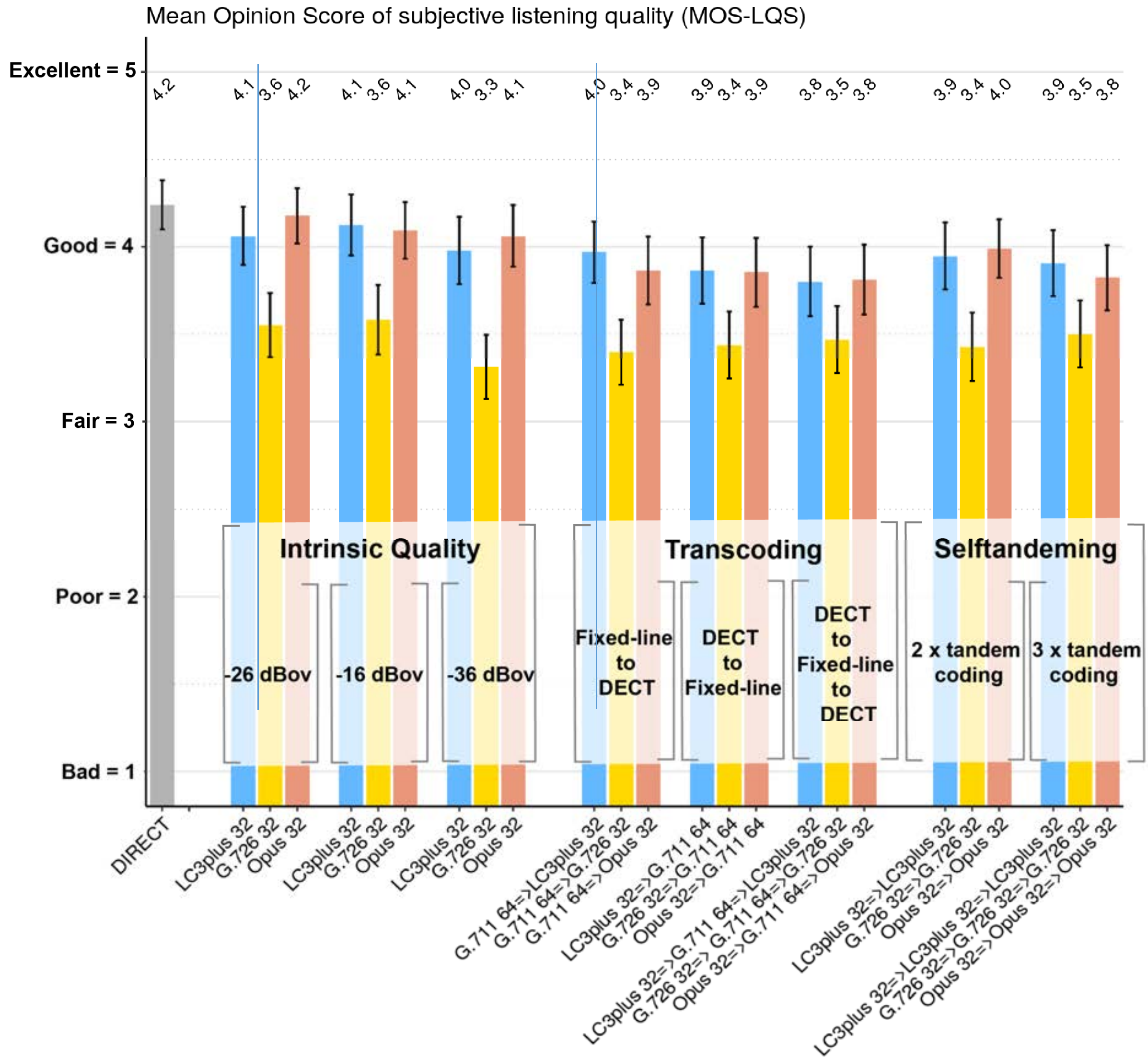
Under the assumption that DECT and VoIP are using identical codecs, the transcoding case becomes an asynchronous self-tandeming case where the same coding process is repeated twice or three times.

As additional performance objective, the Opus codec [i.6] is added to all tests to complete the performance picture.

7.3.2 NB conditions clean speech

To verify the performance of LC3plus in a NB call over DECT using normal slots, an ACR experiment (as per Recommendation ITU-T P.800 [i.1]) has been conducted to compare LC3plus to the legacy DECT codec G.726 [i.3] at 32 kbps and to Opus at 32 kbps [i.6].

For the transcoding cases, G.711 [i.10] at 64 kbps represents the legacy VoIP NB codec.



NOTE: Figure 1 shows mean scores (n = 96) and 95 % confidence intervals of 24 subjects.

Figure 1: P.800 ACR results for DECT use cases with NB clean speech signals

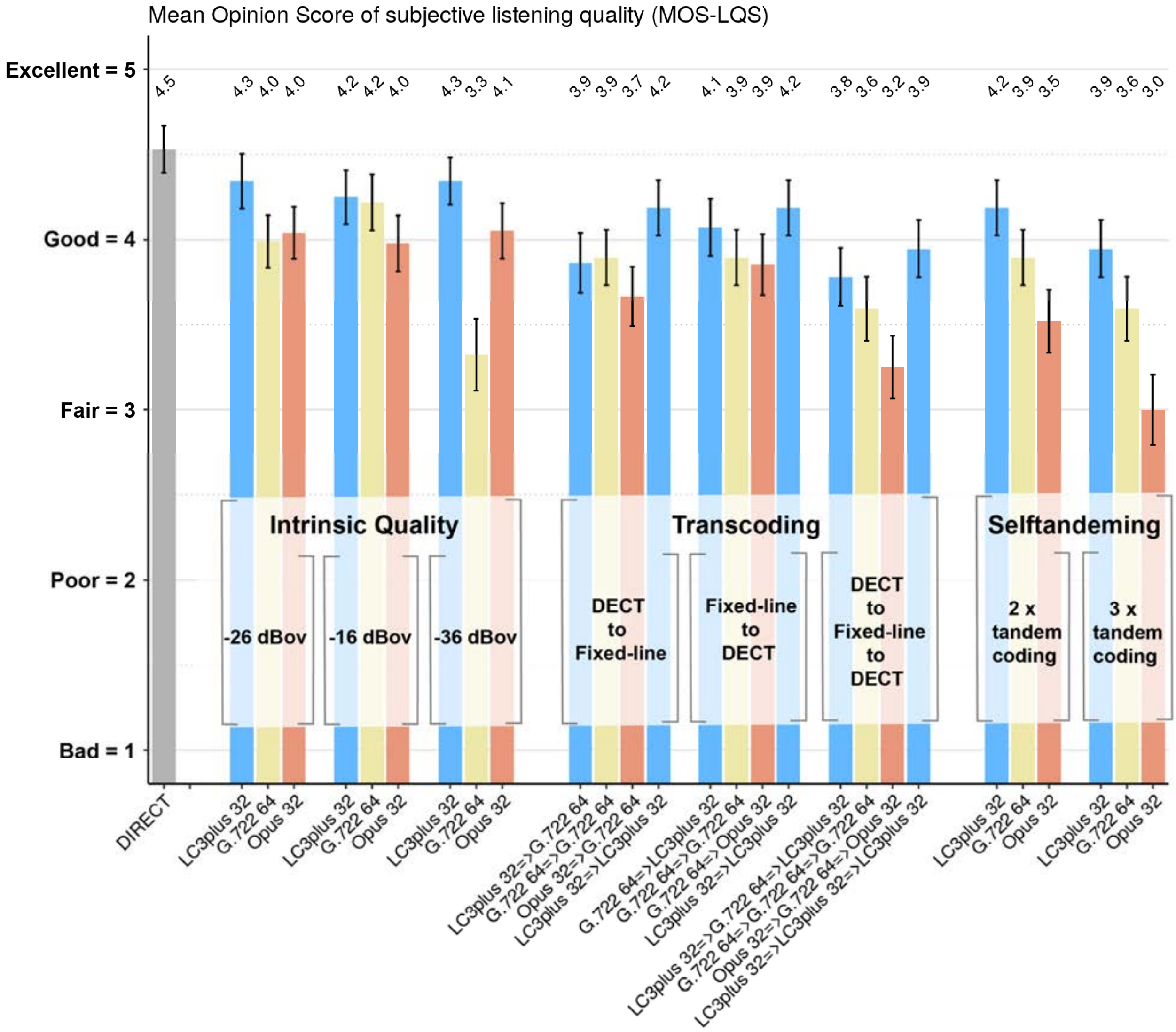
The following conclusions can be made for DECT NB systems:

- The test demonstrates very clearly that LC3plus is significantly better than the DECT legacy codec G.726. This holds true in direct comparison at different amplitude levels as well as for transcoding connections between VoIP-G.711 and DECT.
- The additional performance objectives in comparison to Opus have been achieved since there is no statistically significant difference between LC3plus and Opus conditions.

7.3.3 WB conditions clean speech

To verify the performance of LC3plus at 32 kbps in a WB call over DECT using normal slots an ACR experiment (as per Recommendation ITU-T P.800 [i.1]) has been conducted to compare LC3plus to the legacy DECT codec G.722 [i.11] at 64 kbps (long DECT slot) and to Opus [i.6] at 32 kbps.

For the transcoding cases, G.722 [i.11] at 64 kbps represents the legacy VoIP WB codec.



NOTE: Figure 2 shows mean scores (n = 96) and 95 % confidence intervals of 24 subjects.

Figure 2: P.800 ACR results for DECT use cases with WB clean speech signals

The following conclusions can be made for DECT WB systems:

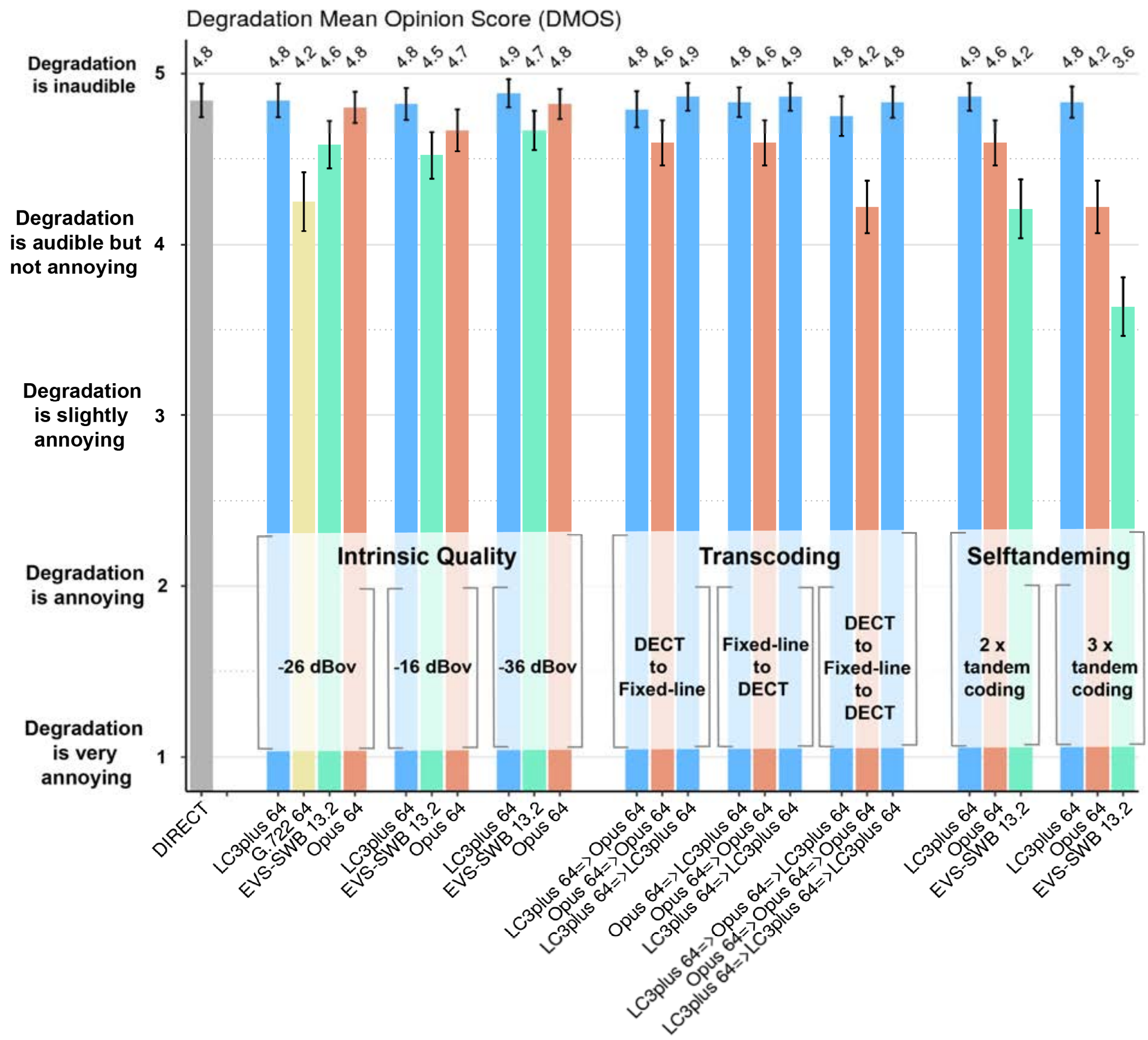
- LC3plus at 32 kbps provides clearly better quality than G.722 at 64 kbps and Opus at 32 kbps for all input levels, except for high input levels where G.722 is on par.
- LC3plus doubles the DECT capacity for WB calls as the codec operated in normal slots (32 kbps) while the legacy G.722 requires long slots (64 kbps).
- Regarding transcoding, LC3plus shows always a higher MOS compared to G.722 and Opus except for the DECT to Fix-line case where LC3plus at 32 kbps is on par with G.722 at 64 kbps.
- LC3plus is significantly more robust for asynchronous self-tandeming conditions than G.722 and Opus.

