ETSI TR 103 590 V1.1.1 (2018-09)

Digital Enhanced Cordless Telecommunications (DECT); Study of Super Wideband Codec in DECT for narrowband, wideband and super-wideband audio communication including options of low delay audio connections (≤ 10 ms framing)

Reference
DTR/DECT-00316

Keywords
codec, DECT, superwideband

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 - NAF 742 C
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° 7803/88

Important notice

The present document can be downloaded from:
http://www.etsi.org/standards-search

The present document may be made available in electronic versions and/or in print. The content of any electronic and/or print versions of the present document shall not be modified without the prior written authorization of ETSI. In case of any existing or perceived difference in contents between such versions and/or in print, the only prevailing document is the print of the Portable Document Format (PDF) version kept on a specific network drive within ETSI Secretariat.

Users of the present document should be aware that the document may be subject to revision or change of status. Information on the current status of this and other ETSI documents is available at
https://portal.etsi.org/TB/ETSIDeliverableStatus.aspx

If you find errors in the present document, please send your comment to one of the following services:
https://portal.etsi.org/People/CommiteeSupportStaff.aspx

Copyright Notification

Contents

Intellectual Property Rights

Essential patents

IPRs essential or potentially essential to normative deliverables may have been declared to ETSI. The information pertaining to these essential IPRs, if any, is 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 IPR Policy, no investigation, 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.

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

Since the introduction of additional codecs in New Generation DECT [i.5] in 2007, wideband services have been widely established for fixed line, mobile and OTT communications networks. This trend is gaining even more momentum by services using cutting edge codecs like 3GPP EVS and the upcoming new Bluetooth™ codec by offering super-wideband audio bandwidth.

NOTE: Bluetooth™ is the trade name of a wireless technology standard for exchanging data over short distances (using short-wavelength UHF radio waves in the ISM band from 2,4 - 2,485 GHz) from fixed and mobile devices, and building personal area networks (PANs), owned by the Bluetooth Special Interest Group. This information is given for the convenience of users of the present document and does not constitute an endorsement by ETSI of the technology named. Equivalent technologies may be used if they can be shown to lead to the same results.

Recent market research from several relevant DECT infrastructure providers indicates a demand for upgrading DECT services and standard with additional features enabled by evolved speech and audio codecs.

The present document collects performance requirements to add a real benefit to current and upcoming DECT applications and evaluates the Low Complexity Communication Codec (LC3) on suitability for this as well as discusses possible adaptations for DECT environments in terms of error protection and signalling.

1 Scope

The present document provides a study of technical updates to the DECT standard to enable super wideband (SWB) audio calls in existing DECT slot formats as well as technical improvements to narrowband (NB) and wideband (WB) calls. All required change requests are listed and defined for the different DECT layers to enable high quality audio communication between DECT FP and PP including DECT repeaters (relays). The study includes an investigation on FEC for block-based codecs. Information is provided on the audio quality in some DECT use cases for NB, WB and SWB and potential improvements by a new audio codec are studied.

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] ETSI EN 300 175-1: "Digital Enhanced Cordless Telecommunications (DECT); Common Interface (CI); Part 1: Overview".

[i.2] ETSI EN 300 175-8: "Digital Enhanced Cordless Telecommunications (DECT); Common Interface (CI); Part 8: Speech and audio coding and transmission".

[i.3] EP2901594B1: "Error Detection for sub-band ADPCM encoded sound signal".

[i.4] ETSI EN 300 175-3: "Digital Enhanced Cordless Telecommunications (DECT); Common Interface (CI); Part 3: Medium Access Control (MAC) layer".

[i.5] ETSI TS 102 527-3: "Digital Enhanced Cordless Telecommunications (DECT); New Generation DECT; Part 3: Extended wideband speech services".

[i.6] ETSI EN 300 700: "Digital Enhanced Cordless Telecommunications (DECT); Wireless Relay Station (WRS)".

[i.7] ETSI EN 300 175-5: "Digital Enhanced Cordless Telecommunications (DECT); Common Interface (CI); Part 5: Network (NWK) layer".

[i.8] ETSI EN 300 176-2: "Digital Enhanced Cordless Telecommunications (DECT); Test specification; Part 2: Audio and speech".

[i.9] Recommendation ITU-T P.863 (09-2014): "Perceptual objective listening quality assessment".

[i.10] Recommendation ITU-T P.800 (08-1996): "Methods for subjective determination of transmission quality".

[i.11] Recommendation ITU-T G.191 (03-1996): "Software tools for speech and audio coding standardization" .

[i.12] Recommendation ITU-T G.722 (09-2012): "7 kHz audio-coding within 64 kbit/s".

[i.13] Recommendation ITU-T G.726 (12-1990): "40, 32, 24, 16 kbit/s Adaptive Differential Pulse Code Modulation (ADPCM)".

[i.14] ETSI TS 126 071: "Digital cellular telecommunications system (Phase 2+) (GSM); Universal Mobile Telecommunications System (UMTS); LTE; Mandatory speech CODEC speech processing functions; AMR speech Codec; General description (3GPP TS 26.071)".

[i.15] ETSI TS 126 171: "Digital cellular telecommunications system (Phase 2+) (GSM); Universal Mobile Telecommunications System (UMTS); LTE; Speech codec speech processing functions; Adaptive Multi-Rate - Wideband (AMR-WB) speech codec; General description (3GPP TS 26.171)".

[i.16] ETSI TS 126 441: "Universal Mobile Telecommunications System (UMTS); LTE; Codec for Enhanced Voice Services (EVS); General overview (3GPP TS 26.441)".

[i.17] SIG Bluetooth™ Hands-Free Profile 1.7.1.

3 Definitions and abbreviations

3.1 Definitions

For the purposes of the present document, the terms and definitions given in ETSI EN 300 175-1 [i.1] and the following apply:

Fullband (FB): speech or audio sampled at 48 kHz

Fullband, compact disc (FBCD): speech or audio sampled at 44,1 kHz

Narrowband (NB): speech or audio sampled at 8 kHz

Semi-Super Wideband (SSWB): speech or audio sampled at 24 kHz

Super Wideband (SWB): speech or audio sampled at 32 kHz

Wideband (WB): speech or audio sampled at 16 kHz

3.2 Abbreviations

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

3GPP 3rd Generation Partnership Project
ACR Absolute Category Rating

NOTE: See Recommendation ITU-T P.800 [i.10].

AMR Adaptive Multi-Rate,

NOTE: See ETSI TS 126 071 [i.14].

AMR-WB Adaptive Multi-Rate Wideband

NOTE: See ETSI TS 126 171 [i.15].

BCH Bose-Chaudhuri-Hocquenghem
BER Bit Error Rate
CELT Constrained Energy Lapped Transform
CMR Codec Mode Request
CNG Comfort Noise Generation
CRC Cyclic Redundancy Check
CS Circuit Switched
DECT Digital Enhanced Cordless Telecommunications

DSP	Digital Signal Processor
DTX	Discontinuous Transmission
EP	Error Protection
ETSI	European Telecommunications Standards Institute
EVS	Enhanced Voice Service

NOTE: See ETSI TS 126 441 [i.16].

FB	Fullband
FBCD	Fullband Compact Disc
FEC	Forward Error Correction
FER	Frame Erasure Rate
FP	Fixed Part (DECT bas station)
FT	Frame Type
GFSK	Gaussian Frequency-Shift Keying
GSM	Global System for Mobile Communications
IETF	Internet Engineering Task Force
I_{NA}	higher layer Information channel (unprotected), minimum delay operation

NOTE: See ETSI TS 102 527-3 [i.5].

I_{NB}	higher layer Information channel (unprotected), normal delay operation

NOTE: See ETSI TS 102 527-3 [i.5].

IP	Internet Protocol
LAN	Local Area Network
LC3	Low Complexity Communication Codec
MIPS	Million Instructions Per Second
MNRU	Modulated Noise Reference Unit
MOS-LQO	Mean Opinion Score - Listening Quality Objective
mSBC	modified Subband Coding

NOTE: See [i.17].

NB	Narrowband
OTT	Over-The-Top content
PLC	Packet Loss Concealment
PLR	Packet Loss Rate
PP	Portable Part (DECT handset)
RAM	Random-Access Memory
RSSI	Received Signal Strength Indicator
RTP	Real-Time Protocol
SSWB	Semi-Super Wideband
STL	Software Tool Library
SWB	Super Wideband
VoIP	Voice over Internet Protocol
WB	Wideband
WMOPS	Weighted Millions of Operations Per Second
WRS	Wireless Relay Station

4 Investigations on an enhanced DECT codec

4.1 Overview

The investigations are organized as follows:

1) Definition of general required codec features (clause 4.2).

2) Study on error profiles and protection schemes for DECT systems (clause 4.3).

3) Characterization of the LC3 as potential candidate (clause 4.4).

4) Definition of required update to DECT specifications (clause 4.5).

4.2 Design constraints/features to be supported

4.2.1 Improved exploitation on DECT slots

Table 1 compares the slot related requirements of the legacy DECT codecs with the proposed new DECT codec.

Table 1: Overview of DECT slot related requirements for new codec

Slot usage	Legacy DECT codecs	Enhanced DECT codec
Normal slots	NB calls (G.726)	NB and WB calls
Long slots	WB calls (G.722)	WB and SWB calls

NB and WB audio quality should be comparable to or better relative to the legacy DECT codecs.

The user experience for error prone channels is expected to be comparable to or better relative to the legacy DECT codecs.

4.2.2 Transmission latency

The codec should operate on 10 ms frame sizes as provided by the DECT transmission slots [i.1]. On top of the framing delay of 10 ms, the additional algorithmic delay should be less than or equal to 2,5 ms.

Additionally, the codec should support frame sizes of 5 ms as well to enable new low delay application besides telephony. For instance, in-room conferencing/amplification or parliament systems require a microphone to loudspeaker delay of less than 20 ms. This guarantees lip-synchronism of the speaker to the amplified signal.

4.2.3 Supported sampling rates, audio bandwidths and sample depths

The codec should support NB, WB, SWB and FB audio bandwidths at the native sample rates of 8 kHz, 16 kHz, 32 kHz and 48 kHz. Additionally, 24 kHz (SSWB) should be supported.

The codec should support the coding of lower audio bandwidths for a given sample rate, e.g. coding of NB signals at 32 kHz. The codec should support the coding of audio samples with 16 bits per samples and may support coding of audio samples with 24 bits per sample.

4.2.4 Support for music streaming

The codec should provide decent audio quality for music streaming services and may provide additional coding features to support stereo music channels.

4.2.5 Packet loss concealment

The codec should support packet loss concealment without adding further algorithmic delay. As the main application is voice, the packet loss concealment should perform well for speech signals.

4.2.6 Low codec complexity

The codec should run with a low computational complexity and low memory footprint to be implementable on typical DECT handheld devices. The complexity should be measured and reported using the latest ITU-T STL complexity measurement toolbox [i.11].

4.3 Investigations on impact of block-based codec

4.3.1 Probability and distribution of bit errors

4.3.1.1 Normal slots - transmission error profile I

An error profile was measured using a real DECT system simulating that a DECT caller is moving through an office building. The caller starts close to the base station and walks away through the office. The measurement recorded the number of bit errors, the position of the bit error, the signal strength (RSSI, in steps of eight) and complete frame losses. Figure 1 outlines the error profile where complete frame losses are indicated by 384 bit errors (A+B field) in combination with signal strength zero.

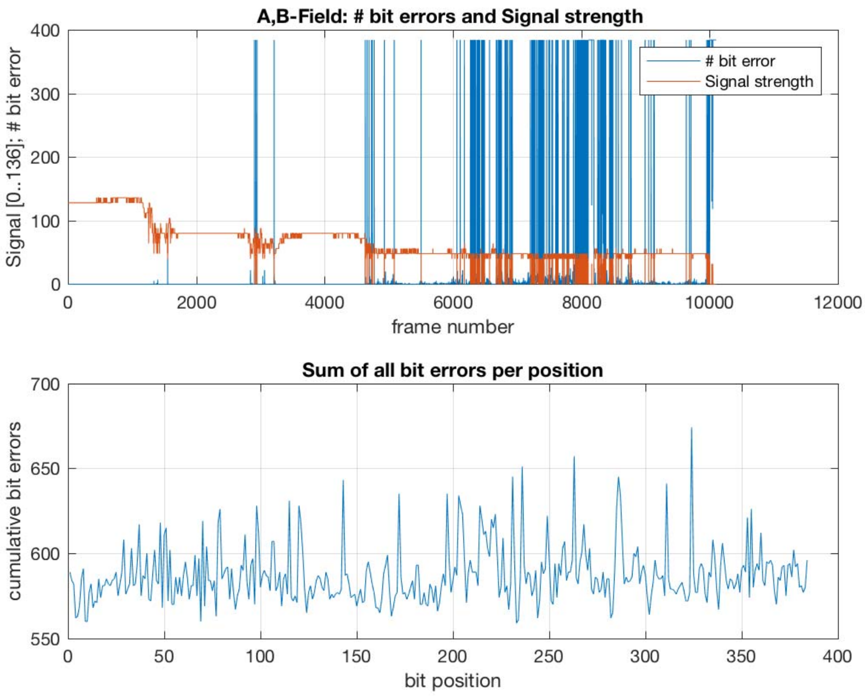

Figure 1: Characterization of Error Profile I

Regarding the position of the bit errors inside the frame, no specific dependency can be found. In order to structure the analysis of the pattern, the bit error profile is further grouped into certain signal strength classes as outlined in Figure 2.

