**3GPP TSG-SA4 Meeting SA4#113-e  *S4-210651***

**Telco, Online, 06 Apr 2021 - 14 Apr 2021 revision of S4-210474**

|  |
| --- |
| *CR-Form-v12.0* |
| **DRAFT CHANGE REQUEST** |
|  |
|  | **26.132** | **CR** | **0000** | **rev** | **-** | **Current version:** | **16.2.0** |  |
|  |
| *For* [***HE******LP***](http://www.3gpp.org/3G_Specs/CRs.htm#_blank)*on using this form: comprehensive instructions can be found at* [*http://www.3gpp.org/Change-Requests*](http://www.3gpp.org/Change-Requests)*.* |
|  |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ***Proposed change affects:*** | UICC apps |  | ME | **X** | Radio Access Network |  | Core Network |  |

|  |
| --- |
|  |
| ***Title:***  | Headset Interface Description |
|  |  |
| ***Source to WG:*** | HEAD acoustics GmbH |
| ***Source to TSG:*** | S4 |
|  |  |
| ***Work item code:*** | HInT |  | ***Date:*** | 2020-09-14 |
|  |  |  |  |  |
| ***Category:*** | **B** |  | ***Release:*** | Rel-17 |
|  | *Use one of the following categories:****F*** *(correction)****A*** *(mirror corresponding to a change in an earlier release)****B*** *(addition of feature),* ***C*** *(functional modification of feature)****D*** *(editorial modification)*Detailed explanations of the above categories canbe found in 3GPP [TR 21.900](http://www.3gpp.org/ftp/Specs/html-info/21900.htm). | *Use one of the following releases:Rel-8 (Release 8)Rel-9 (Release 9)Rel-10 (Release 10)Rel-11 (Release 11)Rel-12 (Release 12)**Rel-13 (Release 13)Rel-14 (Release 14)Rel-15 (Release 15)Rel-16 (Release 16)* |
|  |  |
| ***Reason for change:*** | Within the scope of the work item HInT, it is intended to add new test methods to TS 26.132 for analogue and digital interfaces of UE. As a preparation for these, a detailed specification and description of the introduced interfaces has to be included. |
|  |  |
| ***Summary of change:*** | Introduction of new clauses for analogue and digital interfaces, editorial changes in the existing clauses regarding measurement equipment. |
|  |  |
| ***Consequences if not approved:*** | Testing of the electric (headset) interface not considered. Evaluation of a mobile phone without an associated headset not possible. |
|  |  |
| ***Clauses affected:*** | Introduction, 1, 2, 3, 4, 4.1 (new), 4.2 (new), 4.3 (new), 4.4 (new), 5.1.6 (new), 5.2, 7.2.6 (new), 7.3.0 (new), 7.3, 7.3.3 (new), 7.3.4 (new), 7.4.0 (new), 7.4.7 (new), 7.4.8 (new), 7.5, 7.5.3a (new), 7.7, 7.7.5 (new), 7.8, 7.10.1b (new), 7.10.2b (new), 7.10.3b (new), 7.10.4, 7.11.1, 7.12, 7.12.3 (new), 7.13,8.2.6 (new), 8.3.0 (new), 8.3, 8.3.3 (new), 8.3.4 (new), 8.4.0 (new), 8.4.7 (new), 8.4.8 (new), 8.5, 8.5.3a (new), 8.7, 8.7.5 (new), 8.8, 8.10.1b (new), 8.10.2b (new), 8.10.3b (new), 8.10.4, 8.11.1, 8.12, 8.12.3 (new), 8.13, 9.2.6 (new), 9.3.0 (new), 9.3, 9.3.3 (new), 9.3.4 (new), 9.4.0, 9.4.7 (new), 9.4.8 (new), 9.5.3a (new), 9.7, 9.7.5 (new), 9.8, 9.10.1b (new), 9.10.2b (new), 9.10.3b (new), 9.10.4, 9.12, 9.12.3 (new), 9.13, 10.2.6 (new), 10.3.0 (new), 10.3, 10.3.3 (new), 10.3.4 (new), 10.4.0, 10.4.7 (new), 10.4.8 (new), 10.5.3a (new), 10.7.5 (new), 10.8, 10.10.1b (new), 10.10.2b (new), 10.10.3b (new), 10.10.4, 10.12, 10.12.3 (new), 10.13, Annex D.1 |
|  |  |
|  | **Y** | **N** |  |  |
| ***Other specs*** |  | **X** |  Other core specifications  | TS/TR ... CR ...  |
| ***affected:*** | **x** |  |  Test specifications | TS/TR 26.131 CR NNNN  |
| ***(show related CRs)*** |  | **X** |  O&M Specifications | TS/TR ... CR ...  |
|  |  |
| ***Other comments:*** |  |
|  |  |
| ***This CR's revision history:*** |  |

------------------------- START OF CHANGE 1 -------------------------

# Introduction

The present document specifies test methods to allow the minimum performance requirements for the electro-acoustic characteristics of GSM, 3G, LTE, NR and WLAN terminals when used to provide narrowband, wideband, super-wideband or fullband telephony to be assessed.