- LC3plus when operating in DECT as well as VoIP provides significant better quality compared to the legacy G.722, Opus or any other codec combination for DECT and VoIP.

7.3.4 SWB conditions clean speech

In contrast to the characterization of previous use cases, a DCR experiment (as per Recommendation ITU-T P.800 [i.1]) has been conducted instead of the ACR procedure. Since SWB with clean speech is expected to have high quality in all conditions, the Degradation Category Rating method is chosen because it affords higher sensitivity. DCR uses an annoyance scale and presents a quality reference before each condition to be rated by the listener.

The DCR experiment has been conducted to compare LC3plus at 64 kbps to EVS [i.12] at 13,2 kbps and Opus [i.6] at 64 kbps. For the transcoding cases, Opus [i.6] at 64 kbps is assumed as legacy VoIP SWB codec.



NOTE: Figure 3 shows mean scores (n = 96) and 95 % confidence intervals of 24 subjects.

Figure 3: P.800 DCR results for DECT use cases with SWB clean speech signals

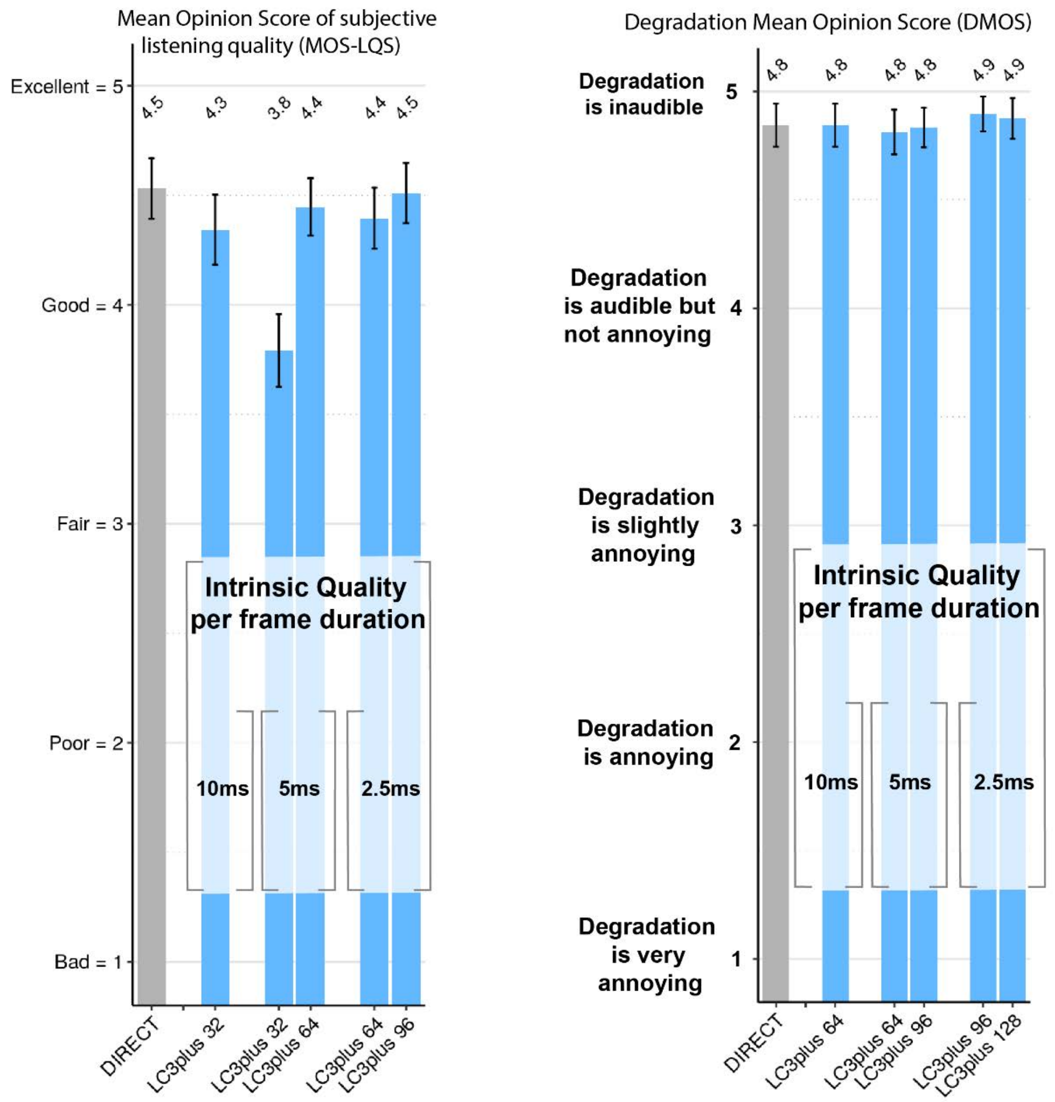
The following conclusions can be made for DECT SWB systems:

- LC3plus at 64 kbps outperforms EVS-SWB at 13,2 kbps significantly on all input levels in intrinsic quality.
- LC3plus at 64 kbps is significantly better than G.722 at 64 kbps at -26 dB input level.
- LC3plus shows the same or better quality level compared to Opus at 64 kbps for intrinsic quality.

- In all transcoding scenarios, LC3plus is significantly better than Opus.
- Even after triple asynchronous transcoding LC3plus maintains the intrinsic quality while EVS-SWB and Opus clearly decrease with every additional coding iteration.

7.3.5 Low Delay modes for clean speech

The experiments in clauses 7.3.2 and 7.3.3 contained LC3plus conditions with 5 ms and 2,5 ms frame durations. Those conditions verify the LC3plus performance for specific low delay use case, e.g. gaming headsets, in the DECT environment. Note that a normal DECT slot provides a maximum slot rate of 64 kbps when operating at 5 ms transmission interval and 128 kbps when operating at 2,5 ms transmission interval.



NOTE: Figure 4 shows mean scores (n = 96) and 95 % confidence intervals of 24 subjects.

Figure 4: P.800 ACR WB (left) and DCR SWB (right) results for LC3plus with different frame durations for clean speech signals

The following conclusions can be made for DECT systems:

- For 5 ms DECT slot interval, LC3plus at 64 kbps transmitted within a DECT normal slot shows no statistical degradation compared the uncoded original (DIRECT). This is true for WB and SWB speech signals.
- For 2,5 ms DECT slot interval, LC3plus at 96 kbps transmitted within a DECT normal slot shows no statistical degradation compared the uncoded original (DIRECT). This is true for WB and SWB speech signals. As LC3plus only requires 75 % of the DECT slot size, additional of channel coder redundancy can help increasing the robustness.
- The plot shows that LC3plus transmitted over a DECT normal slot always provides excellent audio quality.

7.4 DECT scenarios with error prone channels

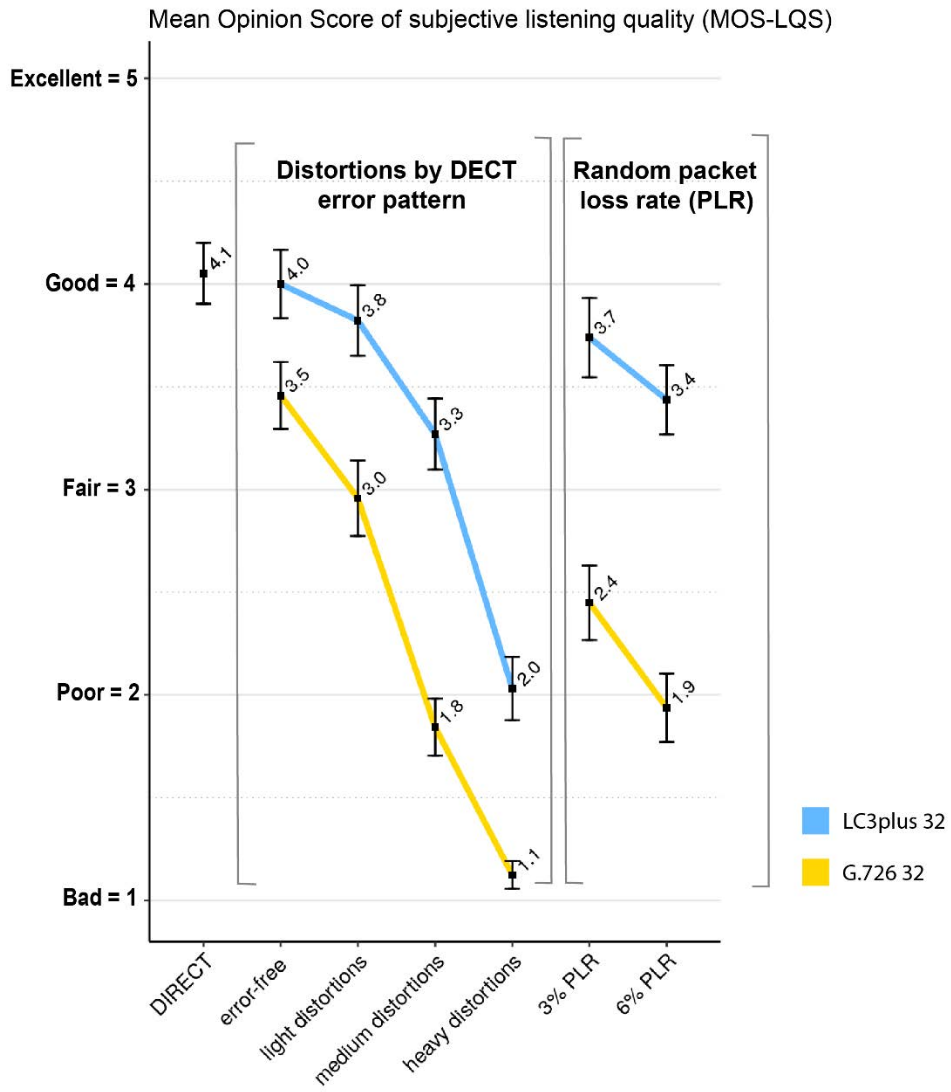
7.4.1 Overview

Two kinds of error profiles have been evaluated. The DECT error patterns are based on real measurements under different channel conditions, i.e. error-free, light distortions, medium distortions, and heavy distortion, which correspond to the definition of Table 4 in ETSI TS 103 624 [i.2]. Those error patterns consist of bit errors and packet losses. LC3plus operates with enabled forward error correction where the protection strength is adapted to the DECT channel condition.

Additionally, random packet loss patterns have been evaluated with Packet Loss Rates (PLR) of 3 % and 6 %.

7.4.2 NB conditions error prone speech

For DECT NB systems, LC3plus at 32 kbps gross rate is compared to G.726 [i.3] at 32 kbps with G.711 PLC as defined in Appendix I of Recommendation ITU-T G.711 [i.10].



NOTE: Figure 5 shows mean scores (n = 96) and 95 % confidence intervals of 24 subjects.

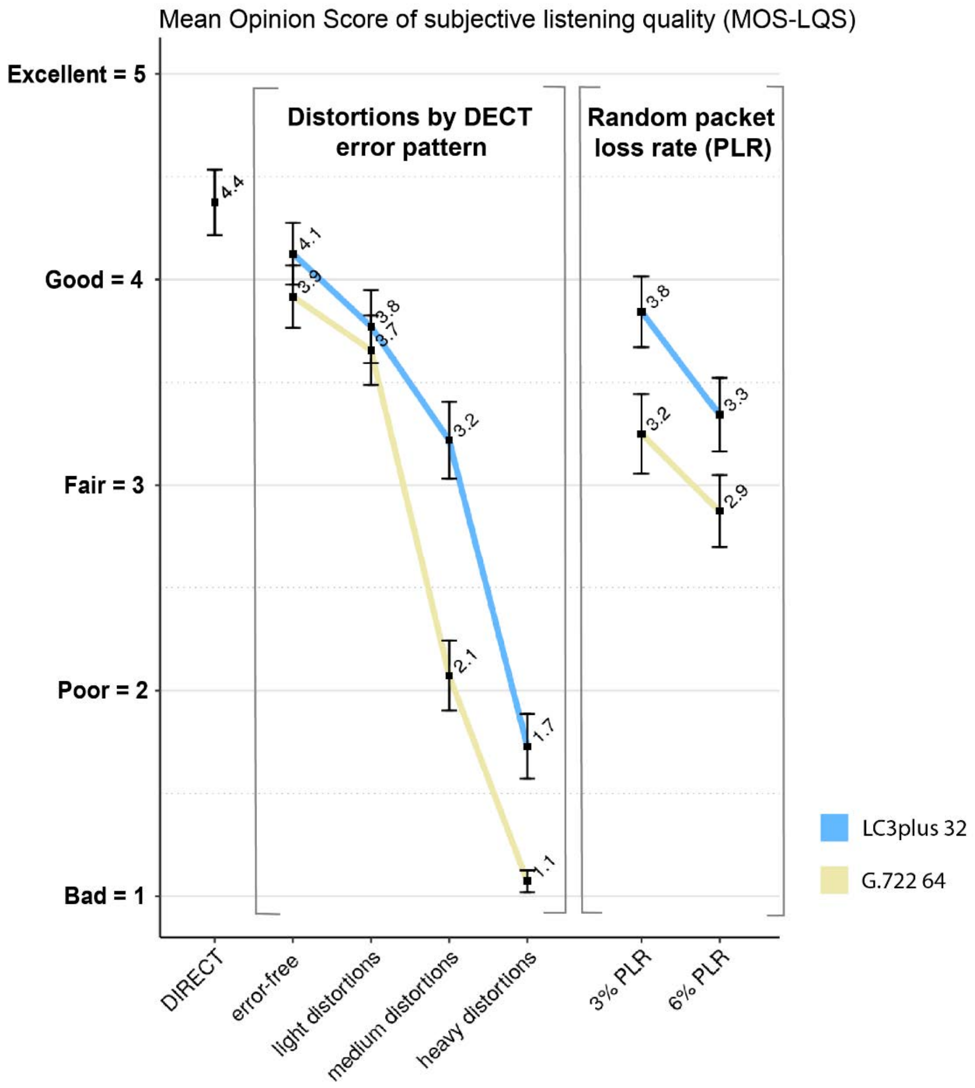
Figure 5: P.800 ACR results for DECT use cases with NB error prone speech signals

The following conclusions can be drawn:

- At all error conditions, LC3plus provides significantly better audio quality than G.726.
- At medium distortions, LC3plus shows a 1,5 higher MOS compared to G.726.
- LC3plus at medium distortions provides similar quality compared to G.726 without distortions.
- LC3plus clearly improves the robustness compared to a NB legacy DECT system.