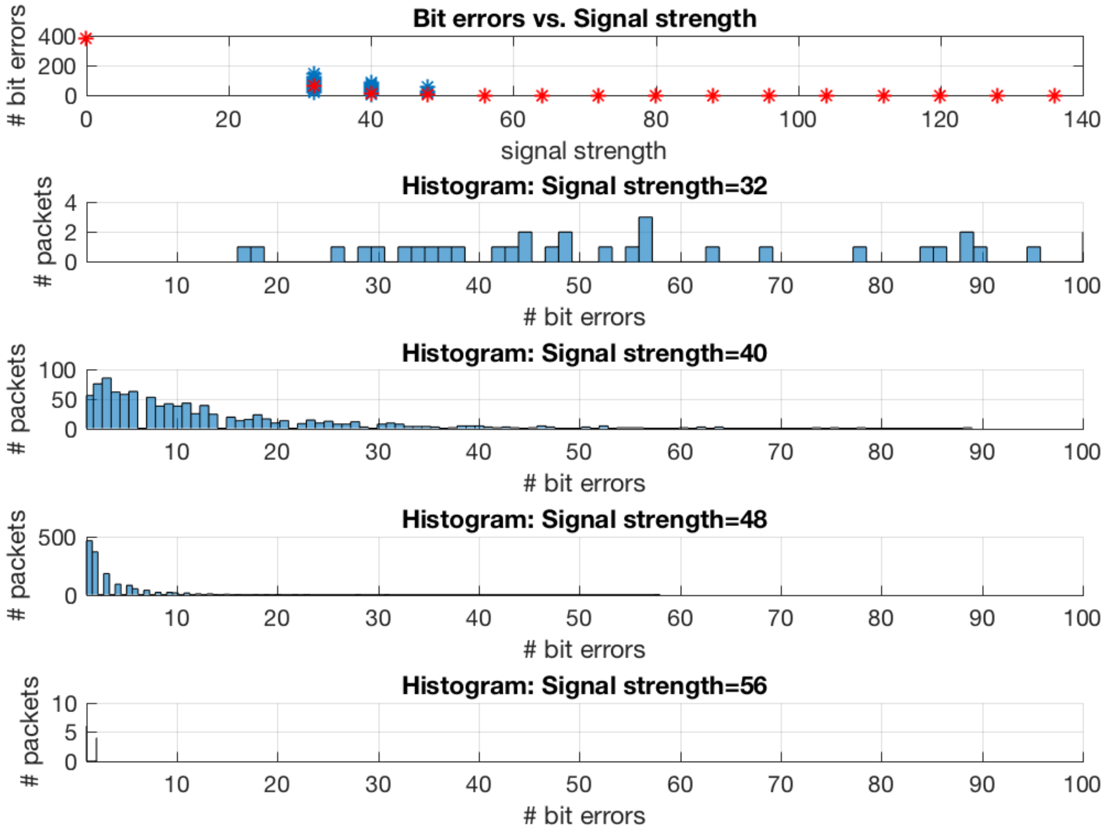

Figure 2: Plot of absolute number (blue) and average number (red) of bit errors for each signal strength (top); Histogram of number of bit errors inside frame for signal strength levels 32 dB, 40 dB, 48 dB and 56 dB

The plots show that:

- For this profile, bit errors only occur for a signal strength level ≤ 56 dB.
- For the signal strength level 48 dB and 56 dB, most packets show less than three bit errors in one packet.
- For signal strength level 40 dB, the bit error rate per packet is significantly higher compared to level 48 dB.
- For signal strength level 32 dB, almost all bits are affected; this level is thus not considered for any recovery activity.
- Four different error protection classes may be appropriate to address the different error characteristics depending on the signal strength, i.e. clean channel, 56 dB, 48 dB, 40 dB.

The packet loss rate (PLR) per signal strength can be estimated by averaging the signal strength over one second. Figure 3 shows packet losses in relation to the averaged signal strength.

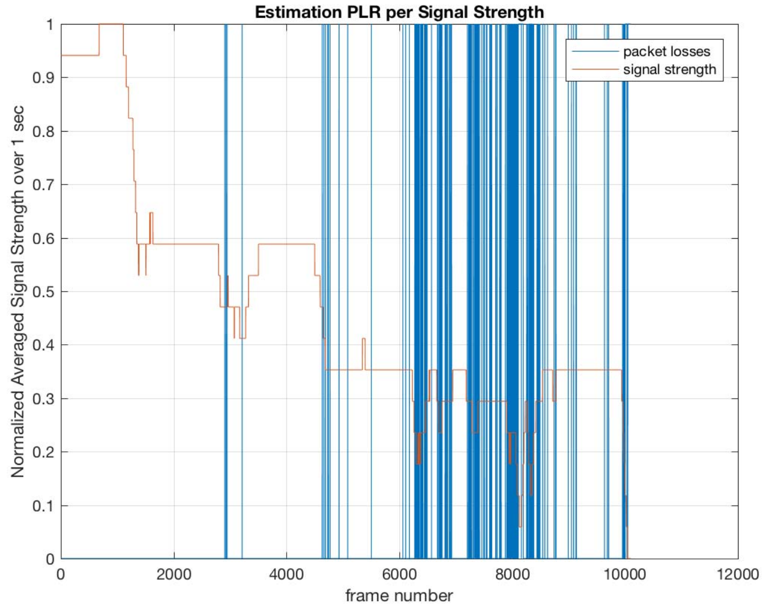

Figure 3: Packet loss rate estimation

According to the given data, a specific PLR can be assigned to a certain signal level as outlined in Table 2.

Table 2: Packet loss and bit error rates

Normalized Averaged signal strength (RSSI)	# Frames	# Packet losses	PLR [%]	PLR rounded [%]	BER rounded [%]
1 (136 dB)	430	0	0	0	
0,94 (128 dB)	728	0	0	0	
0,88 (120 dB)	41	0	0	0	
0,82 (112 dB)	78	0	0	0	
0,76 (104 dB)	20	0	0	0	
0,71 (96 dB)	28	0	0	0	
0,65 (88 dB)	75	0	0	0	
0,59 (80 dB)	2 380	0	0	0	
0,53 (72 dB)	320	0	0	0	
0,47 (64 dB)	436	5	1,15	1	
0,41 (56 dB)	203	2	0,99	1	0,01
0,35 (48 dB)	3 195	28	0,88	1	0,31
0,29 (40 dB)	1 190	88	7,39	7	2,92
0,24 (32 dB)	543	133	24,49	24	17,13

Please note that for realistic simulations packet losses come in addition to any bit errors. Figure 4 shows the cumulative bit error probability as number of bit errors per frame for the relevant RSSIs. This plot provides a compact view of the expected number of bit errors per packet, e.g. at RSSI 48, about 62 % of the packets show no bit error, 70 % show less than 2 errors and 80 % of the packets show less than 4 bit errors.

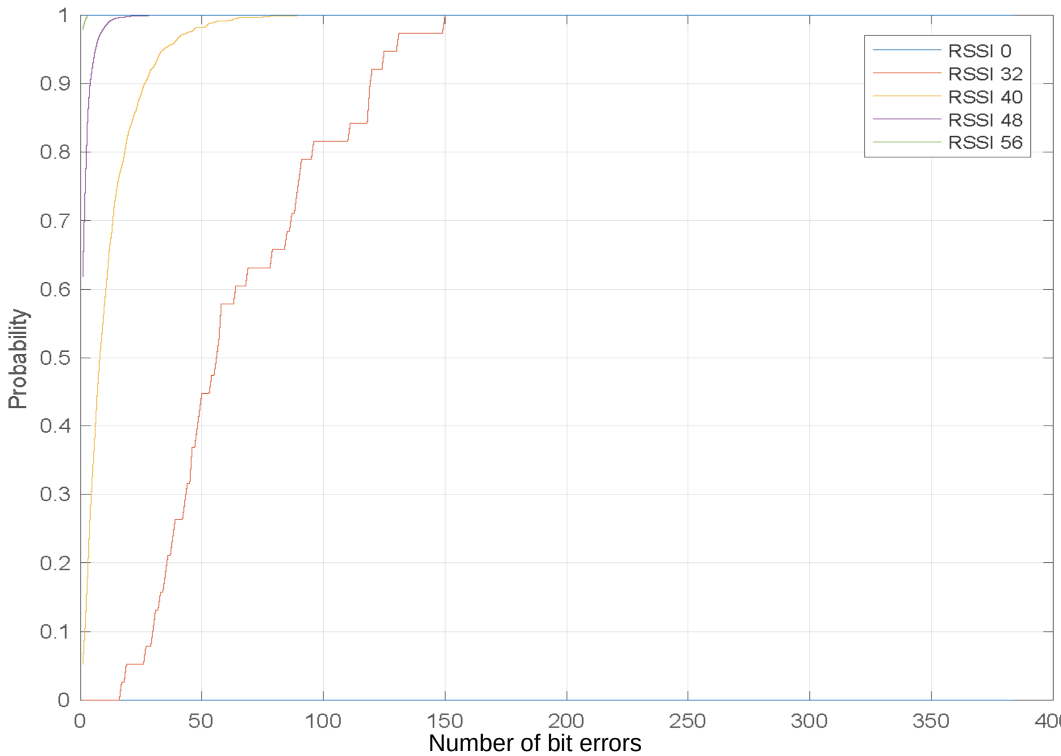

Figure 4: Cumulative probability of bit errors per frame (frame length 40 bytes)

Given the fact that the bit error rate correlates to the signal strength, an adaptively controlled forward error correction code may be beneficial for a DECT system using a block-based audio codec. In order to find the best compromise between voice quality and channel robustness at any signal condition, the codec needs to provide a seamless rate switching scheme to allow the adaptation of codec rate and error protection rate.

NOTE: The parameter signal strength/RSSI might not be usable in a real DECT system and might be replaced by other metrics.

4.3.1.2 Long slots - transmission error profile II

Since a long slot is established based on two normal slots, it is assumed that probability and distribution of bit errors behave similar to those of normal slots.

To prove this assumption, error profiles were derived by Fraunhofer IIS using the following setup:

- An FP (AVM FRITZ!Box 7490) with updated firmware for improved testing and logging capabilities provided by the manufacturer (see note)
- A PP (AVM FRITZ!Fon C4) with updated firmware for improved testing and logging capabilities provided by the manufacturer (see note)
- A computer connected to the PP via LAN, where the following software was used on this computer
- A web browser to configure and monitor the FP
- A soft phone to call the PP
- Wireshark to record the RTP packed G.722 DECT streams (outgoing and incoming)

NOTE: AVM FRITZ!Box 7490 and AVM FRITZ!Fon C4 are examples of suitable products available commercially. This information is given for the convenience of users of the present document and does not constitute an endorsement by ETSI of these products.

In preparation for the measurements, the PP was put into a special mode, in which it responds automatically to an incoming call after 15 s, loops back the incoming signal and terminates the connection after 60 s.

The measurements itself were performed as follows:

- The PP was called from the soft phone.
- A test signal was sent from the soft phone. First, a signal was used where a harmonic part (pitch pipe) and a non-harmonic part (clicks) were sent alternant (which helped for the visual inspection), later a noise signal was used (which allows better cross-correlation).
- The data transfer of the computer was recorded by a network protocol analyser.
- Afterwards the RTP streams containing the G.722 DECT signals (outgoing and incoming), were dumped.
- Dumped streams were subsequently aligned (by correlation measures).
- Statistics were derived with regard to the differences between the outgoing and the incoming stream, which can be treated as transmission errors.
- Successful recording was proved by decoding the streams.

A total of 41 measurements were performed:

- Some test measurements to verify the set-up.
- Some measurements outdoor at different distances.
- Some measurements indoor on a long corridor at different distances.
- Some further measurements indoor on a long corridor but moving some meters into side corridors or rooms.
- Some further measurements indoor, where the PP was moved around.

Figure 5 gives an overview of the cumulative probability of bit errors per frame for the 41 measurements. Note that the derived error statistics reflect two-way transmission.