The objective for narrowband services is to reach a quality as close as possible to ITU-T standards for PSTN circuits. However, due to technical and economic factors, there cannot be full compliance with the general characteristics of international telephone connections and circuits recommended by the ITU-T.

The performance requirements are specified in TS 26.131; the test methods and considerations are specified in the main body of the text.

# 1 Scope

The present document is applicable to any terminal capable of supporting narrowband, wideband, super-wideband or fullband telephony, either as a stand-alone service or as the telephony component of a multimedia service. The present document specifies test methods to allow the minimum performance requirements for the electro-acoustic characteristics of GSM, 3G, LTE, NR and WLAN terminals when used to provide narrowband, wideband, super-wideband or fullband telephony to be assessed.

------------------------- END OF CHANGE 1 -------------------------

------------------------- START OF CHANGE 2 -------------------------

# 2 References

The following documents contain provisions which, through reference in this text, constitute provisions of the present document.

- References are either specific (identified by date of publication, edition number, version number, etc.) or non‑specific.

- For a specific reference, subsequent revisions do not apply.

- For a non-specific reference, the latest version applies. In the case of a reference to a 3GPP document (including a GSM document), a non-specific reference implicitly refers to the latest version of that document *in the same Release as the present document*.

[…]

[53] ITU-T Recommendation P.381: "Technical requirements and test methods for the universal wired headset or headphone interface of digital mobile terminals".

[54] ISO 3: "Preferred numbers — Series of preferred numbers".

[55] ITU-T Recommedation P.DHIP: "Technical requirements and test methods for headsets or headphones with wired or wireless digital interfaces and associated terminals" (draft).

------------------------- END OF CHANGE 2 -------------------------

------------------------- START OF CHANGE 3 -------------------------

# 3 Definitions, symbols and abbreviations

## 3.1 Definitions

For the purposes of the present document the terms *narrowband,* *wideband, super-wideband and fullband* refer to signals associated with the corresponding operating modes of the speech codecs specified in 5.2.

For the purposes of the present document, the terms dB, dBr, dBm0, dBm0p and dBA, shall be interpreted as defined in ITU-T Recommendation G.100 [51]; the term dBPa shall be interpreted as the sound pressure level relative to 1 pascal expressed in dB (0 dBPa is equivalent to 94 dB SPL).

A 3GPP softphone is a telephony system running on a general purpose computer or PDA complying with the 3GPP terminal acoustic requirements (TS 26.131 and 26.132).

For the purposes of the present document the term *clock skew* is defined as the difference between the clock of the device under test (CDUT) and the clock of the reference client (CREF). The skew of CDUT relative to CREF is defined in parts per million (PPM) as: (CDUT -CREF).106/ CREF.

For the purposes of the present document, the term *electrical interface* is defined as an analogue or digital access to an UE, which allows injecting and capturing signals electrically instead of through an acoustical interface. The interface can be either wired (analogue, digital) or wireless (digital). The purpose of this interface is to connect a separate device (typically a headset), which provides a receiver and transmitter for telephony.

## 3.2 Abbreviations

For the purposes of the present document, the abbreviations given in 3GPP TR 21.905 [47] and the following apply. An abbreviation defined in the present document takes precedence over the definition of the same abbreviation, if any, in 3GPP TR 21.905 [47].

[...]

RP Reference point (electrical or acoustical)

[...]

TCL Terminal coupling loss

TCLw Terminal coupling loss (weighted)

[...]

------------------------- END OF CHANGE 3 -------------------------

------------------------- START OF CHANGE 4 -------------------------

# 4 Interface definitions

## 4.1 General

The interfaces required to define terminal electro-acoustic characteristics are the acoustical interfaces, the air interface and the point of interconnect (POI), see Figure 1.

## 4.2 Air interfaces

The Air Interfaces for GSM, 3G, LTE and NR are specified by GSM 05, 3GPP 45, 3GPP 25, 3GPP 36 and 3GPP 38 series specifications, and the Air Interface for WLAN access to EPC is specified by WLAN access to EPC as defined in 3GPP TS 23.402 [48] and TS 24.302 [49].

Measurements can be made using a system simulator (SS) comprising the appropriate radio terminal equipment and speech transcoder. The delays, losses and gains introduced by the test equipment shall be accounted for.

## 4.3 Acoustical interfaces

The POI with the public switched telephone network (PSTN) is considered to have a relative level of 0 dBr.

Five classes of acoustical interface are considered in this specification:

- Handset UE including softphone UE used as a handset;

- Headset UE including softphone UE used with headset;

- Vehicle Mounted Hands-free UE including softphone UE mounted in a vehicle;

- Desktop-mounted hands-free UE including softphone UE with external loudspeaker(s) used in hands-free mode;

- Hand-held hands-free UE including softphone UE with internal loudspeaker(s) used in hands-free mode.