7.4.3 WB conditions error prone speech

For DECT WB systems, LC3plus at 32 kbps gross rate is compared to G.722 at 64 kbps with PLC as defined in Appendix IV of Recommendation ITU-T G.722 [i.11].



NOTE: Figure 6 shows mean scores (n = 96) and 95 % confidence intervals of 24 subjects.

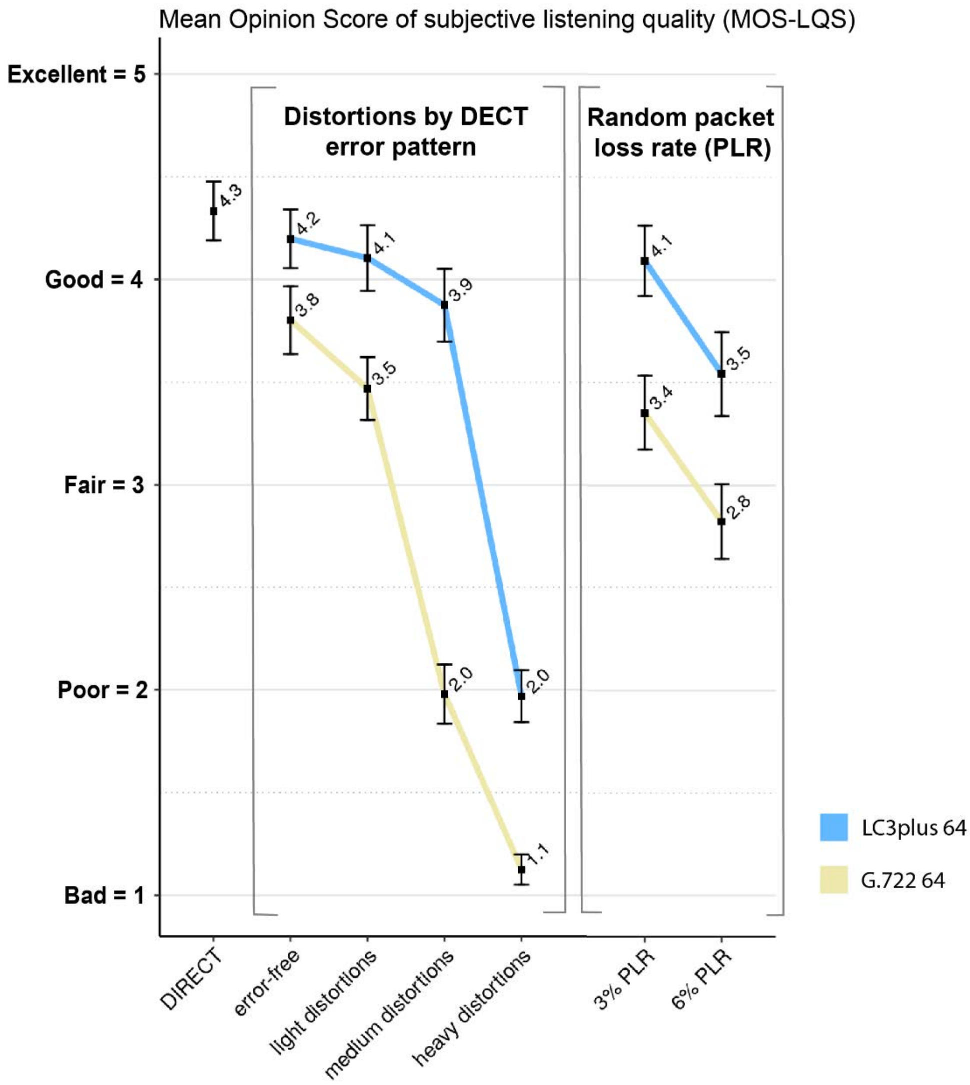
Figure 6: P.800 ACR results for DECT use cases with WB error prone speech signals

The following conclusions can be drawn:

- At all error conditions, LC3plus provides better audio quality than G.722.
- At medium distortions, LC3plus shows a 1,1 higher MOS compared to G.722.
- LC3plus at 6 % PLR is as good as G.722 at 3 % PLR.
- LC3plus clearly improves the robustness compared to a WB legacy DECT system while doubling the capacity.

7.4.4 SWB conditions error prone speech

For DECT SWB systems, LC3plus at 64 kbps gross rate is compared to G.722 at 64 kbps with PLC as defined in Appendix IV of Recommendation ITU-T G.722 [i.11]. Note that G.722 operates in WB to reflect the legacy DECT system.



NOTE: Figure 7 shows mean scores (n = 96) and 95 % confidence intervals of 24 subjects.

Figure 7: P.800 ACR results for DECT use cases with SWB error prone speech signals

The following conclusions can be drawn:

- At all error conditions, LC3plus (SWB) provides significantly better audio quality than G.722 (WB).
- At medium distortions, LC3plus (SWB) shows a 1,9 higher MOS compared to G.722 (WB).
- At medium distortions, LC3plus (SWB) provides similar audio quality compared to G.722 (WB) without distortions. LC3plus, by enabling SWB in DECT, clearly provides a quality boost compared to a legacy WB DECT system.
- LC3plus clearly improves the robustness compared to a WB legacy DECT system.
- LC3plus (SWB) at 3 % FER provides higher average MOS score than G.722 (WB) at 0 % FER.

7.5 VoIP scenarios with error-free channels

7.5.1 Overview

This clause demonstrates that LC3plus used in VoIP systems provides the same or better voice quality and may provide higher efficiency than the VoIP legacy audio codecs. This is also true when VoIP interoperates with other networks, such a mobile, where the following transcoding scenarios are evaluated:

- Voice call from VoIP to mobile device.
- Voice calls from mobile device to VoIP.

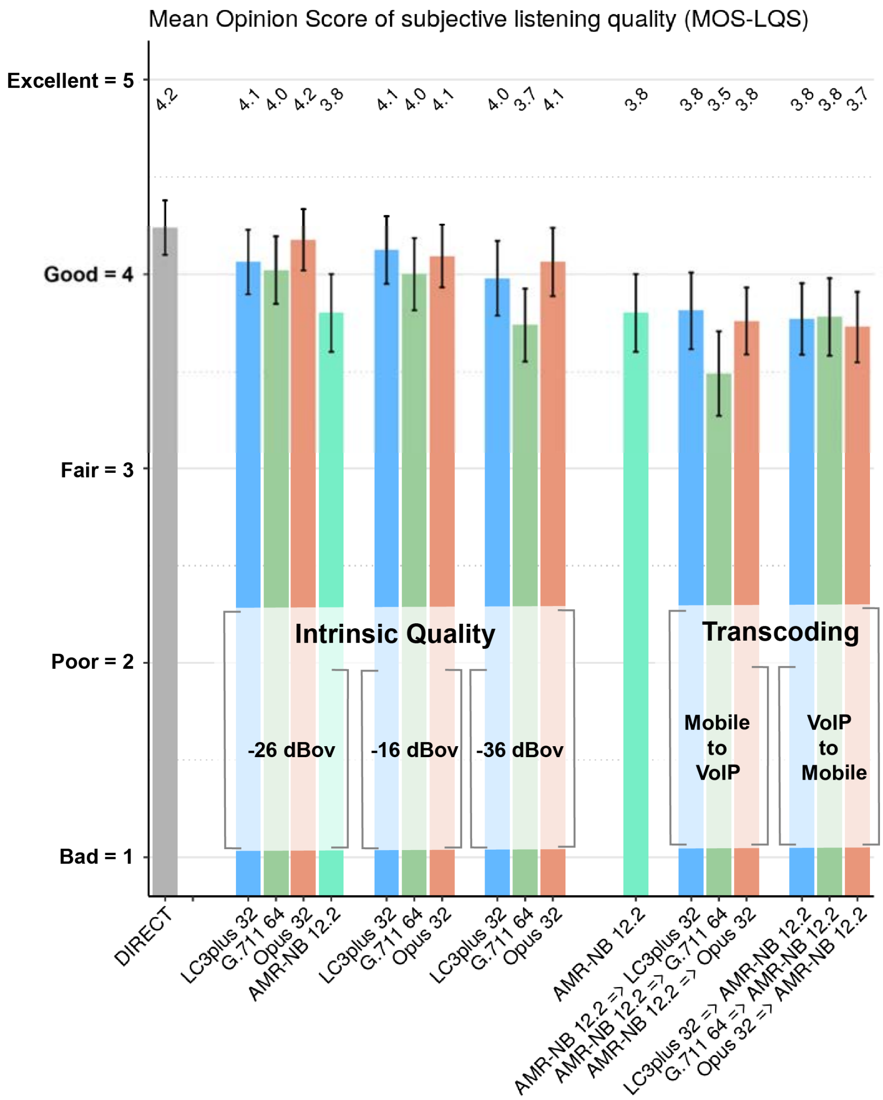
The VoIP legacy codecs are G.711 [i.10] (NB), G.722 [i.11] (WB) and Opus [i.6] (SWB). Legacy mobile terminals utilize AMR-NB [i.13] (NB), AMR-WB [i.5] (WB) and EVS [i.12] (WB and SWB).

As additional performance objective, the Opus [i.6] is added to all tests to complete the performance picture.

7.5.2 NB conditions clean speech

To verify the performance of LC3plus in a NB call over VoIP an ACR experiment (as per Recommendation ITU-T P.800 [i.1]) has been conducted to compare LC3plus at 32 kbps to the legacy VoIP codec G.711 at 64 kbps and to Opus at 32 kbps [i.6].

For the transcoding cases, AMR-NB [i.13] at 12,2 kbps represents the legacy mobile codec.



NOTE: Figure 8 shows mean scores (n = 96) and 95 % confidence intervals of 24 subjects.

Figure 8: P.800 ACR results for VoIP use cases with NB clean speech signals

The following conclusions can be made for VoIP NB systems:

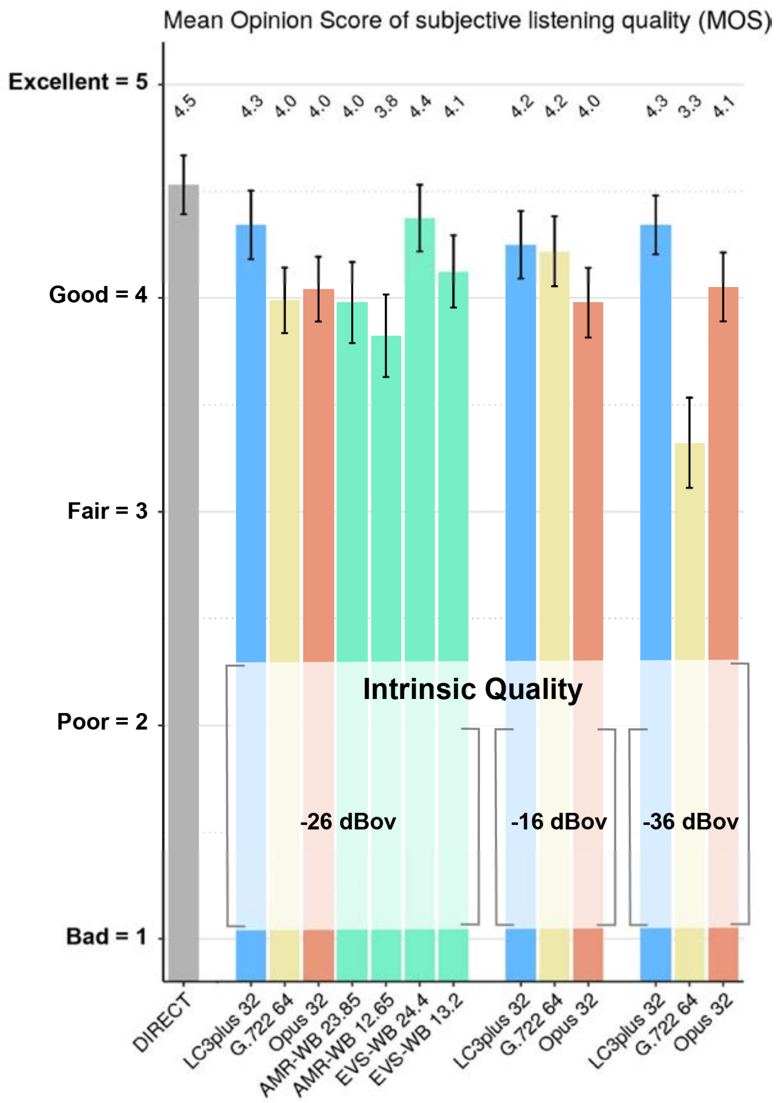
- The test demonstrates that LC3plus operating at 32 kbps is as good as the legacy codec G.711 at 64 kbps. For low input levels and for transcoding from mobile, LC3plus is significantly better than G.711.
- The additional performance objectives in comparison to Opus have been achieved since there is no statistically significant difference between LC3plus and Opus conditions.
- LC3plus and Opus are able to preserve the audio quality of AMR-NB for transcoding.