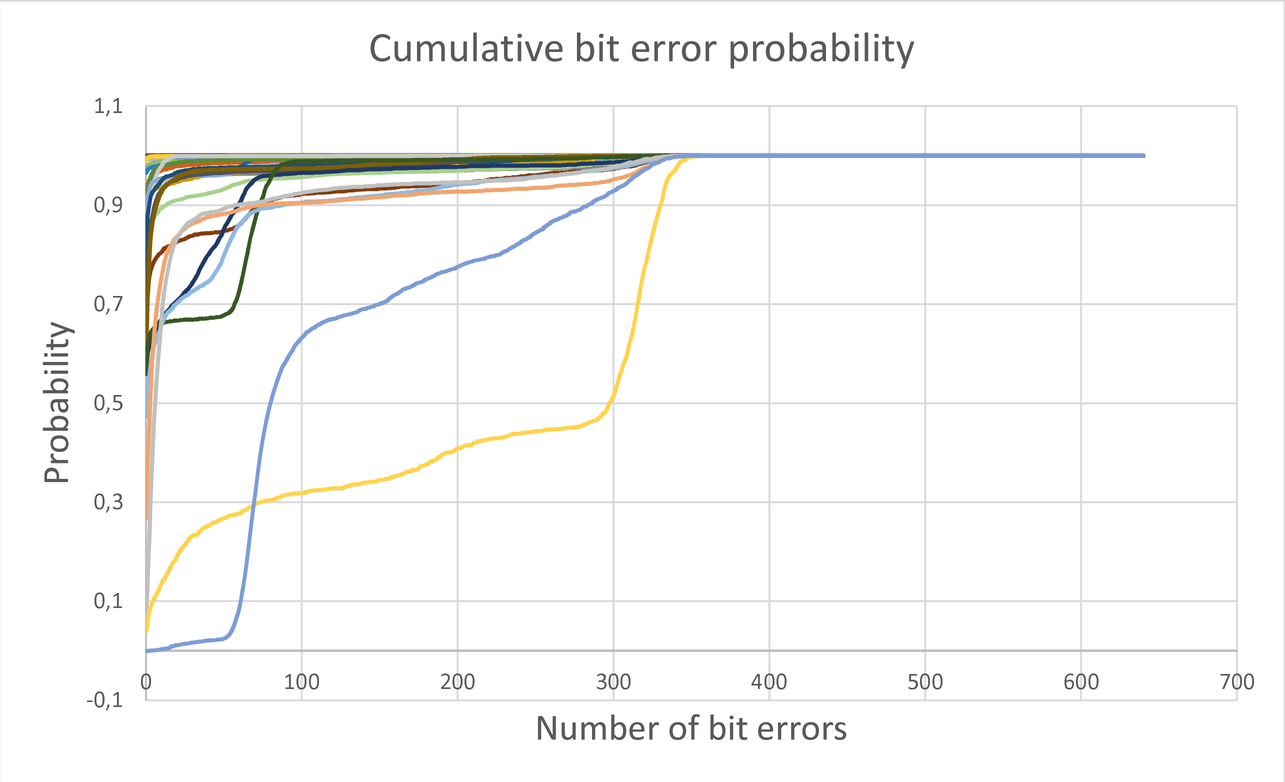

Figure 5: Cumulative probability of bit errors per frame (frame length 80 bytes) - overall figure

Assuming a symmetric channel, one could halve the number of errors for a one-way transmission. The setup provides no option to distinguish between bit errors occurring from PP to FP or from FP to PP.

A legend for the different measurements is left out by intention since it provides no additional information. In fact, there is no strong correlation between the distance and the error rate. Instead, a lot of soft factors (people on the floor, other electromagnetic sources, walls, etc.) seem to have an impact on the channel quality. In contrast to the experiment undertaken for normal slots, no RSSI was available on a frame-by-frame base, thus no RSSI based evaluation can be performed. Having a closer look at the data, a post-screening seems to be appropriate. On one hand there are nine measurements without any error and seven further measurements with an extremely low amount of bit errors. On the other hand, there are two measurements showing a large amount of bit errors per frames, that no meaningful connection is possible anymore. Figure 6 shows just the remaining 23 measurements in a close-up. For comparison reasons, the error profiles described in clause 4.3.1.1 for normal slots are extrapolated to long slots and plotted as well.

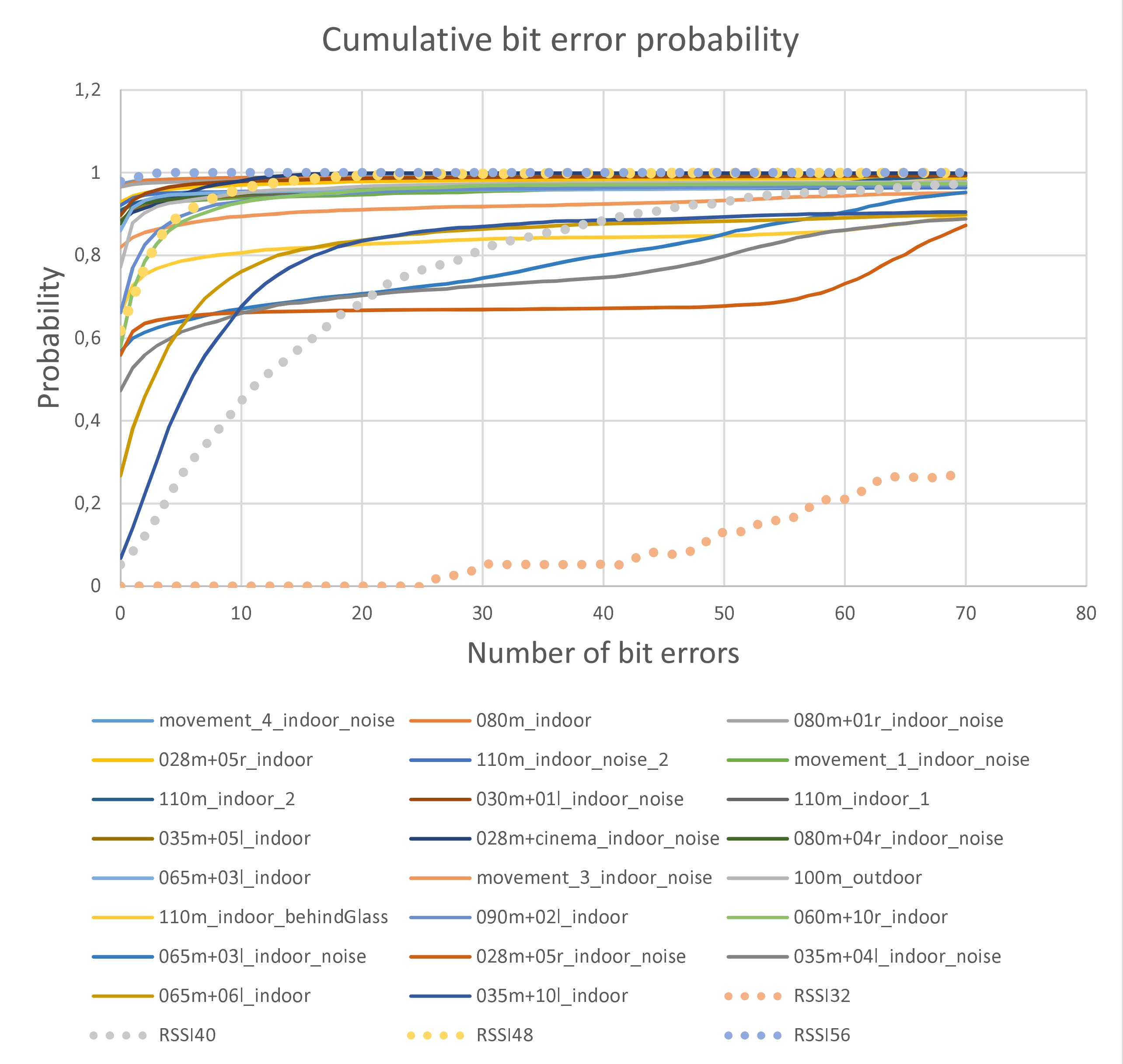

Figure 6: Cumulative probability of bit errors per frame (frame length 80 bytes) - post-screened: Showing only the profiles where FEC may be beneficial

As can be seen from Figure 6, the number of correctly decodable frames can in many situations already be increased significantly by correcting a small number of bits. Afterwards a saturation is visible in almost all measurements, meaning that an increased FEC capability will have almost no positive effect on the number of correctly decodable frames. In some rare cases there is a second area (between 30 and 70 bit errors), in which the number of correctly decodable frames could be increased, but this would require an extremely strong FEC.

In general, the extrapolated error profiles for normal slots show a very similar slope compared to the real measurements for long slots.

4.3.2 Study of CRC/FEC protection schemes

4.3.2.1 Static rates

Assessments based on Recommendation ITU-T P.863 [i.9] were carried out for 18 speech items (clean speech, duration 58 s each). The following encoder configurations were assessed:

- narrowband (8 kHz sampling rate) at 32 kbps
- wideband (16 kHz sampling rate) at 32 kbps
- wideband (16 kHz sampling rate) at 64 kbps
- super wideband (32 kHz sampling rate) at 64 kbps

The following forward error correction schemes were assessed:

- no protection;
- Bose-Chaudhuri-Hocquenghem (BCH) codes with a max correction capability starting from 1 bit and going up to the maximum possible correction rate, which depends on the minimum supported audio payload for the tested configuration; note that the used correction capability is 1 bit less than the max correction capability in order to have a sufficient error detection capability.

Figure 7 shows the LC3 net bitrate relative to the applied forward error protection scheme.

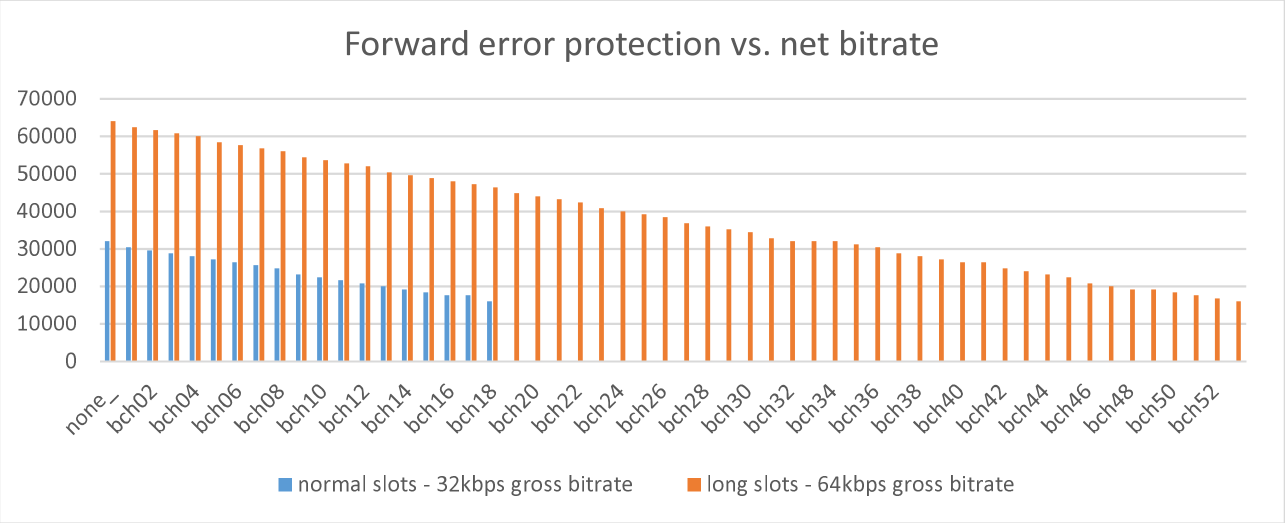

Figure 7: Forward error protection vs. LC3 net bitrate

The following error profiles ware applied (for PLR and BER see Table 2):

- clean
- RSSI 56 dB
- RSSI 48 dB
- RSSI 30 dB

The averaged MOS-LQO scores for the mentioned configurations (as given by the Recommendation ITU-T P.863 [i.9] assessment) are shown in Figure 8, Figure 9, Figure 10 and Figure 11. The bars with different colours refer to the different signal strength classes, indicated by their RSSI. It turns out that a BCH code can improve the perceived audio quality substantially if it is adjusted to the signal strength.

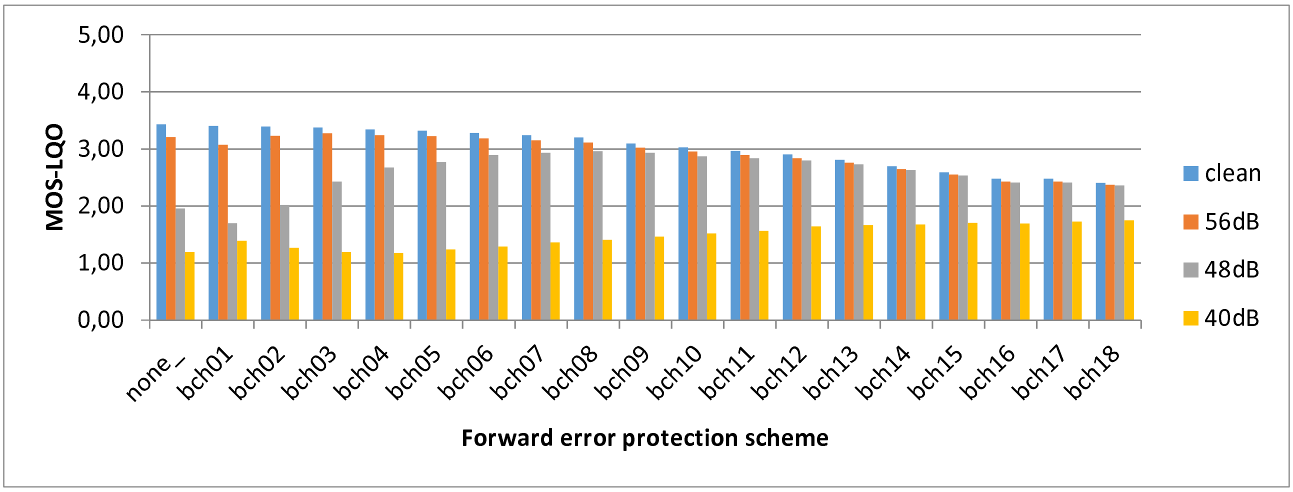

Figure 8: Averaged MOS-LQO scores for NB at 32 kbps gross bitrate

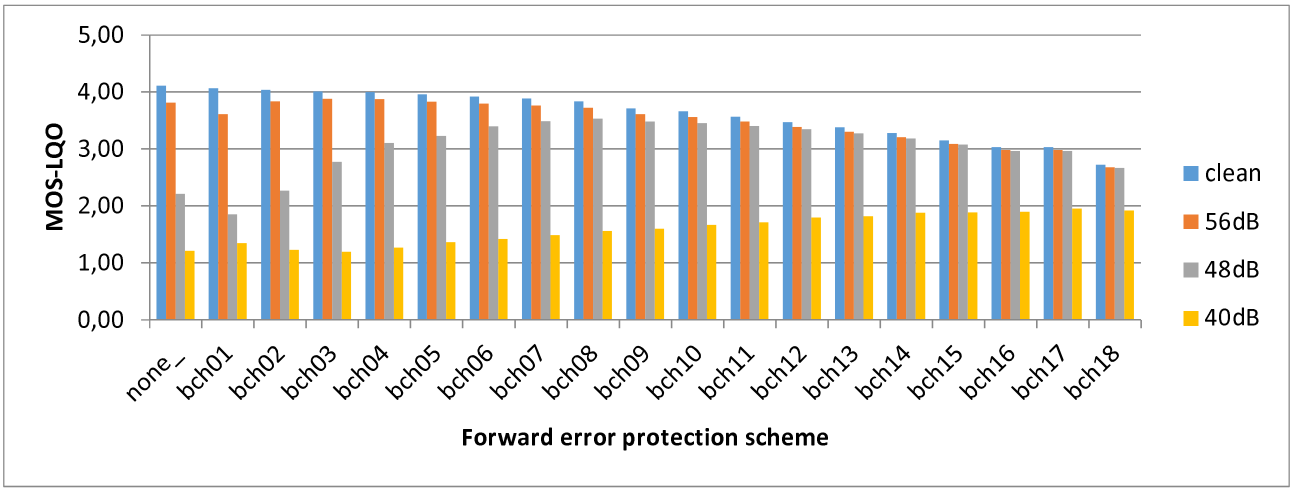

Figure 9: Averaged MOS-LQO scores for WB at 32 kbps gross bitrate

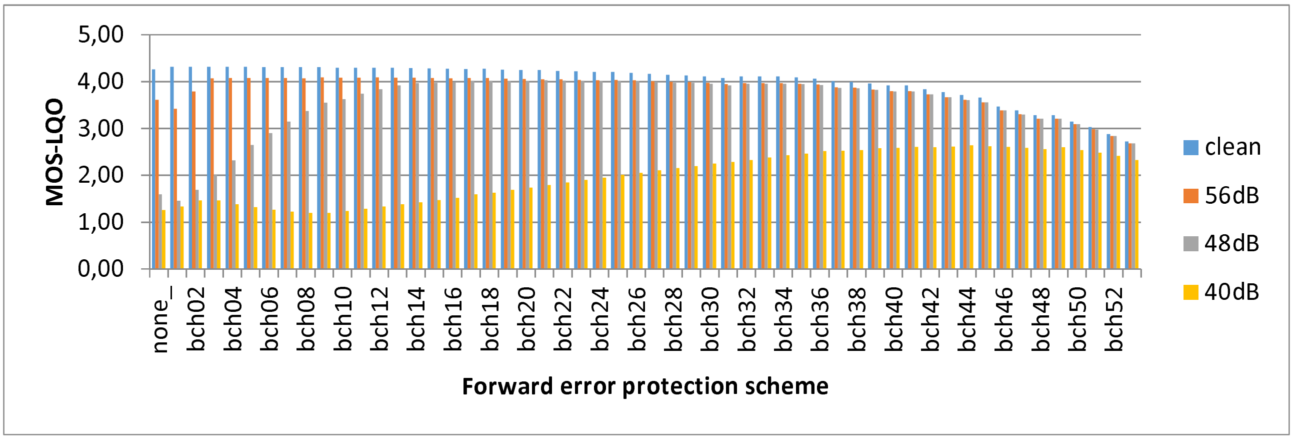

Figure 10: Averaged MOS-LQO scores for WB at 64 kbps gross bitrate

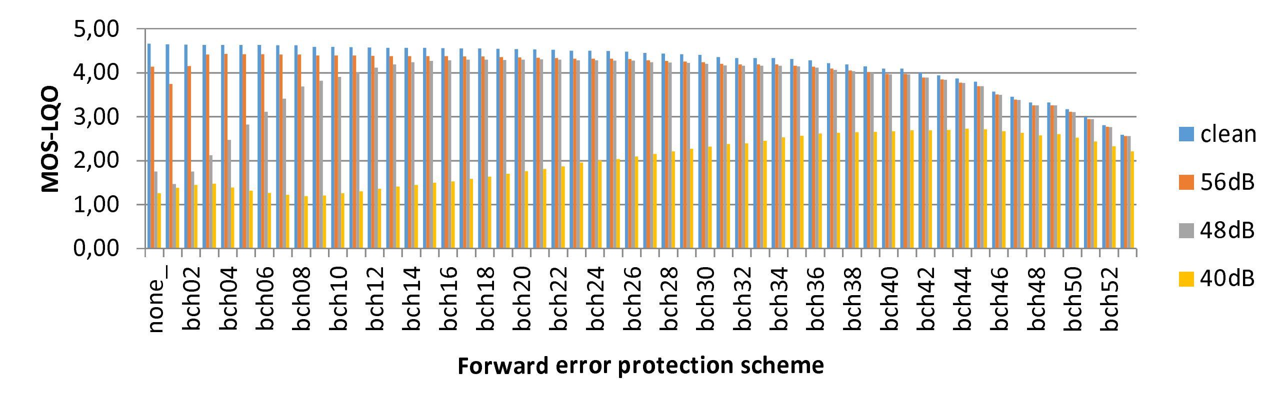

Figure 11: Averaged MOS-LQO scores for SWB at 64 kbps gross bitrate

The sweet spots, i.e. the best objective quality for the different channel conditions relative to the FEC schemes are given in Table 3, Table 4, Table 5 and Table 6.

Table 3: Best objective quality for different channel conditions relative to FEC scheme for NB @ 32 kbps

NB @ 32 kbps	Clean	56 dB	48 dB	40 dB
max MOS-LQO score	3,43	3,28	2,96	1,75
FEC	none	bch03	bch08	bch18
net bitrate	32,0 kbps	28,8 kbps	24,8 kbps	16,0 kbps

Table 4: Best objective quality for different channel conditions relative to FEC scheme for WB @ 32 kbps

WB @ 32 kbps	Clean	56 dB	48 dB	40 dB
max MOS-LQO score	4,11	3,88	3,53	1,95
FEC	none	bch03	bch08	bch17
net bitrate	32,0 kbps	28,8 kbps	24,8 kbps	17,6 kbps

Table 5: Best objective quality for different channel conditions relative to FEC scheme for WB @ 64 kbps

WB @ 64 kbps	Clean	56 dB	48 dB	40 dB
max MOS-LQO score	4,32	4,09	4,03	2,64
FEC	bch01	bch09	bch21	bch44
net bitrate	62,4 kbps	54,4 kbps	43,2 kbps	23,2 kbps

Table 6: Best objective quality for different channel conditions relative to FEC scheme for SWB @ 64 kbps

SWB @ 64 kbps	Clean	56 dB	48 dB	40 dB
max MOS-LQO score	4,66	4,42	4,30	2,73
FEC	none	bch04	bch21	bch44
net bitrate	64,0 kbps	60,0 kbps	43,2 kbps	23,2 kbps

Figure 12 outlines the achievable quality (by means of the MOS-LQO scores) for the investigated codec configurations and signal strengths in a graphical manner.

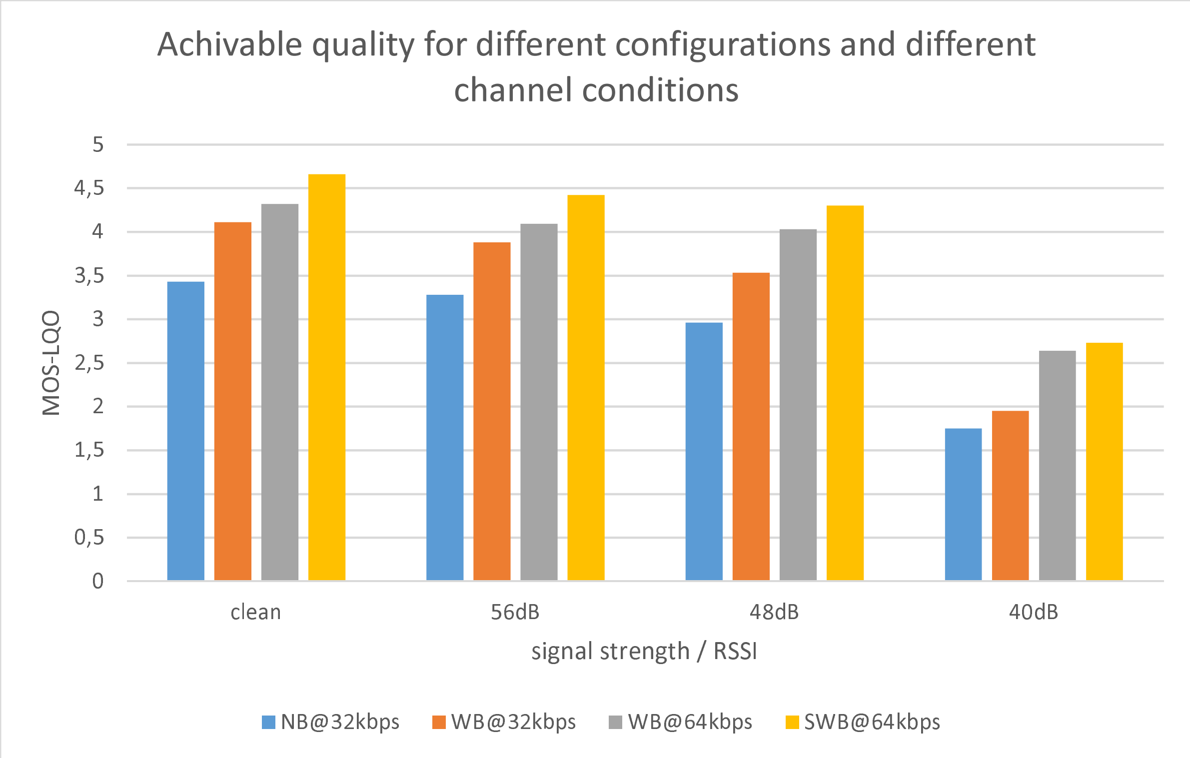

Figure 12: Achievable quality for different configurations at various channel conditions

According to those results, the following recommendations can be made:

- 64 kbps allows for all tested conditions a better quality than 32 kbps.
- For 64 kbps, the SWB mode provides better quality than the WB mode for all channel conditions.
- For 32 kbps, the WB mode provides better quality than the NB mode for all channel conditions.

4.3.2.2 Dynamic rate switching of source and channel coder

4.3.2.2.1 General

Based on the given results, a bit allocation of source and channel coding dependent on the present channel condition is advisable. As a consequence, the source coder needs to support seamless rate switching, as e.g. LC3 does.