7.5.3 WB conditions clean speech

To verify the performance of LC3plus in a WB call over VoIP an ACR experiment (as per Recommendation ITU-T P.800 [i.1]) has been conducted to compare LC3plus at 32 kbps to the legacy VoIP codec G.722 at 64 kbps and to Opus at 32 kbps [i.6].

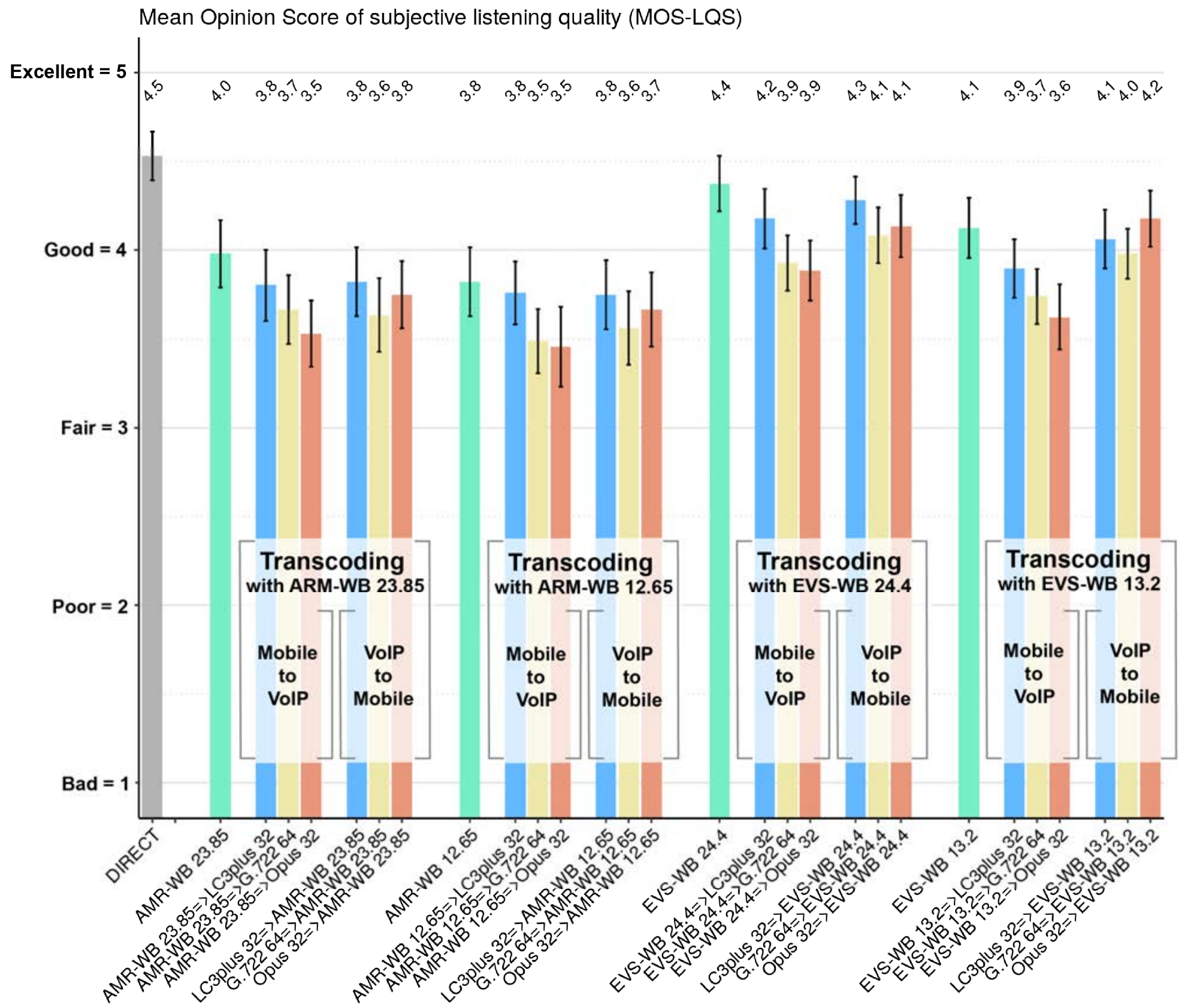
For the transcoding cases, AMR-WB [i.5] at 12,65 kbps and 23,85 kbps as well as EVS [i.12] at 24,4 kbps and 13,2 kbps represent the legacy mobile codecs.

Due to the large number of conditions, the intrinsic quality is shown in Figure 9 while Figure 10 shows all transcoding cases.



NOTE: Figure 9 shows mean scores (n = 96) and 95 % confidence intervals of 24 subjects.

Figure 9: P.800 ACR results for VoIP use cases with WB clean speech signals



NOTE: Figure 10 shows mean scores (n = 96) and 95 % confidence intervals of 24 subjects.

Figure 10: P.800 ACR results for VoIP transcoding use cases with WB clean speech signals

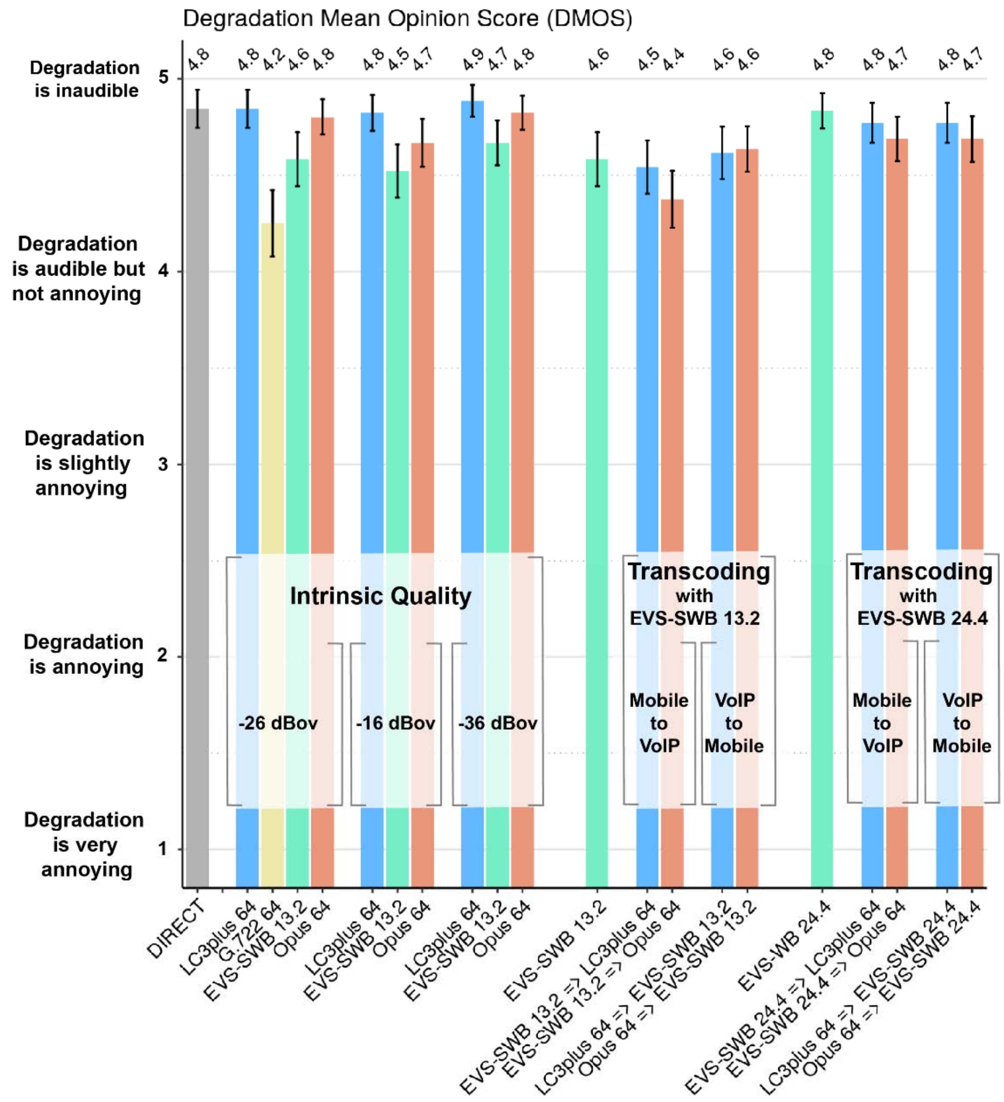
The following conclusions can be made for VoIP WB systems:

- LC3plus at 32 kbps provides clearly better quality than G.722 at 64 kbps and Opus at 32 kbps for all input levels, except for high input levels where G.722 is on par.
- LC3plus doubles the VoIP capacity for WB calls as the codec operated at 32 kbps while the legacy G.722 requires 64 kbps.
- LC3plus 64 kbps provides statistically the same quality as EVS 24,4 kbps.
- For all transcoding cases, LC3plus is always on par or better than the reference codecs G.722 and Opus.
- In 6 out of 8 cases, LC3plus statistically preserves the audio quality of the mobile codec for transcoding. Opus preserves in 2 out of 8 cases, and G.722 in 1 out of 8 cases the audio quality of the mobile codec for transcoding.

7.5.4 SWB conditions clean speech

To verify the performance of LC3plus in a SWB call over VoIP a DCR experiment (as per Recommendation ITU-T P.800 [i.1]) has been conducted to compare LC3plus at 64 kbps to the legacy VoIP codec G.722 [i.11] at 64 kbps, Opus [i.6] at 64 kbps and EVS [i.12] at 13,2 kbps.

For the transcoding cases, EVS at 13,2 kbps and 24,4 kbps represent the legacy mobile codecs.



NOTE: Figure 11 shows mean scores (n = 96) and 95 % confidence intervals of 24 subjects.

Figure 11: P.800 DCR results for VoIP transcoding use cases with SWB clean speech signals

The following conclusions can be made for VoIP SWB systems:

- LC3plus at 64 kbps provides clearly better quality than G.722 at 64 kbps which confirms the benefit of SWB compared to WB.
- LC3plus at 64 kbps is significantly better than EVS at 13,2 kbps and significantly better than Opus at 64 kbps for high level input. For regular and amplified input, LC3plus is on par with Opus.
- For transcoding, LC3plus shows a higher MOS compared to Opus when interoperating with mobile codecs.

7.6 VoIP scenarios with error prone channels

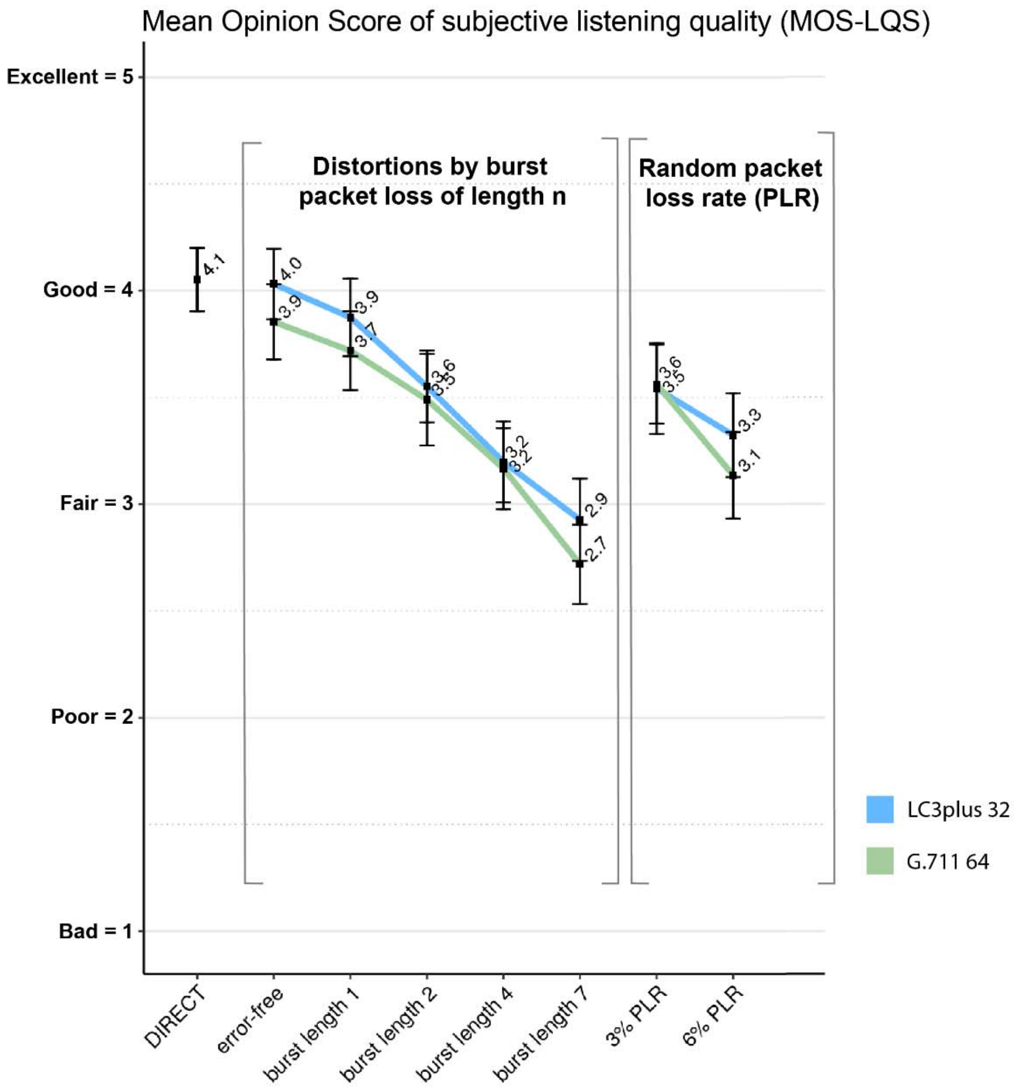
7.6.1 Overview

Two kinds of error profiles have been evaluated. The VoIP burst error patterns are based on 1,43 % PLR using a random pattern of single frame losses which have been extended to 2 (2,87 % PLR), 4 (5,74 % PLR) and 7 (10,05 % PLR) consecutive frame losses to form the burst pattern.