4.3.2.2.2 Graceful degradation at DECT range limit

For the tested channel conditions (clean channel, 56 dB, 48 dB, 40 dB), a graceful degradation toward the range limit can be confirmed. For 32 dB, no meaningful output can be achieved anymore.

This behaviour is also confirmed by Recommendation ITU-T P.800 [i.10] experiment documented in clause 4.4.2.3.

4.3.2.2.3 Audio bandwidth switching

According to Figure 12, no bandwidth switching seems advisable for 32 kbps. This is also confirmed by the P.800 [i.10] experiment documented in clause 4.4.2.3.

For 64 kbps a bandwidth switching from SWB to WB may lead to a slight improvement for the 40 dB channel condition.

4.3.2.2.4 Potential channel coder configuration

The DECT channel analysis provided in clause 4.3.2.1 indicates that four error protection (EP) classes could be used for a real DECT system for normal slots. Table 7 outlines a potential configuration of the EP classes for LC3, operating for WB signals at 32 kbps.

Table 7: EP configuration for DECT normal slots

EP class	Codec rate	EP rate	Correction capability	Detection capability (note 1)	Complexity (note 2)
ep_class_1	28 800	3 200	0 bits	99,999 995 %	4,31
ep_class_2	27 200	4 800	2 bits	99,999 995 %	7,77
ep_class_3	24 800	7 200	7 bits	99,999 995 %	7,76
ep_class_4	17 600	14 400	16 bits	99,999 995 %	9,12
NOTE 1: Detection capability is only relevant if bit errors occur. It includes the cases where a correctable frame is marked as un-correctable as well as an un-correctable frame is decoded as a valid frame. NOTE 2: Complexity in million cycles per second measured with an Arm® Cortex®-A9 simulator (simulating the A9MPx1) (see note 3). Code only moderately optimized, especially for low protection classes. NOTE 3: Arm® Cortex®-A9 simulator is an example of a suitable product available commercially. This information is given for the convenience of users of the present document and does not constitute an endorsement by ETSI of this product."					

This EP configuration is also used for the listening experiment documented in clause 4.4.2.3.

4.3.3 Comparison to current DECT codecs

DECT utilizes so far Recommendation ITU-T G.726 [i.13] for NB and Recommendation ITU-T G.722 [i.12] for WB. Table 8 shows the delay and complexity estimations for those codecs within the DECT system and compares them with the parameters of the LC3 codec.

Table 8: Comparison between Recommendation ITU-T G.726, Recommendation ITU-T G.722 and LC3 within the DECT system

	G.726	G.722	LC3 NB	LC3 WB	LC3 SWB
Sampling rate	8 kHz	16 kHz	8 kHz	16 kHz	32 kHz
Framing size (samples)	1	2	80	160	320
DECT framing (samples)	80	160	80	160	320
Algorithmic delay (samples)	1	22	20	40	80
Total delay (samples)	81	182	100	200	400
Framing size (ms)	0,125 ms	0,125 ms	10 ms	10 ms	10 ms
DECT framing (ms)	10 ms	10 ms	10 ms	10 ms	10 ms
Algorithmic delay (ms)	0,125 ms	1,375 ms	2,5 ms	2,5 ms	2,5 ms
Total delay (ms)	10,125 ms	11,375 ms	12,5 ms	12,5 ms	12,5 ms
Complexity (worst frame)	9,2 WMOPS (see note 2)	7,7 WMOPS (see note 1)	7,2 WMOPS	10,5 WMOPS	19,3 WMOPS
Complexity (average)	8,8 WMOPS (see note 2)	7,3 WMOPS (see note 1)	5,9 WMOPS	8,4 WMOPS	14,3 WMOPS
RAM	0,3 kB	1 kB	19,3 kB	19,3 kB	29 kB
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.2] for G.726 and G.722.					

4.3.4 Usability of I_N_minimum_delay service

The minimum delay service defined in ETSI EN 300 175-3 [i.4], I_N_minimum_delay service (I_{NA}) may in general not be used for block/frame based speech coding algorithms.

It is suggested to use the following available DECT slot formats:

- For NB/WB, 32 kbps block-based speech coding: 32 kbps, unprotected: 2 Level GFSK Modulation, Physical packet P32, carrying I_N_normal_delay service (I_{NB}).
- For SWB/FB 64 kbps block-based speech coding: 64 kbps, unprotected: 2 Level GFSK Modulation, Physical packet P64, carrying I_N_normal_delay service (I_{NB}).

4.4 Evaluation results of LC3

4.4.1 Audio quality evaluation for clean channels

4.4.1.1 NB/WB experiment tandeming I

4.4.1.1.1 Setup

The following listening experiment was carried out to verify the performance of the LC3 codec compared to Recommendation ITU-T G.722 [i.12] and mSBC [i.17], mainly in the context of Bluetooth™ scenarios. The experiment aims to assess the self-tandeming capabilities and the interop capabilities to mobile networks. The experiment parameters are given in Table 9.

Table 9: Experiment Recommendation ITU-T P.800 [i.10] NB/WB Tandeming I

Parameter	Value
Type	P.800 ACR
Content	Clean speech, German, 4 Talker, 6 samples each
Listeners	24 naive listeners, German natives
Max. audio bandwidth	WB, 50 Hz - 8 kHz
Lab	Fraunhofer IIS
Transcoders	G.722, LC3, Opus
Mobile Phone	AMR [i.14], AMR-WB [i.15], EVS-WB [i.16]
Self-tandeming	G.722, LC3
Comments	Transcoding scenarios • Phone → Bluetooth™ • Bluetooth™ → Phone • Bluetooth™ → Phone → Bluetooth™

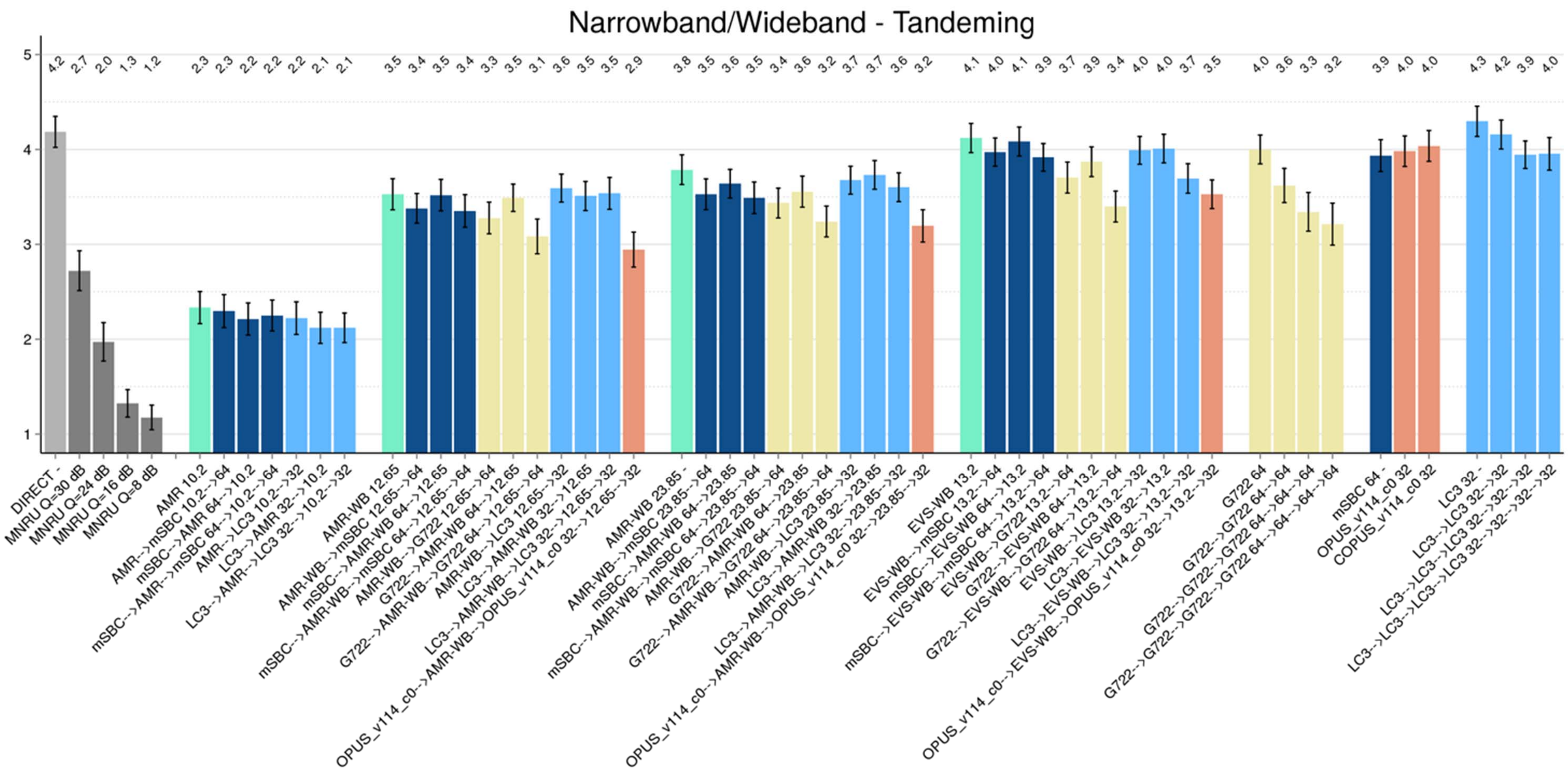

Figure 13: Recommendation ITU-T P.800 [i.10] ACR experiment, WB, clean speech

4.4.1.1.2 Observations

The following observations can be made according to Figure 13:

- LC3 at 32 kbps (LC3 32) provides significantly better audio quality than G.722 at 64 kbps (G722 64).
- LC3 shows much less quality degradation for self-tandeming conditions than G.722 (LC3→LC3→… vs. G722→G722→ …).
- LC3 at 32 kbps (LC3 32) provides significantly better audio quality than Opus-CELT at 32 kbps and complexity level 0 (OPUS_v114_c0 and COPUS_v114_c0).
- LC3 exhibits no problems for interoperability to any mobile network codec condition.

4.4.1.2 SSWB/SWB experiment tandeming I

4.4.1.2.1 Setup

The following listening experiment was carried out to verify the performance of the LC3 codec compared in SWB mode. The experiment aims to assess the self-tandeming capabilities and the interop capabilities to mobile networks. Another aspect is to check the benefit for SWB audio quality compared to legacy DECT WB quality. The experiment parameters are given in Table 10.

Table 10: Experiment Recommendation ITU-T P.800 [i.10] SSWB/SWB Tandeming I

Parameter	Value
Type	P.800 ACR
Content	Clean speech, German, 4 Talker, 6 samples each
Listeners	24 naive listeners, German natives
Max. audio bandwidth	SWB, WB, 50 Hz - 16 kHz
Lab	Fraunhofer IIS
Transcoders	LC3
Mobile Phone	EVS-SWB
Self-tandeming	LC3
References, Anchors	LC3sswb (24 kHz), G.722sswb (24 kHz), G.722 (16 kHz)

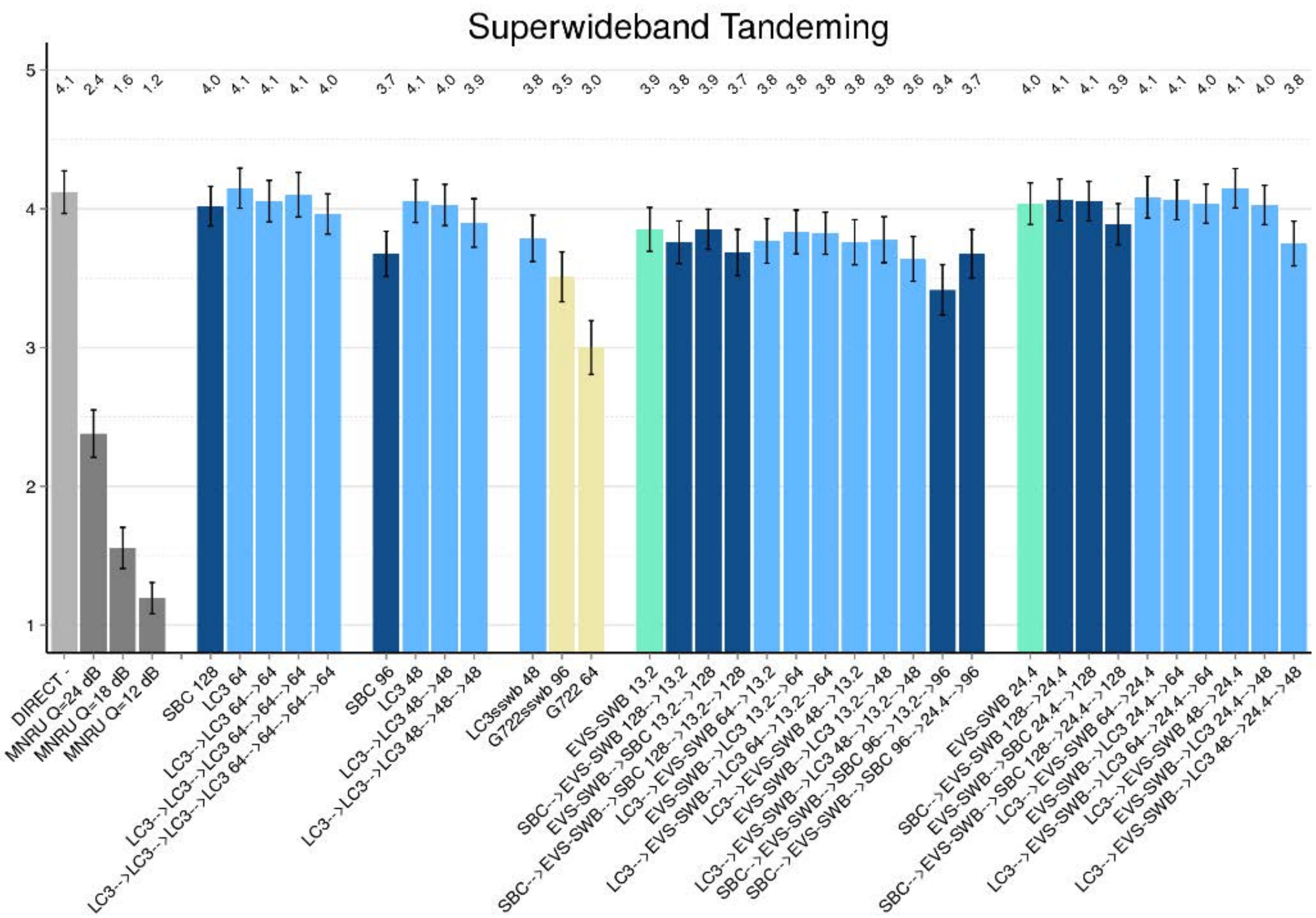

Figure 14: Recommendation ITU-T P.800 [i.10] ACR experiment, SWB, clean speech

4.4.1.2.2 Observations

The following observations can be made according to Figure 14:

- SWB audio bandwidth with LC3 at 64 kbps (LC3 64) or 48 kbps (LC3 48) provides significantly better audio quality than WB audio bandwidth with G.722 at 64 kbps (G722 64).
- LC3 shows almost no quality degradation for self-tandeming conditions (LC3→LC3→…).
- LC3 exhibits no problems for interoperability to any mobile network codec condition.

4.4.1.3 WB/SWB experiment on short frame size

4.4.1.3.1 Setup

The following listening experiment was carried out to compare the performance of LC3 operating at 5 ms frame size to the regular LC3 mode operating at 10 ms frame size. The experiment parameters are given in Table 11.

Table 11: Experiment Recommendation ITU-T P.800 [i.10] WB/SWB on short frames

Parameter	Value
Type	P.800 ACR
Content	Clean speech, German, 4 Talker, 6 samples each
Listeners	23 naive listeners, German natives
Max. audio bandwidth	SWB, 50 Hz - 16 kHz
Lab	Fraunhofer IIS
CuT	LC3 (frame size 5 ms, total delay 7,5 ms)
Reference	• LC3 (frame size 10 ms, total delay 12,5 ms) • EVS 13,2 kbps and 24,4 kbps

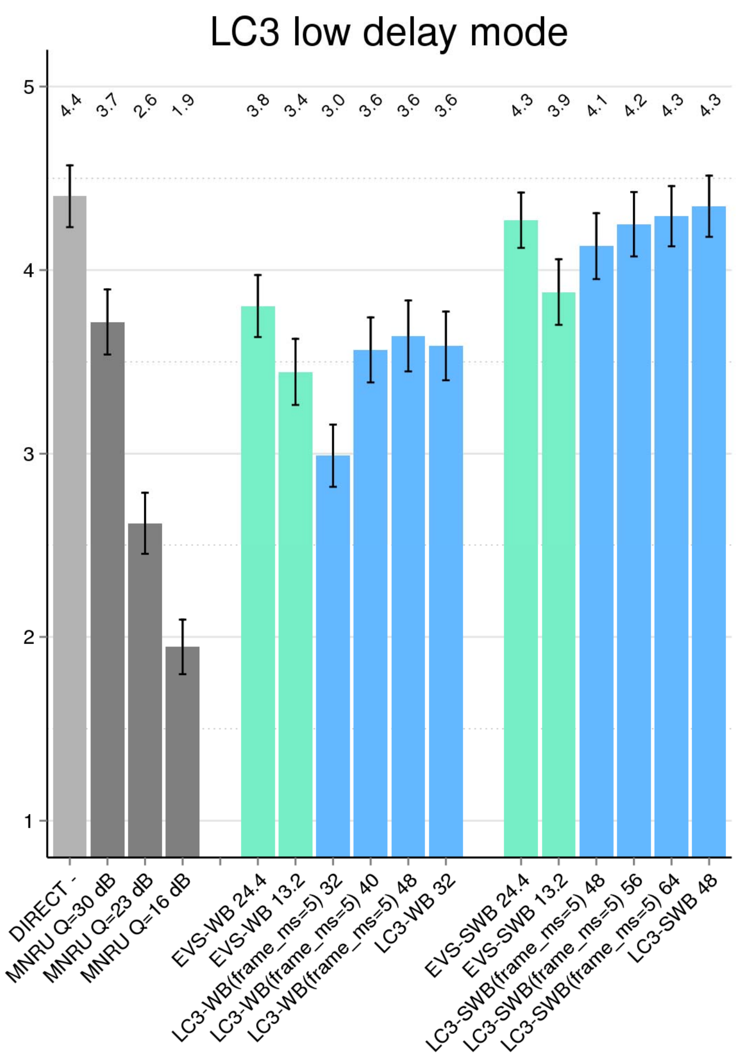

Figure 15: Recommendation ITU-T P.800 [i.10] ACR experiment on 5 ms frame size

4.4.1.3.2 Observations

The following observations can be made according to Figure 15:

- For WB, LC3 operating in 5 ms frame size requires 40 kbps to provide similar quality to the regular LC3 at 32 kbps.
- For SWB, LC3 operating in 5 ms frame size requires 56 kbps to provide similar quality to the regular LC3 at 48 kbps.
- This speech experiment indicates that an additional data rate of 8 kbps is required to compensate for the shorter frame size, i.e. 5 ms instead of 10 ms.

4.4.2 Audio quality evaluation for error prone channels

4.4.2.1 Reference conditions

Table 12 shows the considered reference conditions.

Table 12: Considered reference conditions

codec	Clean channel	Packet loss	Packet loss and bit errors
G.726	X	-	-
G.722	X	• Appendix III • Appendix IV	• Appendix IV • Appendix IV with 1-bit parity • Appendix IV with 2-bit parity
Opus (CELT)	X	X	With FEC similar to LC3

The 1-bit and 2-bit parity are implemented along the lines in EP2901594B1 [i.3], including mute-release means in the presence of multiple subsequent bit errors. While the bitrate available to code the audio signal is 64 kbps for G.722 without any parity bit, it gets reduced to 56 kbps with 1-bit parity and to 48 kbps with 2-bit parity.

4.4.2.2 Codec characterization depending on PLR

4.4.2.2.1 Setup

The following listening experiment was carried out to verify the packet loss concealment performance of LC3 compared to Opus and the EVS codec. The experiment aims to assess the maximum possible packet loss rate for the LC3 codec which still provides a reasonable quality level. The experiment parameters are given in Table 13.

Table 13: Experiment Recommendation ITU-T P.800 [i.10] WB PLC

Parameter	Value
Type	P.800 ACR
Content	Clean speech, German, 2 Talker, 6 samples each
Listeners	24 naive listeners, German natives
Max. audio bandwidth	WB, 50 Hz - 8 kHz
Lab	Fraunhofer IIS
Packet loss rates (PLR)	0 %, 1 %, 3 %, 6 %, 10 % (synchronized error pattern across codecs)
Codecs	LC3 @ 32 kbps, Opus @ 32 kbps, EVS @ 24,4 kbps
LC3 PLC methods	Standard and advanced
Anchors	LC3 with muting as PLC (PLC_meth=5), MNRU 10 dB and 20 dB

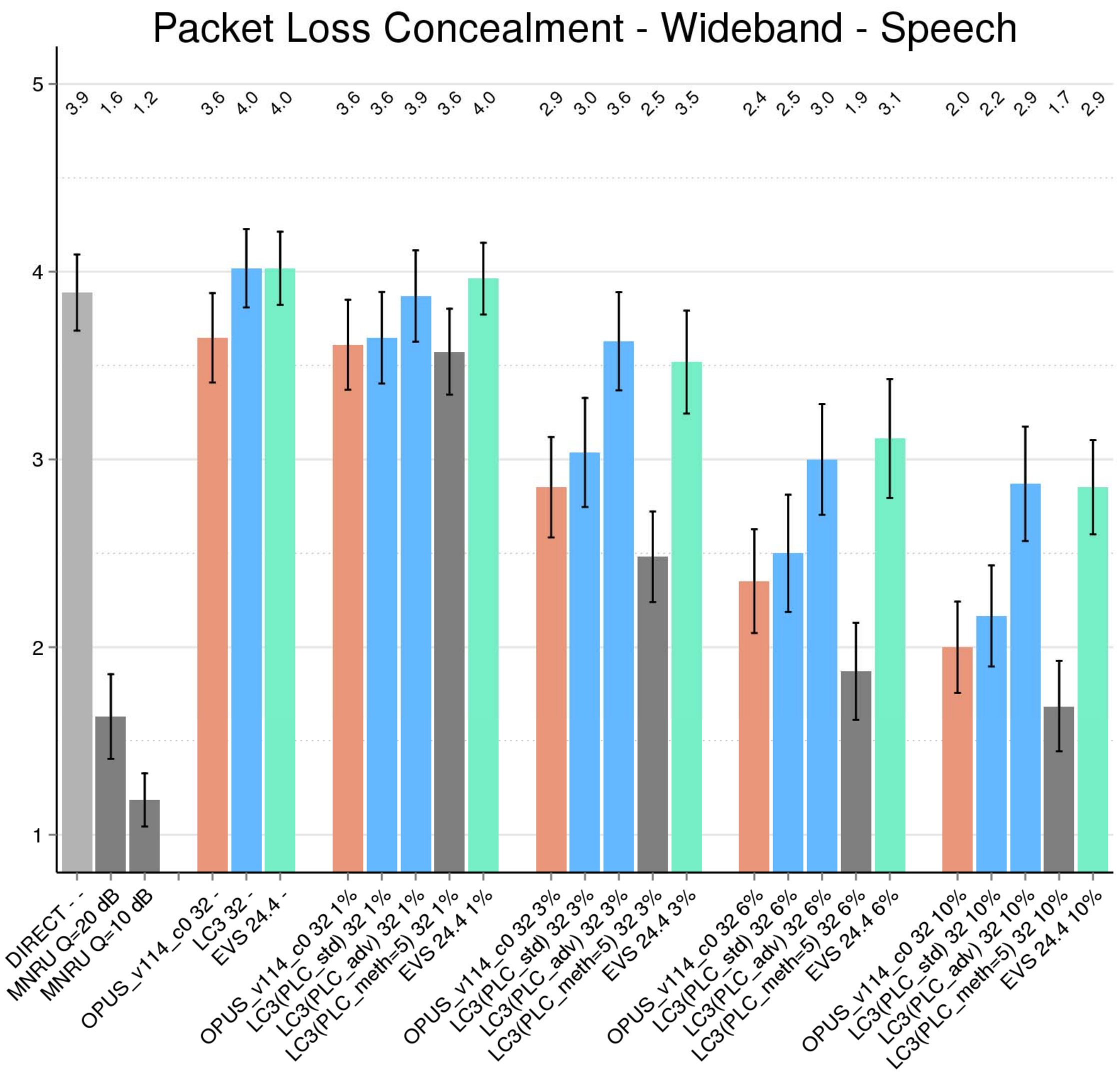

Figure 16: Recommendation ITU-T P.800 [i.10] ACR experiment on packet loss conditions

4.4.2.2.2 Observations

The following observation can be made according to Figure 16:

- The standard concealment method of LC3 provides the same audio quality compared to Opus at any PLR.
- The advanced concealment method of LC3 provides significantly better quality compared to Opus and the standard PLC method at any PLR and performs similar to the EVS concealment.