Additionally, random packet loss patterns have been evaluated with Packet Loss Rates (PLR) of 3 % and 6 %.

7.6.2 NB conditions error prone speech

For VoIP NB systems, LC3plus at 32 kbps rate is compared to G.711 at 64 kbps with PLC as defined in Appendix I of Recommendation ITU-T G.711 [i.10].



NOTE: Figure 12 shows mean scores (n = 96) and 95 % confidence intervals of 24 subjects.

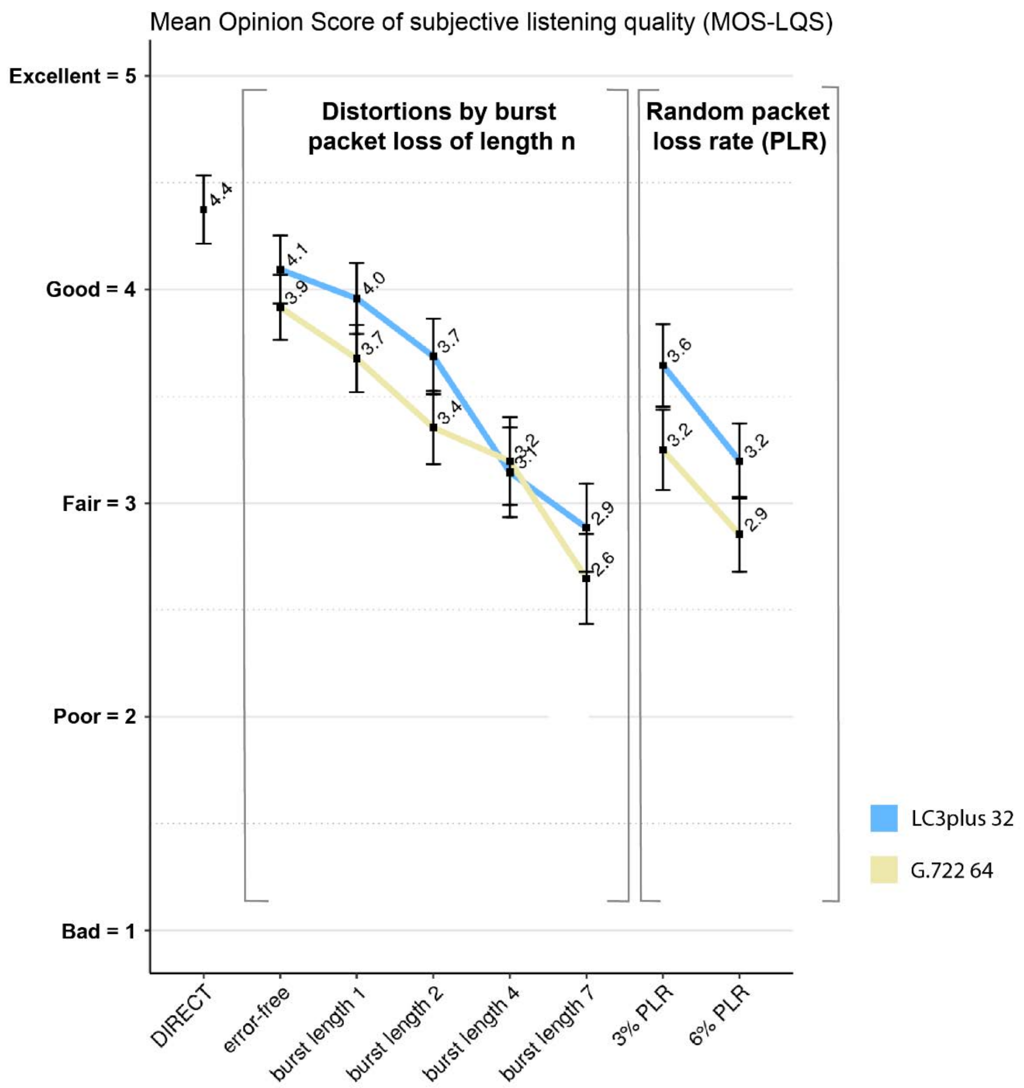
Figure 12: P.800 ACR results for VoIP use cases with NB error prone speech signals

The following conclusions can be made for VoIP NB systems:

- The test demonstrates that LC3plus operating at 32 kbps shows a higher or at least the same MOS compared to the legacy codec G.711 at 64 kbps. This holds true for burst and random frame losses.

7.6.3 WB conditions error prone speech

For VoIP WB systems, LC3plus at 32 kbps rate is compared to G.722 at 64 kbps with PLC as defined in Appendix IV of Recommendation ITU-T G.722 [i.11].



NOTE: Figure 13 shows mean scores (n = 96) and 95 % confidence intervals of 24 subjects.

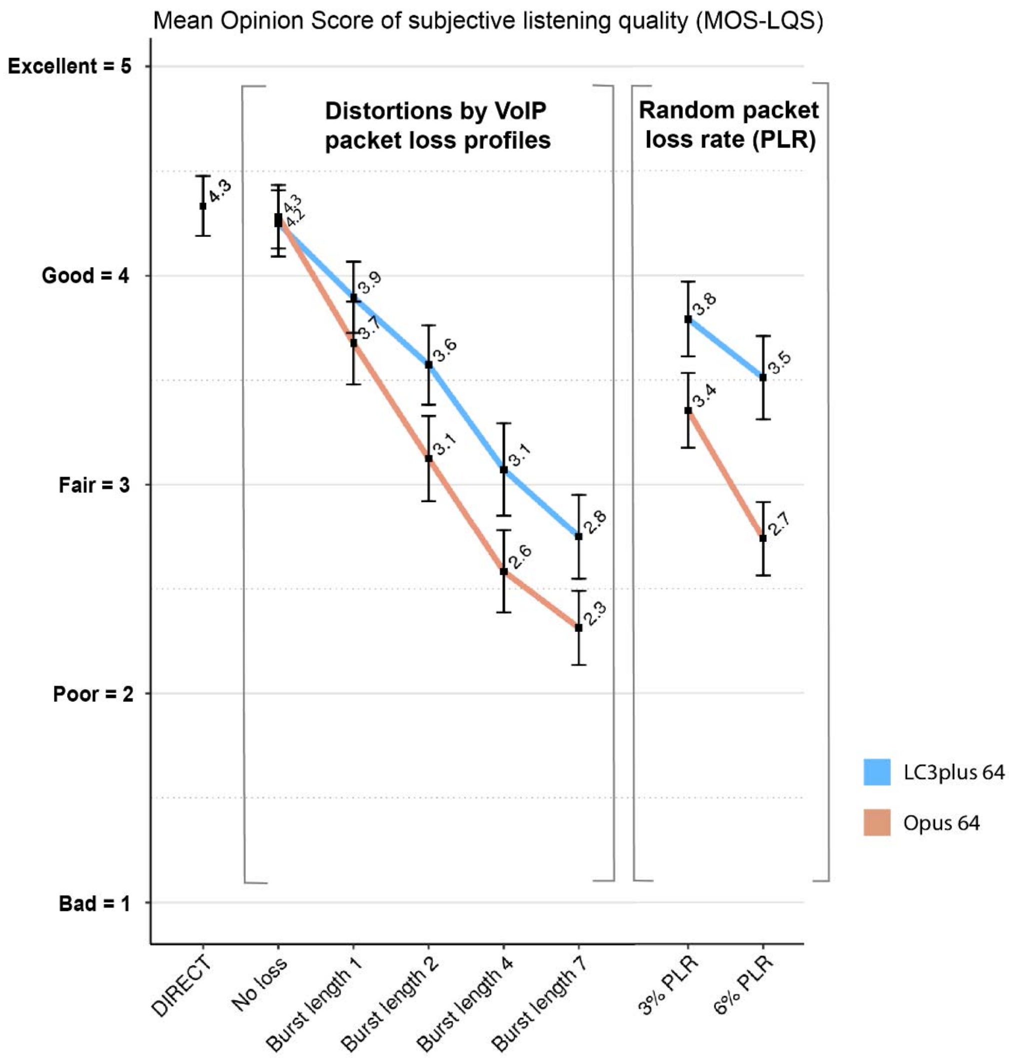
Figure 13: P.800 ACR results for VoIP use cases with WB error prone speech signals

The following conclusions can be made for VoIP WB systems:

- The test demonstrates that LC3plus operating at 32 kbps shows in 6 out of 7 conditions a higher MOS compared to the legacy codec G.722 at 64 kbps.
- For random frame losses, LC3plus performs significantly better than G.722.

7.6.4 SWB conditions error prone speech

For VoIP SWB systems, LC3plus at 64 kbps is compared to Opus [i.6] at 64 kbps.



NOTE: Figure 14 shows mean scores (n = 96) and 95 % confidence intervals of 24 subjects.

Figure 14: P.800 ACR results for VoIP use cases with SWB error prone speech signals

The following conclusions can be made for VoIP SWB systems:

- The test demonstrates that LC3plus outperforms Opus when operating over error-prone channels including random and burst errors.

8 Objective Measurements

8.1 ODG Measurements

The music quality of LC3plus was assessed using the advanced model of Recommendation ITU-R BS.1387 [i.14] (PEAQ). The test corpus consists of 12 critical music items [i.9] at 48 kHz sampling rate which serve as standard test set in other standardization bodies such as MPEG. Figure 15 shows the Objective Difference Grade (ODG) of LC3plus in relation to the used bit rate and frame duration. ODG corresponds to the subjective difference grade (SDG) for a Recommendation ITU-R BS.1116 [i.15] test method, meaning ODG of 0 corresponds to 5 (Imperceptible) of the Recommendation ITU-R BS.1116 [i.15] scale.

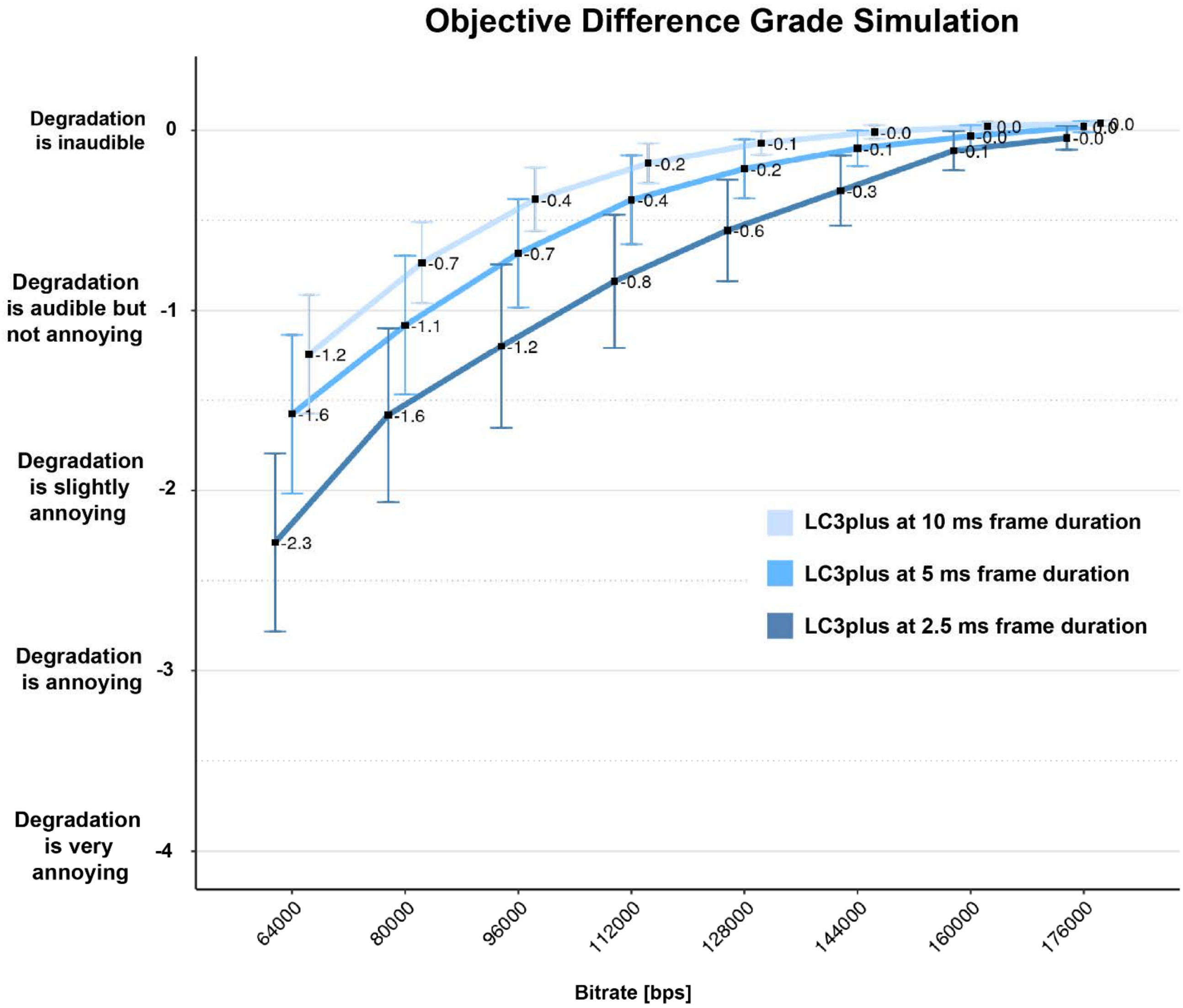


Figure 15: ODG estimation for 48 kHz mono signals depending on bit rate and frame duration

For items close to transparency, subjects need to guess the correct original item and therefore, an SDG of 5,0 is impossible. To reflect this behaviour, an average ODG of -0.1 is assumed to identify the bit rate where LC3plus provides an audio quality close to perceptual transparency.