4.4.2.3 Codec characterization depending on PLR and BER

4.4.2.3.1 Setup

The following listening experiment was carried out to verify the performance of LC3 compared to G.722 and Opus over simulated DECT channels. On one hand this experiment compares the performance of LC3 to Recommendation ITU-T G.722 [i.12] Appendix III and Appendix IV for frame loss. On the other hand, it evaluates the performance of LC3 compared to Recommendation ITU-T G.722 [i.12] and Opus for DECT channels. For DECT channels, LC3 and Opus payloads are protected by the BCH coding schemes. Furthermore, the NB quality of LC3 is verified in comparison to Recommendation ITU-T G.726 [i.13]. The experiment parameters are given in Table 14.

Table 14: Experiment Recommendation ITU-T P.800 [i.10] WB DECT

Parameter	Value
Type	P.800 ACR
Content	Clean speech, German, 2 Talker, 6 samples each
Listeners	23 naive listeners, German natives
Max. audio bandwidth	WB, 50 Hz - 8 kHz
Lab	Fraunhofer IIS
Packet loss rates (PLR)	0 %, 1 %, 3 %, 6 %, 10 % (synchronized error pattern across codecs)
RSSI	40, 48, 56, 80 (see Table 2 for PLR and BER)
Codecs	LC3 @ 32 kbps, Opus @ 32 kbps, G.726 @ 32 kbps, G.722 Appendix IV @ 64 kbps (regular, 1-bit parity, 2-bit parity)
Error Protection schemes	epmode 1..4 according to EP class definition in Table 7

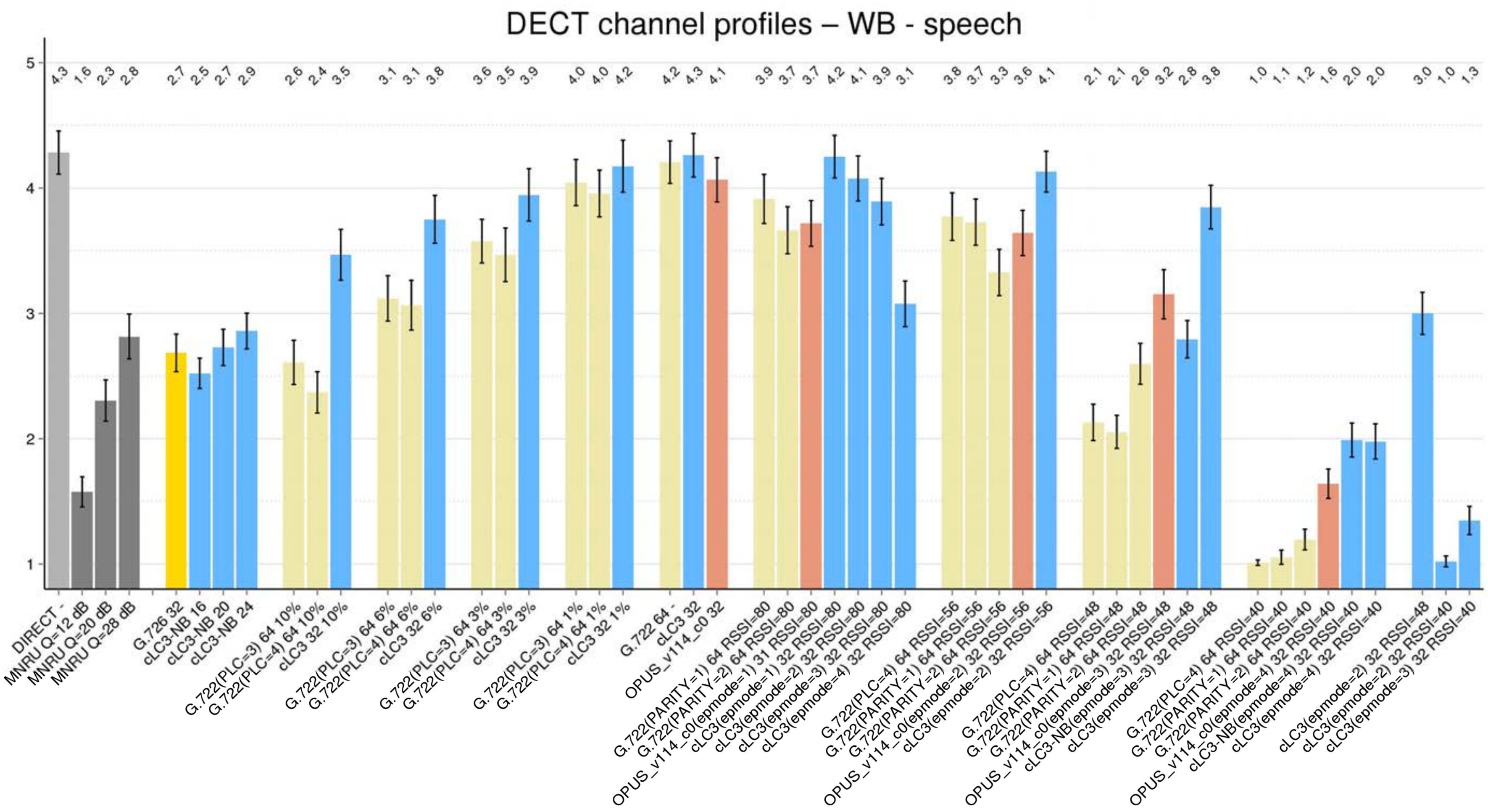

Figure 17: Recommendation ITU-T P.800 [i.10] ACR WB test on packet loss and DECT channel conditions

4.4.2.3.2 Observations

The following observations can be made according to Figure 17:

- For NB, LC3 requires 20 kbps to provide similar audio quality compared to G.726 operating at 32 kbps.
- For PLC, LC3 provides significantly better performance for lost frame concealment compared to G.722 Appendix III and IV.
- LC3 performs significantly better than G.722, G.722 with 1 parity, G.722 with 2 bit parity and Opus for each individual DECT signal strength, i.e. RSSI 80, 56, 48, 40, for optimal EP class usage.
- At RSSI 40, all G.722 variants and Opus show a quality below any usable service requirements. Only LC3 might provide a service quality good enough to continue the call under extreme error conditions.

4.5 Integration of LC3 in existing DECT infrastructure

4.5.1 General

This clause outlines a number of issues that need to be addressed when introducing a new codec in the DECT specifications. In most cases there are several alternatives for how an issue may be resolved. The clauses below discuss the foreseen alternatives and, where possible, propose which alternative should be selected.

The current description assumes that the LC3 codec is used in 2-way communication, e.g. traditional telephony use cases. One-way communication, e.g. listening to a presentation and streaming use cases, is not yet considered.

4.5.2 High-level codec description

4.5.2.1 Codec overview

This clause gives a high-level description of the Low Complexity Communication Codec (LC3). LC3 obtains very high audio quality at medium bitrates utilizing a low computational complexity. The main features of LC3 are as follows.

Table 15: Codec summary

Feature	Supported Range
Frame length	10 ms or 5 ms
Inherent algorithmic delay	2,5 ms
Total algorithmic delay	12,5 ms or 7,5 ms
Supported sampling rates	8, 16, 24, 32 and 48 kHz
Supported bitrates	16 - 320 kbps, 0,8 kbps increments
Supported bits per audio sample	16 or 24 bits per sample

4.5.2.2 Audio channels

LC3 encodes a single audio cannel. Stereo or multi-channel coding is supported with multiple mono streams.

4.5.2.3 Complexity

Complexity is described in clause 4.3.3.

4.5.2.4 Audio bandwidth detection

The LC3 encoder has a built-in audio bandwidth detector which determines if the input signal is bandwidth limited to frequencies lower than the Nyquist frequency. The encoder may decide, on a frame-by-frame basis, to encode only a lower audio bandwidth, thereby spending the bits on only the relevant signal components, e.g. the speech. The remaining frequencies are instead re-inserted in the decoder by generating low-level noise.

For example, when operating at 32 kHz sampling frequency (0 - 16 kHz audio bandwidth), if the audio bandwidth detector finds that the input signal contains mainly speech at lower frequencies, then the encoder may encode frames with WB or even NB.

This audio bandwidth switching is only related to the internal processing of the codec. It does not affect the audio front-end processing such as the microphone, speakers, echo cancellation, etc.

4.5.3 High-level options for LC3 integration into the DECT specifications

4.5.3.1 Fixed bitrate, multi-mode or adaptive multi-mode operation

The flexibility of the LC3 codec means that it is possible to use the codec in several ways:

A) Fixed mode rate operation, where only a single audio bandwidth range and a single bitrate are used for the entire session.

B) Multi-mode operation, where the audio bandwidth and/or bitrate are changed rarely, for example when switching channel or base station. Changing audio bandwidth and/or bitrate requires re-configuring the connection, similar to performing a channel re-selection or handover. Frequent channel re-configuration is usually not possible because this would increase the control signalling too much.

C) Adaptive multi-mode operation, where the audio bandwidth and/or bitrate may change at virtually any time during the call, even from frame to frame. The receiver monitors the performance (e.g. the Frame Erasure Rate, (FER)), determines if it would be better to switch to a different audio bandwidth and/or bitrate and sends a Codec Mode Request (CMR) back to the sender. When the sender receives the CMR it switches to the requested mode as soon as possible, usually at the next frame border.

The main pros and cons with the different alternatives are summarized in Table 16.

Table 16: Summary of advantages and disadvantages

Alternative	Pros	Cons
A (fixed mode operation)	Simple implementation and verification No bits need to be allocated to Frame Type indication and Codec Mode Request	Not possible to optimize performance for different and/or varying operating conditions Transcoding always required when connecting systems using different modes
B (multi-mode operation)	Can optimize quality to different but more or less stationary operating conditions With proper codec configuration, transcoding can be avoided when connecting different systems using different modes	Cannot optimize quality to varying operating conditions Some bits need to be reserved for inband Frame Type indication Additional Layer 2 signalling for Codec Mode Request
C (adaptive multi-mode operation)	Possible to optimize quality for different and varying operating conditions With proper codec configuration, transcoding can be avoided when connecting different systems using different modes	Some bits need to be reserved for codec mode indication and codec mode request, i.e. some overhead Some bits may need to be reserved for inband Frame Type indication Some bits need to be reserved for inband Codec Mode Request

The evaluation of different combinations of codec bitrate and channel protection schemes in clause 4.3 indicates that a suitable compromise between optimal performance for all channel conditions and complexity would be to use LC3 with 4 modes:

- Clean channel: All bits used for the codec, no bits used for channel protection.

- Good channel (~56 dB): Approximately 90:10 percent split between bits used for speech coding and bits used for channel protection.
- Degraded channel (~48 dB): Approximately 75:25 percent split between bits used for speech coding and bits used for channel protection.
- Poor channel (~40 dB): Approximately 50:50 percent split between bits used for speech coding and bits used for channel protection.

The principle of using rate adaptation to optimize the quality for different and varying channel conditions is similar to how the AMR and AMR-WB codecs are used in GSM.

Proposal:

For best overall performance, LC3 should be used with several codec modes/rates. About 4 modes/rates seem to be sufficient to cover the foreseeable range of channel conditions that may happen.

Exact bitrate allocation for speech encoding and channel encoding, respectively, needs to be determined for the final specification.

4.5.3.2 Codec negotiation at session setup

Codec selection is performed at connection setup by first indicating which codecs are supported and then FP decides which codec is to be used. A Codec List is defined in ETSI EN 300 175-5 [i.7] which includes code points for each defined codec.

Code points in the codec list need to be defined for the LC3 codec. It is probably suitable to define one code point for each audio bandwidth.

Table 17: Codec identifiers for LC3 (example)

Codec identifier	Meaning
<tbd>	LC3 NB
<tbd>	LC3 WB
<tbd>	LC3 SSWB
<tbd>	LC3 SWB
<tbd>	LC3 FBCD (see note)
<tbd>	LC3 FB
NOTE: The mode FBCD (44,1 kHz sampling rate) might not be supported in DECT.	

A single Codec Identifier may also be sufficient. However, in that case, the sampling rate would either have to be defined with other session setup signalling or it would need to be signalled for every speech frame. It is expected that the sampling rate will not change very frequently during a session, if at all. Therefore, signalling at session setup is preferable.

If the audio bandwidth needs to be changed during an ongoing session, for whatever reason, then a session re-negotiation or channel re-selection should be sufficient.

Proposal:

One or several Codec Identifiers need to be defined. It is likely suitable to define different codec identifiers for different audio bandwidths.

Codec negotiation for future air interface requires more details on the session negotiation procedures that will be used.

4.5.3.3 Codec mode set negotiation

To be able to use a multi-mode or multi-rate codec in a session, it is necessary to define which modes/rates are allowed for that session. This can, in general, be done in a number of ways, for example:

A) The specification decides which modes/rates are to be used.

B) Session setup signalling decides which modes/rates are to be used.

C) The specification decides the maximum mode/rate allowed for the session, i.e. all lower modes/rates are allowed.

D) Session setup signalling decides the maximum mode/rate allowed for the session, i.e. all lower modes/rates are allowed.

Options A and C are not very flexible and require good understanding of the channel conditions already before the specification is frozen. Options B and D are more flexible and can be modified for different scenarios that are not known at the time of preparing the specification.

Options C and D enables transcoding-free interworking between different network configurations, even if the networks have different maximum modes/rates. With options A and B, transcoding would be necessary unless the network configurations are identical or if one configuration is a sub-set of the other configuration.

For options C and D, it may also be challenging to design channel coding for every single bitrate given that the bitrate granularity is 0,8 kbps, which would result in a huge number of channel coding schemes.

For options C and D, it should be noted that the sender may not need to support every single bitrate below the maximum allowed bitrate, but the receiver should support receiving and decoding of any bitrate. Development and verification for these options may therefore be challenging.

When deploying a system with several codec modes, the number of codec modes that should be used may also depend on the deployment environment. For example, in a home environment with few users and short distances between the FP and the PPs, then the interference is likely low, and it may be sufficient with just one or two codec modes. However, for an office environment with many users, the interference level may be significant, and it may then be necessary to use more codec modes to allow for adaptation over a larger bitrate range.

The suitable operating range for each codec mode also depends on the receiver performance, which may be different for different vendors. For these reason, it is beneficial if the adaptation thresholds are also negotiated at session setup.

The adaptation thresholds should also use hysteresis to avoid too frequent mode changes. An example of this is shown in Table 18.

Table 18: Example of adaptation thresholds

Operating bitrate	Threshold for switch to next lower bitrate	Threshold for switching to next higher bitrate
Bitrate 1 (highest)	Thr1	N/A
Bitrate 2	Thr2	Thr1 + hyst1
Bitrate 3	Thr3	Thr2 + hyst2
Bitrate 4 (lowest)	N/A	Thr3 + hyst3

There are several options for which performance metrics should be used to determine when adaptation should be made. Some examples are: estimating the channel quality (C/I, C/N); calculating the Frame Erasure Rate (FER); or estimating the Bit Error Rate (BER).

Proposal:

The best solution is to negotiate the allowed codec mode set at session setup, i.e. option B or D. New signalling is likely needed.

The session setup signalling should include adaptation thresholds when the PP is expected to switch from one mode to the next. The adaptation thresholds should use hysteresis.

Codec mode set negotiation for future air interface requires more details on the session negotiation procedures that will be used.

4.5.3.4 Adaptation signalling

When using an adaptive scheme, it is necessary to signal both (similar to how it is done for AMR, AMR-WB and EVS codecs):

- Frame Type (FT), indicating which mode (codec bitrate and channel protection scheme) that was used. This needs to be signalled for each frame in the forward direction (from sender to receiver).

- Codec Mode Request (CMR), receiver indicates which mode it wants to receive. For fastest adaptation and best robustness, this needs to be signalled in each frame in the backward direction (from receiver to sender).

This Frame Type signalling needs to be significantly better protected than the speech data bits since errors in the FT will result in erroneous decoding of the frame.

The Codec Mode Request also needs to be robust, but this may be solved by repeating the CMR in every frame.

Experiences from AMR and AMR-WB on GSM channels suggests that, assuming a 2-bit FT or CMR, about 1 byte (8 bits) should be sufficient for the FT and CMR signalling, respectively.

It may be possible to send FT and CMR in the A or Z fields, thus allowing the B field to be used entirely for the speech data and the associated channel encoding.

It may also be possible to send CMRs in lower-layer (network) signalling, and not within the codec data. However, this likely means that the frequency of the CMR signalling cannot be very high, which gives slower adaptation to varying channel conditions.

Higher order modulation may also be beneficial because this would give more flexibility to adjust the channel coding used for FT/CMR and speech data.

Proposal:

It should be studied if the A or Z fields can be used for CMR signalling and possibly also for FT signalling.

4.5.3.5 Audio bandwidth selection/adaptation

The audio bandwidth, e.g. WB or SWB, can probably be fixed during the session. If the audio bandwidth needs to be changed then a channel re-negotiation (channel re-selection, handover) could be used.

A more advanced solution would be to have adaptation also for the audio bandwidth, similar to what is done in the EVS codec.

Audio bandwidth adaptation may be more interesting when interworking with other systems, e.g. Bluetooth™, fixed networks and mobile networks.

Adapting the audio bandwidth during a session does not necessarily mean that the entire audio chain needs to be reconfigured, e.g. to use 16 kHz sampling frequency instead of 32 kHz, since a low-pass filter could be used instead of changing the sampling frequency. This is however an implementation consideration.

Proposal:

For current air interface, the audio bandwidth can be selected at session setup. If the audio bandwidth needs to be changed during an ongoing session, then a session re-negotiation should be sufficient.

For future air interface, audio bandwidth adaptation during an ongoing session may be more interesting.

4.5.3.6 DTX

In CS systems, DTX is typically used for two reasons:

- Reducing the interference level and thereby improving the performance for other users.
- Reducing the average complexity and increasing the battery lifetime.

In VoIP, there is a third reason for using DTX. The silence/background noise periods offer an opportunity to adjust the play-out time since it is very easy to adjust. The Comfort Noise Generation can easily generate additional noise frames to delay the play-out time for the subsequent frames, thereby creating additional margin for large packet jitter. Correspondingly, the CNG can also generate fewer noise frames to gain back delay when detecting that the packet jitter has been reduced.

There are more advanced, and more complex, methods for doing jitter buffer managements. However, DTX offers a simple way for doing jitter buffer management in a low-complex way that is usually quite sufficient.

The usage of DTX may be different for different installations. For example, for a home environment, DTX may be unnecessary. However, for an office environment with many users it may be very important to reduce the interference level and then DTX may be highly desirable. It may therefore be desirable to be able to signal at session setup whether DTX should or should not be used.

Proposal:

For the current air interface, the use of DTX may perhaps not be very important. However, DTX is an interesting feature for future VoIP applications, e.g. in DECT-2020.

Session setup signalling may be needed for enabling or disabling DTX.

4.5.3.7 Acoustic front-end

To use the LC3 codec efficiently and to ensure interworking with external systems, it is necessary to ensure that the acoustic front-end is well-designed and suitable for SWB. ETSI EN 300 176-2 [i.8] defines audio testing for NB and WB but SWB is to a large degree left as "For Future Studies".

New test procedures need to be defined for SWB. It is likely suitable to base this work on similar work done in other standardization organizations, for example 3GPP/SA4 and ITU-T.

Proposal:

Audio testing for SWB should be defined.

4.5.3.8 VoIP, RTP payload format

An RTP payload format needs to be developed for VoIP solutions in DECT and for interworking between DECT systems over backbone IP networks. It is foreseen that this will be standardized in IETF.

Since the codec is very flexible, the RTP payload format needs to support all possible variants, including: all audio bandwidths, all possible bitrates, adaptation, multiple channels, etc.

It is important to ensure that the RTP payload format is designed such that the codec usage in DECT is efficiently supported.

Proposal:

An RTP payload format description should be developed.

4.5.3.9 Backwards compatibility

Backwards compatibility and coexistence with legacy equipment is always important when introducing a new major feature like a new codec. Maintaining legacy codecs and implementing proper codec negotiation at session setup negotiation and re-negotiation should solve these issues.

It is however important to ensure that the introduction of LC3 does not increase control plane signalling too much as this could starve out other devices from using the system.

Proposal:

Careful design of the control plane signalling needed for LC3 is necessary.

4.5.3.10 Interworking with external networks and devices

As long as transcoding is used for communication to external networks and devices, no particular issues are foreseen. However, matching the audio bandwidth used by the LC3 codec may be necessary.

When receiving speech from an external network or device, the LC3 codec will use the built-in audio bandwidth detector (see codec description above) to avoid wasting bits for non-existing frequencies. I.e. even if the DECT session is setup to use SWB but the incoming speech is only WB, then the LC3 codec will encode only WB.

When sending speech to an external network or device, if the DECT session is setup to use SWB but the external system uses only WB, then LC3 will encode SWB but the higher frequencies will be filtered away when re-encoding to the externally used codec. Therefore, if the audio bandwidth is not matched with the external system, then LC3 will likely waste some bits on encoding frequencies that will just be removed in the transcoding process, which leaves fewer bits for encoding those frequencies that will actually be heard by the remote user. Hence, if the audio bandwidths are not properly matched then this may result in sub-optimal quality.

This may be an argument for audio bandwidth adaptation during an ongoing session. However, proper audio bandwidth matching at session setup and session re-negotiation should be sufficient to handle most interworking scenarios.

Proposal:

The audio bandwidth should not be permanently configured for the system but rather configured at session setup.

4.5.3.11 Repeater operation, relays

When Wireless Relay Stations (WRS) are used then the application layer data is sent transparently through the WRS and the WRS cannot know which codec that is used, see clause 7.4.7 of ETSI EN 300 700 [i.6]. This offers a simple method for introducing the LC3 codec with rate adaptation where the sum of the speech codec bits and the channel codec bits add up to a fixed bitrate. The bit errors will then accumulate for every hop in the path between the FP and the PP and the receiver should adapt in the same way as if the same amount of bit errors would occur for a single hop.

A more advanced solution is to perform channel encoding and decoding on a per-hop basis. This may offer improved performance for hops with few bit error that could be fully corrected before forwarding the frame. However, this may break the layering principle and may have significant implementation impacts.

Proposal:

For simplicity, it is suggested to treat the LC3 speech codec bits and the LC3 channel codec bits as application layer data and to send this transparently through the WRS.

A more advanced solution may be considered at a later stage.

5 Conclusions

The present document studies the benefits of a potential DECT upgrade to the LC3 codec in order to support higher audio bandwidth and maximize the DECT capacity for legacy transport slots. This includes a characterization of the DECT channels and complementary solutions for robust channel protection schemes applicable for block-based codecs.

In summary, the study identifies the following potential improvements for DECT systems:

- For WB calls transmitted over normal DECT slots, LC3 is able to provide the same or better audio quality compared to the legacy G.722 codec transmitted over long slots. In this case, the DECT capacity can be doubled (see clause 4.4.2.3). This holds true for all tested channel conditions.
- The superior SWB quality compared to WB quality has been clearly pointed out (see clause 4.4.1.2).
- LC3 provides a significant better performing concealment for packet loss conditions compared to G.722 (see clause 4.4.2.3). The concealment is comparable to state-of-the-art communication codecs like EVS (see clause 4.4.2.2). This makes the codec future proof for the next packet based DECT generation.
- LC3 in combination with a proposed channel protection scheme provides an acceptable speech quality even under extreme weak channel conditions where G.722 based systems fail (see clause 4.4.2.3). This indicates that the coverage range of DECT systems can possible be extended with LC3.
- Regarding tandeming operation with other legacy mobile or fixed line codecs, several listening experiments have shown that LC3 is able to transcode calls without any significant quality degradation (see clauses 4.4.1.1 and 4.4.1.2). For self-tandeming conditions, LC3 shows a more robust behaviour than G.722 (see clause 4.4.1.1) in WB conditions. For SWB conditions, LC3 shows a very robust behaviour for self-tandeming as well.

- LC3 enables low delay applications by offering 5 ms frame size coding at a slightly increased bit budget (see clause 4.4.1.3).

In order to enable the listed improvements above, the following specification updates and work packages need to be completed:

1) Codec standardization of LC3:
   a) Encoder/Decoder specification
   b) LC3 Source Code
   c) Test vectors and conformance
   d) Codec characterization
2) Codec integration into DECT:
   a) Change Requests to DECT specifications (ETSI EN 300 175, applicable parts need to be determined)
   b) Change Requests to DECT test specifications (ETSI EN 300 176, applicable parts need to be determined)
   c) Change Requests to New Generation DECT specifications (ETSI TS 102 527, applicable parts need to be determined)
   d) DECT Channel coder
   e) Codec configuration
   f) Codec signalling mechanism

Annex A: Change History

Date	Version	Information about changes
21st August 2017	0.0.2	Document skeleton
30th November 2017	0.0.3	Provided error profiles for normal slots, listening results for base quality of LC3 in WB and SWB and for packet loss conditions
1st March 2018	0.0.4	Completed normal slot investigations and low delay audio quality check; completed information about DECT error profiles
8th March 2018	0.0.5	Added comments collected during presentation; upgraded to stable draft
11th April 2018	0.0.6	Resolved comments; version for approval
28th May 2018	0.0.7	Corrected figures 8 - 11
7th June 2018	0.0.8	Editorial updates to follow ETSI drafting rules
7th June 2018	0.0.9	Ready for approval

History

Document history		
V1.1.1	September 2018	Publication