Therefore, the following conclusion can be drawn from Figure 15:

- LC3plus scales with higher bitrate towards a perceptual transparent audio quality.
- For 10 ms frame duration, LC3plus achieves an audio quality close to perceptual transparency at 128 kbps.
- For 5 ms frame duration, LC3plus achieves an audio quality close to perceptual transparency at 144 kbps.
- For 2,5 ms frame duration, LC3plus achieves an audio quality close to perceptual transparency at 160 kbps.
- Compared to 10 ms frame duration, LC3plus at 5 ms frame duration requires ca. 20 % higher bitrate to achieve the same level of audio quality.
- Compared to 10 ms frame duration, LC3plus at 2,5 ms frame duration requires ca. 60 % higher bitrate to achieve the same level of audio quality.

8.2 Distortion Metrics

8.2.1 General

The quality of high-resolution audio representation is usually described by Total Harmonic Distortion plus Noise (THD+N) and Signal to Noise Ratio (SNR). Those metrics quantify the distortion compared to the original representation of 24 bits per sample and 48 kHz or 96 kHz sampling rate.

In the following, LC3plus coded signals (in floating point arithmetic) are analysed and the THD+N and SNR distortions at 1 kHz sine wave, stimulus level -3 dB, 24 and 48 kHz measured audio bandwidth are plotted with respect to the used bit rate. To outline the benefit of the High-Resolution mode of LC3plus, the plots compare the LC3plus High-Resolution mode to the regular coding mode. A detailed description of the THD+N metric can be found in clause 7.3.5.4.2 of ETSI TS 103 634 [i.4] and the description of the SNR metric can be found in clause 7.3.5.4.3 of ETSI TS 103 634 [i.4].

8.2.2 Total Harmonic Distortion plus Noise (THD+N)

Figure 16 shows the THD+N for 48 kHz sampling rate and Figure 17 shows the THD+N for 96 kHz sampling rate.

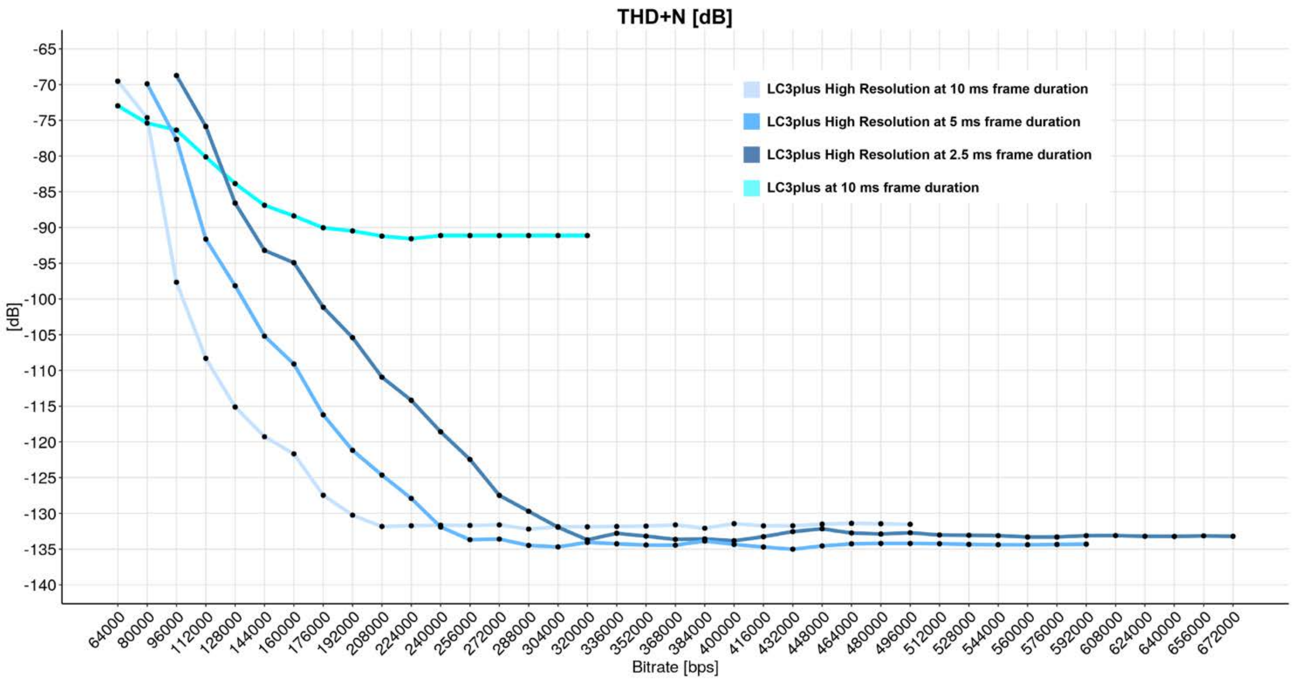


Figure 16: THD+N measurement at stimulus level -3 dB, 48 kHz sampling rate, 1 kHz sine wave, 24 kHz measured bandwidth at different frame durations for LC3plus high resolution and regular mode

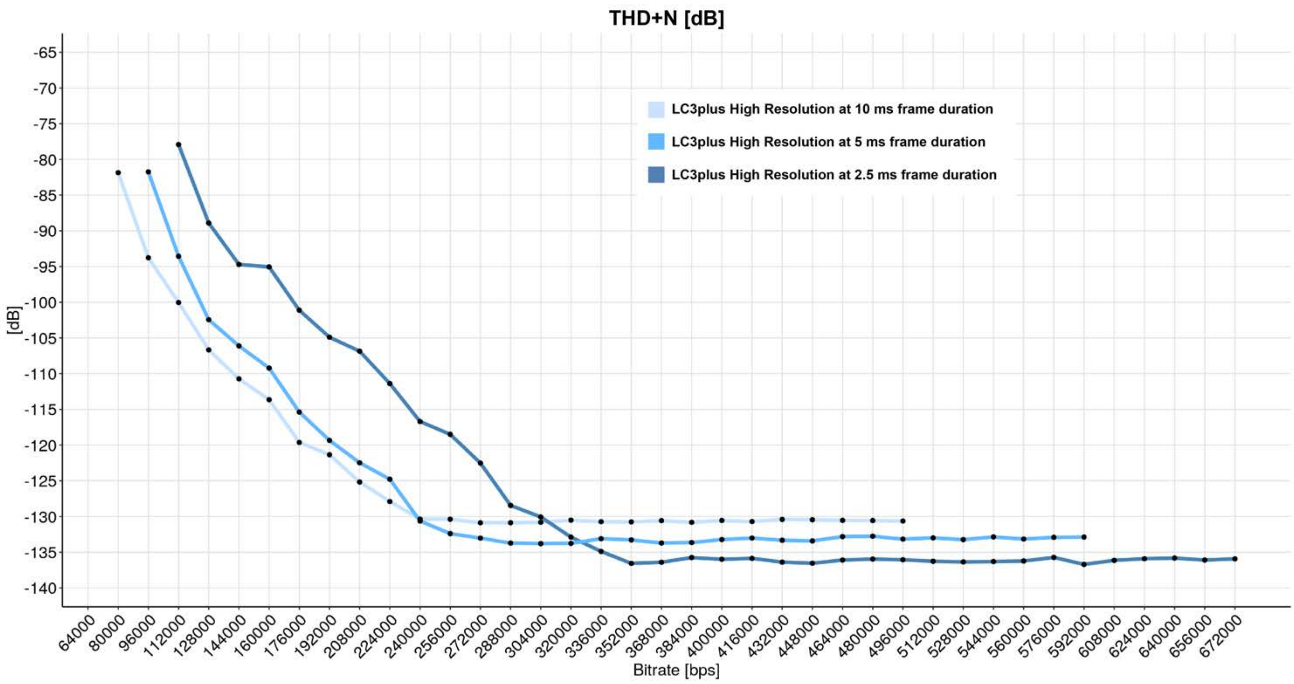


Figure 17: THD+N measurement at stimulus level -3 dB, 96 kHz sampling rate, 1 kHz sine wave, 48 kHz measured bandwidth at different frame durations for LC3plus high resolution

The High-Resolution mode can clearly provide less distortions compared to the regular LC3plus coding mode (at 48 kHz sampling rate) and can achieve a THD+N of less than -130 dB for all frame sizes (10 ms, 5 ms and 2,5 ms).

8.2.3 Signal to Noise Ratio (SNR)

Figure 18 shows the SNR for 48 kHz sampling rate and Figure 19 shows the SNR for 96 kHz sampling rate.

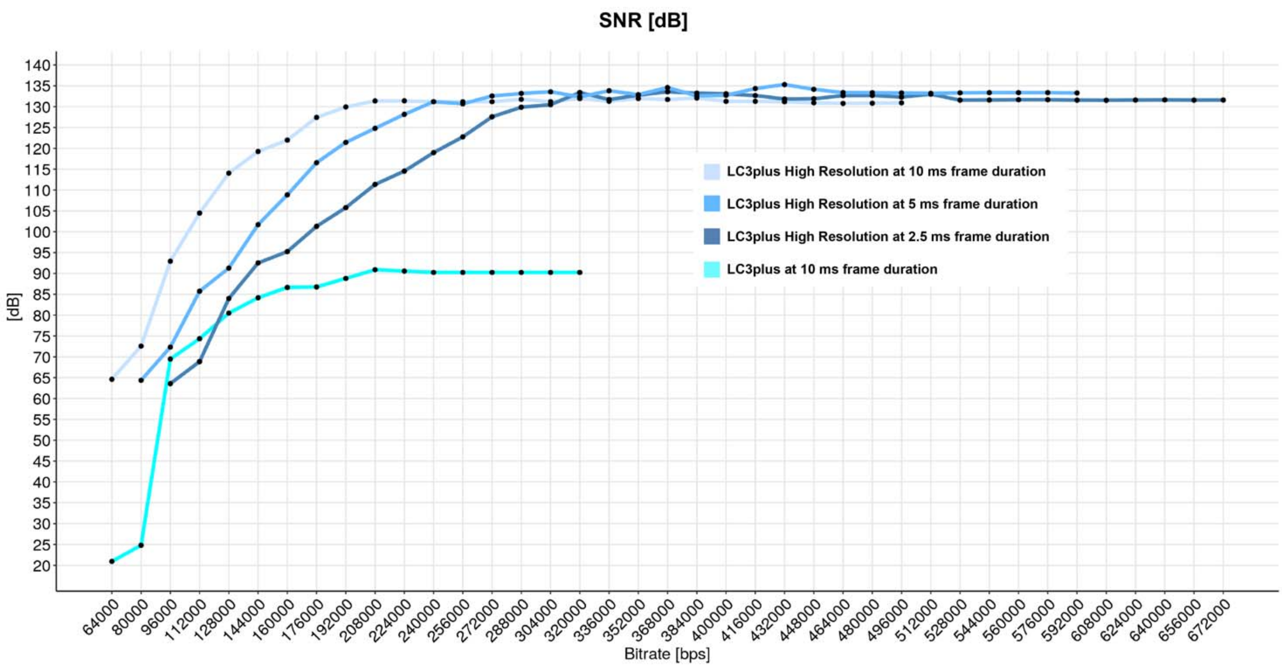


Figure 18: SNR measurement at 1 kHz sine wave and 48 kHz sampling rate at different frame durations for LC3plus high resolution and regular mode

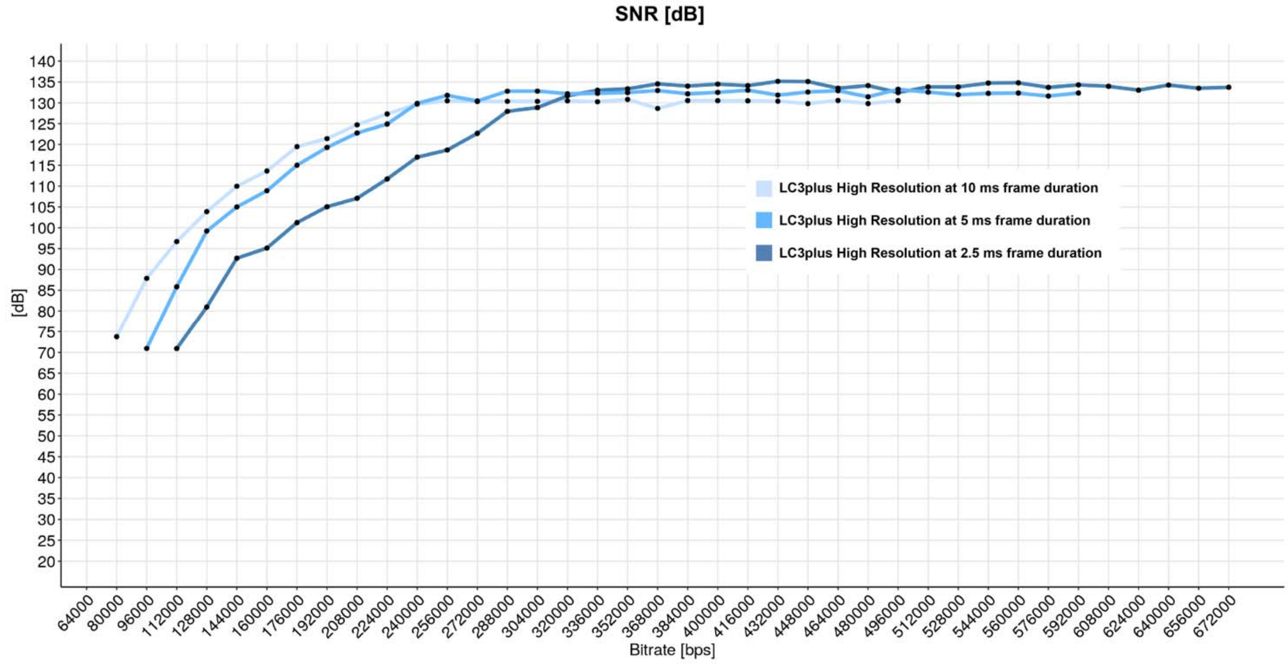


Figure 19: SNR measurement at 1 kHz sine wave and 96 kHz sampling rate at different frame durations for LC3plus high resolution

The High-Resolution mode can clearly provide a lower distortion compared to the regular LC3plus coding mode and can achieve an SNR of more than 130 dB for all frame sizes (10 ms, 5 ms and 2,5 ms).

8.2.4 Conclusion

The data in clauses 8.2.2 and 8.2.3 show that LC3plus in High-Resolution mode is able to transmit high-resolution audio content with very low distortions, even for small frame durations of 5 ms and 2,5 ms.

8.3 Complexity/Memory Usage

Table 2 lists the complexity and memory data points of G.726, G.722 and LC3plus for all sample rate as provided in ETSI EN 300 175-8 [i.8].

Table 2: Overview memory and complexity

	G.726	G.722	LC3plus NB	LC3plus WB	LC3plus SWB	LC3plus FB
Bitrate	32 kbps	64 kbps	32 kbps	32 kbps	64 kbps	128 kbps
Samplerate	8 kHz	16 kHz	8 kHz	16 kHz	32 kHz	48 kHz
Static RAM	≤ 1 kbyte	≤ 1 kbyte	5,6 kbyte	7,3 kbyte	12,1 kbyte	16,9 kbyte
Dynamic RAM			3,1 kbyte	5,8 kbyte	11,2 kbyte	16,6 kbyte
Table ROM	1,1 kbyte	4,5 kbyte	84,2 kbyte	88,7 kbyte	91,6 kbyte	95,9 kbyte
PROM	17,5 kbyte	24,5 kbyte	238,3 kbyte	238,1 kbyte	241,3 kbyte	247,5 kbyte
Complexity worst frame	9,2 WMOPS (see note 2)	7,7 WMOPS (see note 1)	9,7 WMOPS	14,0 WMOPS	24,4 WMOPS	32,4 WMOPS
Complexity average	8,8 WMOPS (see note 2)	7,3 WMOPS (see note 1)	8,3 WMOPS	11,6 WMOPS	19,1 WMOPS	26,4 WMOPS
NOTE 1: G.722 complexity is measured using STL 2009 implementation and the same file as for LC3. NOTE 2: The WMOPS complexity of G.726 is estimated based on the MIPS numbers provided in ETSI EN 300 175-8 [i.8] for G.726 and G.722.						

For LC3plus, the complexity scales up with increasing sample rate and audio bandwidth.

For WB and NB, the complexity of LC3plus is close to G.726 and G.722.

The RAM consumption is higher for LC3plus compared to the legacy codecs, however the codec fits typical DSP architectures.

NOTE: LC3plus and G.722/G.726 use different STL BASOP counters for estimating the complexity, reflecting the processor capability at time of development. Therefore, G.722/G.726 might require less computational complexity on modern processors than LC3plus.

9 Conclusion

The present document shows that LC3plus fulfils all design goals listed in clause 4.

LC3plus successfully enables SWB voice services for DECT and VoIP and provides significantly better audio quality compared to legacy WB services, see clauses 7.3.4 and 7.5.4.

LC3plus doubles the capacity for DECT (and VoIP) WB voice service by reducing the required bit rate from 64 kbps to 32 kbps compared to the legacy G.722. At the same time, the quality is even significantly improved, see clauses 7.3.3 and 7.5.3.

LC3plus provides significantly better error robustness over DECT channels compared to legacy codecs (see clauses 7.4.2 to 7.4.4). For VoIP scenarios, LC3plus shows in 50 % of all test cases a significant improvement compared to the reference codecs and in 50 % of all test cases, LC3plus is on par with the reference codec (see clauses 7.6.2 to 7.6.4).

LC3plus consistently matches or exceeds the transcoding performance with legacy VoIP and mobile codecs compared to reference codecs, see clauses 7.3.2 to 7.3.4 and 7.5.2 to 7.5.4. Especially for self-tandeming conditions, LC3plus shows in 10 out of 12 conditions a significant improvement.

LC3plus scales with increasing bitrate towards a transparent music mode, see clause 8.1. At 128 kbps per channel and 10 ms frame duration, LC3plus achieves an audio quality level close to perceptual transparency.

LC3plus provides the possibility to transmit High-Resolution audio content with a THD+N of less than -130 dB and SNR higher than 130 dB, see clause 8.2. LC3plus starts to provide this audio quality level at 240 kbps per channel.

LC3plus can operate with frame durations of 10 ms, 5 ms and 2,5 ms which results in a total codec delay of 12,5 ms, 7,5 ms and 5 ms. This enables besides DECT voice and VoIP applications also new ones such as gaming headsets. Smaller frame sizes require a slightly higher bit budget to achieve the same level of quality, see clauses 7.3.5 and 8.1.

LC3plus offers a computational complexity close to G.726 and G.722 when operating at NB or WB, see clause 8.3. The RAM and ROM demand of LC3plus is higher than G.722 and G.726, however the codec is implementable on common DSPs.

Annex A: Statistical evaluation

A.1 General

A.1.1 Statistical evaluation

To verify if the requirements for the mentioned use cases were met, selected CuT conditions have been compared to their corresponding reference points Ref by one-sided two sample independent t-tests (also known as Student t-tests). This statistical hypothesis test verifies if the null hypothesis of equal mean values of two conditions can be rejected. For this purpose, the p-value, derived from the test statistic *t* can be calculated. $t = \frac{|\mu_x - \mu_y|}{\sqrt{\frac{\sigma_x{}^2 - \sigma_y{}^2}{96}}}$. The p-value measures the probability of the null hypothesis being true. A p-value ≥ 0,05 means that there is **no** significant difference between the tested data groups. Whereas p < 0,05 means the null hypothesis can be rejected i.e. condition A is significantly better than condition B on a significance level of 5 %.

In the following, statistical results for each use case are listed in p-value tables. The corresponding plots are presented in clause 7.

A.1.2 Data source

The statistical data was created in course of TTF 005 and reviewed by ETSI TC STQ. All data is available in an electronic attachment to ETSI TS 103 624 [i.2].

A.2 Detailed results

A.2.1 DECT NB clean speech

Table A.1: Statistical significance between CuT and Ref for DECT use cases with NB clean speech signals according to Figure 1

CuT	Ref	P-value	Interpretation
LC3plus 32	G.726 32	0.000044	CuT is significantly better
LC3plus 32 (-16 dBoV)	G.726 32(-16 dBoV)	0,000049	CuT is significantly better
LC3plus 32 (-36 dBoV)	G.726 32(-36 dBoV)	0,000002	CuT is significantly better
LC3plus 32	Opus 32	0,161795	No significant difference
LC3plus 32 (-16 dBoV)	Opus 32 (-16 dBoV)	0,397296	No significant difference
LC3plus 32 (-36 dBoV)	Opus 32 (-36 dBoV)	0,263816	No significant difference
G.711 64=>LC3plus 32	G.711 64=>G.726 32	0,000012	CuT is significantly better
LC3plus 32=>G.711 64	G.726 32=>G.711 64	0,001131	CuT is significantly better
LC3plus 32=>G.711 64=>LC3plus 32	G.726 32=>G.711 64=>G.726 32	0,009369	CuT is significantly better
G.711 64=>LC3plus 32	G.711 64=>Opus 32	0,215446	No significant difference
LC3plus 32=>G.711 64	Opus 32=>G.711 64	0,469953	No significant difference
LC3plus 32=>G.711 64=>LC3plus 32	Opus 32=>G.711 64=>Opus 32	0,470836	No significant difference
LC3plus 32=>LC3plus 32	G.726 32=>G.726 32	0,000143	CuT is significantly better
LC3plus 32=>LC3plus 32=>LC3plus 32	G.726 32=>G.726 32=>G.726 32	0,001767	CuT is significantly better
LC3plus 32=>LC3plus 32	Opus 32=>Opus 32	0,372877	No significant difference
LC3plus 32=>LC3plus 32=>LC3plus 32	Opus 32=>Opus 32=>Opus 32	0,266743	No significant difference
LC3plus 32=>LC3plus 32	G.711 64=>G.711 64	0,408308	No significant difference
LC3plus 32=>LC3plus 32=>LC3plus 32	G.711 64=>G.711 64=>G.711 64	0,468691	No significant difference

A.2.2 DECT WB clean speech

Table A.2: Statistical significance between CuT and Ref for DECT use cases with WB clean speech signals according to Figure 2 and Figure 4

CuT	Ref	P-value	Interpretation
LC3plus 32	G.722 64	0,001062	CuT is significantly better
LC3plus 32(-16 dBoV)	G.722 64 (-16 dBoV)	0,393117	no significant difference
LC3plus 32 (-36 dBoV)	G.722 64 (-36 dBoV)	0,000000	CuT is significantly better
LC3plus 32	Opus 32	0,004014	CuT is significantly better
LC3plus 32 (-16 dBoV)	Opus 32 (-16 dBoV)	0,010122	CuT is significantly better
LC3plus 32 (-36 dBoV)	Opus 32 (-36 dBoV)	0,003816	CuT is significantly better
LC3plus 32	LC3plus 32 (5 ms frame duration)	0,000004	CuT 10 ms is significantly better than CuT 5 ms
LC3plus 32	LC3plus 48 (5 ms frame duration)	0,160456	no significant difference
LC3plus 32	LC3plus 64 (2,5 ms frame duration)	0,313939	no significant difference
LC3plus 32	LC3plus 96 (2,5 ms frame duration)	0,060642	no significant difference
LC3plus 32=>G.722 64	G.722 64=>G.722 64	0,397883	no significant difference
G.722 64=>LC3plus 32	G.722 64=>G.722 64	0,067196	no significant difference
LC3plus 32=>G.722 64=>LC3plus 32	G.722 64=>G.722 64=>G.722 64	0,074557	no significant difference
LC3plus 32=>G.722 64	Opus 32=>G.722 64	0,058381	no significant difference
G.722 64=>LC3plus 32	G.722 64=>Opus 32	0,039515	CuT is significantly better
LC3plus 32=>G.722 64=>LC3plus 32	Opus 32=>G.722 64=>Opus 32	0,000031	CuT is significantly better

A.2.3 DECT SWB clean speech

Table A.3: Statistical significance between CuT and Ref for DECT use cases with SWB clean speech signals according to Figure 3 and Figure 4

CuT	Ref	P-value	Interpretation
LC3plus 64	EVS-SWB 13,2	0,001642	CuT is significantly better
LC3plus 64 (-16 dBov)	EVS-SWB 13,2 (-16 dBov)	0,000248	CuT is significantly better
LC3plus 64 (-36 dBov)	EVS-SWB 13,2 (-36 dBov)	0,001506	CuT is significantly better
LC3plus 64	G.722 64	0,000000	CuT is significantly better
LC3plus 64	Opus 64	0,270094	No significant difference
LC3plus 64 (-16 dBov)	Opus 64 (-16 dBov)	0,023841	CuT is significantly better
LC3plus 64 (-36 dBov)	Opus 64 (-36 dBov)	0,153346	No significant difference
LC3plus 64	LC3plus 64 (5 ms frame duration)	0,332632	No significant difference
LC3plus 64	LC3plus 96 (5 ms frame duration)	0,439158	No significant difference
LC3plus 64	LC3plus 96 (2,5 ms frame duration)	0,209698	No significant difference
LC3plus 64	LC3plus 128 (2,5 ms frame duration)	0,325331	No significant difference
LC3plus 64=>LC3plus 64	LC3plus 64=>Opus 64	0,139994	No significant difference
LC3plus 64=>LC3plus 64	Opus 64=>LC3plus 64	0,301130	No significant difference
LC3plus 64=>LC3plus 64	Opus 64=>Opus 64	0,000427	CuT is significantly better
LC3plus 64=>LC3plus 64	EVS-SWB 13.2=>EVS-SWB 13.2	0,000000	CuT is significantly better
LC3plus 64=>LC3plus 64=>LC3plus 64	LC3plus 64=>Opus 64=>LC3plus 64	0,134740	No significant difference
LC3plus 64=>LC3plus 64=>LC3plus 64	Opus 64=>Opus 64=>Opus 64	0,000000	CuT is significantly better
LC3plus 64=>LC3plus 64=>LC3plus 64	EVS-SWB 13.2=>EVS-SWB 13.2=>EVS-SWB 13.2	0,000000	CuT is significantly better

A.2.4 DECT NB error prone speech

Table A.4: Statistical significance between CuT and Ref objective for DECT use cases with NB error prone speech signals according to Figure 5

CuT	Ref	P-value	Interpretation
LC3plus 32 (epmode=1) DP0	G.726 32 DP0	0,000006	CuT is significantly better
LC3plus 32 (epmode=2) DP1	G.726 32 DP1	0,000000	CuT is significantly better
LC3plus 32 (epmode=3) DP2	G.726 32 DP2	0,000000	CuT is significantly better
LC3plus 32 (epmode=4) DP3	G.726 32 DP3	0,000000	CuT is significantly better
LC3plus 32 3 % (EPFsize 10 ms)	G.726 32 3 % (EPFsize 10 ms)	0,000000	CuT is significantly better
LC3plus 32 6 % (EPFsize 10 ms)	G.726 32 6 % (EPFsize 10 ms)	0,000000	CuT is significantly better

A.2.5 DECT WB error prone speech

Table A.5: Statistical significance between CuT and Ref for DECT use cases with WB error prone speech signals according to Figure 6

CuT	Ref	P-value	Interpretation
LC3plus 32 (epmode=1) DP0	G.722 64 DP0	0,028048	CuT is significantly better
LC3plus 32 (epmode=2) DP1	G.722 64 DP1	0,177512	No significant difference
LC3plus 32 (epmode=3) DP2	G.722 64 DP2	0,000000	CuT is significantly better
LC3plus 32 (epmode=4) DP3	G.722 64 DP3	0,000000	CuT is significantly better
LC3plus 32 3 % (EPFsize 10 ms)	G.722 64 3 % (EPFsize 10 ms)	0,000008	CuT is significantly better
LC3plus 32 6 % (EPFsize 10 ms)	G.722 64 6 % (EPFsize 10 ms)	0,000164	CuT is significantly better

A.2.6 DECT SWB error prone speech

Table A.6: Statistical significance between CuT and Ref for DECT use cases with SWB error prone speech signals according to Figure 7

CuT	Ref	P-value	Interpretation
LC3plus 64 (epmode=1) DP0	G.722 64 DP0	0,000257	CuT is significantly better
LC3plus 64 (epmode=2) DP1	G.722 64 DP1	0,000000	CuT is significantly better
LC3plus 64 (epmode=3) DP2	G.722 64 DP2	0,000000	CuT is significantly better
LC3plus 64 (epmode=4) DP3	G.722 64 DP3	0,000000	CuT is significantly better
LC3plus 64 3 % (EPFsize 10 ms)	G.722 64 3 % (EPFsize 10 ms)	0,000000	CuT is significantly better
LC3plus 64 6 % (EPFsize 10 ms)	G.722 64 6 % (EPFsize 10 ms)	0,000001	CuT is significantly better

A.2.7 VoIP NB clean speech

Table A.7: Statistical significance between CuT and Ref for VoIP use cases with NB clean speech signals according to Figure 8

CuT	Ref	P-value	Interpretation
LC3plus 32	G.711	0,365699	No significant difference
LC3plus 32 (-16 dBov)	G.711 (-16 dBov)	0,166498	No significant difference
LC3plus 32 (-36 dBov)	G.711 (-36 dBov)	0,039681	CuT is significantly better
LC3plus 32	Opus 32	0,161795	No significant difference
LC3plus 32 (-16 dBov)	Opus 32 (-16 dBov)	0,397296	No significant difference
LC3plus 32 (-36 dBov)	Opus 32 (-36 dBov)	0,263816	No significant difference
LC3plus 32	AMR-NB	0,024264	CuT is significantly better
AMR-NB 12.2=>LC3plus 32	AMR-NB 12.2=>G.711 64	0,015142	CuT is significantly better
LC3plus 32=>AMR-NB 12.2	G.711 64=>AMR-NB 12.2	0,469435	No significant difference
AMR-NB 12.2=>LC3plus 64	AMR-NB 12.2=>Opus 32	0,345731	No significant difference
LC3plus 64=>AMR-NB 12.2	Opus 32=>AMR-NB 12.2	0,373764	No significant difference
AMR-NB 12.2	AMR-NB 12.2=>LC3plus 64	0,296789	Transcoding preserves quality
AMR-NB 12.2	AMR-NB 12.2=>G.711 64	0,042635	Transcoding degrades quality
AMR-NB 12.2	AMR-NB 12.2=>Opus 32	0,435216	Transcoding preserves quality
AMR-NB 12.2	LC3plus 32=>AMR-NB 12.2	0,406260	Transcoding preserves quality
AMR-NB 12.2	G.711 64=>AMR-NB 12.2	0,380751	Transcoding preserves quality
AMR-NB 12.2	Opus 32=>AMR-NB 12.2	0,468287	Transcoding preserves quality

A.2.8 VoIP WB clean speech

Table A.8: Statistical significance between CuT and Ref for VoIP use cases with WB clean speech signals according to Figure 9 and Figure 10

CuT	Ref	P-value	Interpretation
LC3plus 32	G.722 64	0,001062	CuT is significantly better
LC3plus 32	G.722 64	0,393117	No significant difference
LC3plus 32	G.722 64	0,000000	CuT is significantly better
LC3plus 32	Opus 32	0,004014	CuT is significantly better
LC3plus 32	Opus 32	0,010122	CuT is significantly better
LC3plus 32	Opus 32	0,003816	CuT is significantly better
AMR-WB 23.85=>LC3plus 32	AMR-WB 23.85=>G.722 64	0,166460	No significant difference
LC3plus 32=>AMR-WB 23.85	G.722 64=>AMR-WB 23.85	0,095132	No significant difference
AMR-WB 12.65=>LC3plus 32	AMR-WB 12.65=>G.722 64	0,017887	CuT is significantly better
LC3plus 32=>AMR-WB 12.65	G.722 64=>AMR-WB 12.65	0,094772	No significant difference
EVS-WB 24.4=>LC3plus 32	EVS-WB 24.4=>G.722 64	0,016753	CuT is significantly better
LC3plus 32=>EVS-WB 24.4	G.722 64=>EVS-WB 24.4	0,030094	CuT is significantly better
EVS-WB 13.2=>LC3plus 32	EVS-WB 13.2=>G.722 64	0,085893	No significant difference
LC3plus 32=>EVS-WB 13.2	G.722 64=>EVS-WB 13.2	0,224500	No significant difference
AMR-WB 23.85=>LC3plus 32	AMR-WB 23.85=>Opus 32	0,025468	CuT is significantly better
LC3plus 32=>AMR-WB 23.85	Opus 32=>AMR-WB 23.85	0,296173	No significant difference
AMR-WB 12.65=>LC3plus 32	AMR-WB 12.65=>Opus 32	0,019342	CuT is significantly better
LC3plus 32=>AMR-WB 12.65	Opus 32=>AMR-WB 12.65	0,280110	No significant difference
EVS-WB 24.4=>LC3plus 32	EVS-WB 24.4=>Opus 32	0,008558	CuT is significantly better
LC3plus 32=>EVS-WB 24.4	Opus 32=>EVS-WB 24.4	0,096309	No significant difference
EVS-WB 13.2=>LC3plus 32	EVS-WB 13.2=>Opus 32	0,015480	CuT is significantly better
LC3plus 32=>EVS-WB 13.2	Opus 32=>EVS-WB 13.2	0,161795	No significant difference
AMR-WB 23.85	AMR-WB 23.85=> LC3plus 32	0,102244	Transcoding preserves quality
AMR-WB 23.85	AMR-WB 23.85=> G.722 64	0,012007	Transcoding degrades quality
AMR-WB 23.85	AMR-WB 23.85=> Opus 32	0,000596	Transcoding degrades quality
AMR-WB 23.85	LC3plus 32=>AMR-WB 23.85	0,127622	Transcoding preserves quality
AMR-WB 23.85	G.722 64=>AMR-WB 23.85	0,008444	Transcoding degrades quality
AMR-WB 23.85	Opus 32=>AMR-WB 23.85	0,046312	Transcoding degrades quality
AMR-WB 12.65	AMR-WB 12.65=> LC3plus 32	0,317716	Transcoding preserves quality
AMR-WB 12.65	AMR-WB 12.65=> G.722 64	0,007045	Transcoding degrades quality
AMR-WB 12.65	AMR-WB 12.65=> Opus 32	0,008298	Transcoding degrades quality
AMR-WB 12.65	LC3plus 32=>AMR-WB 12.65	0,298338	Transcoding preserves quality
AMR-WB 12.65	G.722 64=>AMR-WB 12.65	0,007045	Transcoding degrades quality
AMR-WB 12.65	Opus 32=>AMR-WB 12.65	0,137901	Transcoding preserves quality
EVS-WB 24.4	EVS-WB 24.4=> LC3plus 32	0,045472	Transcoding degrades quality
EVS-WB 24.4	EVS-WB 24.4=> G.722 64	0,000057	Transcoding degrades quality
EVS-WB 24.4	EVS-WB 24.4=> Opus 32	0,000027	Transcoding degrades quality
EVS-WB 24.4	LC3plus 32=> EVS-WB 24.4	0,183924	Transcoding preserves quality
EVS-WB 24.4	G.722 64=> EVS-WB 24.4	0,005251	Transcoding degrades quality
EVS-WB 24.4	Opus 32=> EVS-WB 24.4	0,022869	Transcoding degrades quality
EVS-WB 13.2	EVS-WB 13.2=> LC3plus 32	0,028708	Transcoding degrades quality
EVS-WB 13.2	EVS-WB 13.2=> G.722 64	0,000587	Transcoding degrades quality
EVS-WB 13.2	EVS-WB 13.2=> Opus 32	0,000064	Transcoding degrades quality
EVS-WB 13.2	LC3plus 32=> EVS-WB 13.2	0,301014	Transcoding preserves quality
EVS-WB 13.2	G.722 64=> EVS-WB 13.2	0,096125	Transcoding preserves quality
EVS-WB 13.2	Opus 32=> EVS-WB 13.2	0,328382	Transcoding preserves quality

A.2.9 VoIP SWB clean speech

Table A.9: Statistical significance between CuT and Ref for VoIP use cases with SWB clean speech signals according to Figure 11

CuT	Ref	P-Value	Interpretation
LC3plus 64	EVS-SWB 13,2	0,001643	Transcoding degrades quality
LC3plus 64(-16 dBov)	EVS-SWB 13,2(-36 dBov)	0,000248	CuT is significantly better
LC3plus 64(-36 dBov)	EVS-SWB 13,2(-36 dBov)	0,001506	CuT is significantly better
LC3plus 64	Opus 64	0,270094	No significant difference
LC3plus 64(-16 dBov)	Opus 64(-16 dBov)	0,023841	CuT is significantly better
LC3plus 64(-36 dBov)	Opus 64(-36 dBov)	0,153346	No significant difference
EVS-SWB 13.2=>LC3plus 64	EVS-SWB 13.2=>Opus 64	0,052286	No significant difference
LC3plus 64=>EVS-SWB 13.2	Opus 64=>EVS-SWB 13.2	0,409384	No significant difference
EVS-SWB 24.4=>LC3plus 64	EVS-SWB 24.4=>Opus 64	0,134740	No significant difference
LC3plus 64=>EVS-SWB 24.4	Opus 64=>EVS-SWB 24.4	0,148479	No significant difference
EVS-SWB 13.2	EVS-SWB 13.2=>LC3plus 64	0,416149	Transcoding preserves quality
EVS-SWB 13.2	LC3plus 64=>EVS-SWB 13.2	0,169363	Transcoding preserves quality
EVS-SWB 13.2	EVS-SWB 13.2=>Opus 64	0,077576	Transcoding preserves quality
EVS-SWB 13.2	Opus 64=>EVS-SWB 13.2	0,106538	Transcoding preserves quality
EVS-SWB 24.4	EVS-SWB 24.4=>LC3plus 64	0,500000	Transcoding preserves quality
EVS-SWB 24.4	LC3plus 64=>EVS-SWB 24.4	0,500000	Transcoding preserves quality
EVS-SWB 24.4	EVS-SWB 24.4=>Opus 64	0,139854	Transcoding preserves quality
EVS-SWB 24.4	Opus 64=>EVS-SWB 24.4	0,144239	Transcoding preserves quality

A.2.10 VoIP NB error prone speech

Table A.10: Statistical significance between CuT and Ref for VoIP use cases with NB error prone speech signals according to Figure 12

CuT	Ref	P-Value	Interpretation
LC3plus 32	G.711 64	0,074315	No significant difference
LC3plus 32 PLP1	G.711 64 PLP1	0,117521	No significant difference
LC3plus 32 PLP2	G.711 64 PLP2	0,324907	No significant difference
LC3plus 32 PLP4	G.711 64 PLP4	0,408817	No significant difference
LC3plus 32 PLP7	G.711 64 PLP7	0,062874	No significant difference
LC3plus 32 3 % (EPFsize=20)	G.711 64 3 % (EPFsize=20)	0,441766	No significant difference
LC3plus 32 6 % (EPFsize=20)	G.711 64 6 % (EPFsize=20)	0,095132	No significant difference

A.2.11 VoIP WB error prone speech

Table A.11: Statistical significance between CuT and Ref for VoIP use cases with WB error prone speech signals according to Figure 13

CuT	Ref	P-Value	Interpretation
LC3plus 32	G.722 64	0,008196	CuT is significantly better
LC3plus 32 PLP1	G.722 64 PLP1	0,004282	CuT is significantly better
LC3plus 32 PLP2	G.722 64 PLP2	0,362654	No significant difference
LC3plus 32 PLP4	G.722 64 PLP4	0,056116	No significant difference
LC3plus 32 PLP7	G.722 64 PLP7	0,062874	No significant difference
LC3plus 32 3 % (EPFsize=20)	G.722 64 3 % (EPFsize=20)	0,002206	CuT is significantly better
LC3plus 32 6 % (EPFsize=20)	G.722 64 6 % (EPFsize=20)	0,003636	CuT is significantly better

A.2.12 VoIP SWB error prone speech

Table A.12: Statistical significance between CuT and Ref for VoIP use cases with SWB error prone speech signals according to Figure 14

CuT	Ref	P-Value	Interpretation
LC3plus 64	Opus 64	0,388994	No significant difference
LC3plus 64 PLP1	Opus 64 PLP1	0,049882	CuT is significantly better
LC3plus 64 PLP2	Opus 64 PLP2	0,000969	CuT is significantly better
LC3plus 64 PLP4	Opus 64 PLP4	0,000701	CuT is significantly better
LC3plus 64 PLP7	Opus 64 PLP7	0,000833	CuT is significantly better
LC3plus 64 3 % (EPFsize=20)	Opus 64 3 % (EPFsize=20)	0,000438	CuT is significantly better
LC3plus 64 6 % (EPFsize=20)	Opus 64 6 % (EPFsize=20)	0,000000	CuT is significantly better

Annex B:
Change History

Date	Version	Information about changes
12-2022	1.1.1	First publication

History

Document history		
V1.1.1	December 2022	Publication