(See definition of softphone in Clause 3.1)

NOTE: The test setup for a softphone UE shall be derived according to the following rules:

- When using a softphone UE as a handset: the test setup shall correspond to handset mode.

- When using a softphone UE with headset: the test setup shall correspond to headset mode.

- When a softphone UE is mounted in a vehicle: the test setup shall correspond to vehicle-mounted hands-free mode.

- When using a softphone UE in hands-free mode:

- When using internal loudspeaker(s), the test setup shall correspond to hand-held hands-free.

- When using external loudspeaker(s), the test setup shall correspond to desktop-mounted hands-free.

## 4.4 Electrical interfaces

An electrical interface is considered in this specification and details on standardized analogue (wired) and digital (wired and wireless) headset interfaces can be found in clause 5.1.6. For the electrical interface, the POI in sending / receiving direction is respectively defined as the input / output of the reference speech coder of the system simulator.

Any of the UE types mentioned in clause 4.3 providing an electrical interface can be considered as Electrical Interface UE.

------------------------- END OF CHANGE 4 -------------------------

------------------------- START OF CHANGE 5 -------------------------

### 5.1.6 Test setup for electrical interfaces

#### 5.1.6.1 Wired analogue connection

UE testing via analogue connection shall be carried out with a universal wired headset interface, which complies with the electrical and physical characteristics described in clause 6 of ITU‑T P.381 [53]. Different types of socket designs can be used as well. Other implementations of analogue electrical interfaces (wired or wireless) are out of scope.

Figure 15a5b illustrates the setup required for testing analogue electrical interfaces. The electric output impedance of the reference interface of the test equipment shall be in the range of 1 Ω and 10 kΩ. The corresponding electric input impedance shall be 32 Ω +/- 2 Ω. The common ground impedance (between sending and receiving sides) for the test system shall be ≤ 0.05 Ω.

If not specified otherwise, the nominal signal levels are:

- -60 dBV in send direction (input to electrical interface UE), which corresponds to an acoustic level of -4.7 dBPa at the MRP, i.e., a default sensitivity of ~55 dBV/Pa.

- -39 dBV in receive direction (for an electrical interface UE providing stereo/diotic output), for a nominal volume setting (if present).

For testing echo and double talk scenarios, an artificial feedback of the receive signal into the sending path shall be used. This echo path can be realized in an analogue (e.g, a stand-alone device) or digital (e.g., part of the test system) way. For measurements without artificial echo loss, the feedback path is disabled.

NOTE: It is assumed that mainly passive third-party devices are connected via analogue electrical interface to the UE, which do not contain any typical signal processing capabilities (like e.g., echo cancellation or noise reduction). Thus, all tests specified for this interface are comparable to handset UE or headset UE, i.e. they expect that any possible signal processing is applied in the UE.



Figure 15a5b: Test setup for analogue electrical headset interface

#### 5.1.6.2 Digital connection

Figure 15a5c illustrates the setup required for testing digital electrical interfaces of a UE. Such testing is possible with different types of wired and wireless technologies and requires a digital reference interface in the test system. The present document is only applicable to USB and Bluetooth interfaces, other digital interfaces are out of scope.

For testing echo and double talk scenarios, an artificial feedback of the receive signal into the sending path shall be used. This echo path shall be realized as part of the test system. The received and decoded signal from the UE is fed back into the sending direction, in advance to the encoding/protocol/hardware layer. For measurements without artificial echo loss, the feedback path is disabled.

Digital levels for send and receive direction are specified in dBm0, referring to the same definition as for the input/output of the terminal and the system simulator (see clause 5.2).

In contrast to the analogue interface, devices connected to the UE via digital interface may either provide active (includes signal processing for send and/or receive direction) or passive (comparable to analogue headsets, see clause 5.1.6.1) functionalities. Tests are only applicable in case of typical signal processing for telecommunication (e.g., noise reduction, echo cancellation) takes places only in the UE and not in the equipment to be connected. If necessary and if the digital interface and the associated protocols support the exchange of commands/meta-information, the electrical interface UE shall be configured in such away that carries out its own signal processing.

Tests are not applicable in case the signal processing in the UE cannot be enabled in any way for the electrical interface. The UE should then behave as a transparent gateway, i.e. shall not apply any signal processing at all. Appropriate transparency tests can be found in Recommendation ITU-T P.DHIP [55], but are out of scope for the present document.



Figure 15a5c: Test setup for digital electrical headset interface

------------------------- END OF CHANGE 5 -------------------------

------------------------- START OF CHANGE 6 -------------------------

## 5.2 Setup of the electrical interfaces of test equipment

### 5.2.1 Codec approach and specification

In this approach, a codec is used to convert the digital input/output bit-stream of the system simulator to the equivalent analogue values. With this approach a system simulator simulating the radio link to the terminal under controlled and error-free conditions is required, unless otherwise specified for the respective test. The system simulator has to be equipped with a high-quality codec with characteristics as close as possible to ideal.

Definition of 0 dBr point:

D/A converter - a Digital Test Sequence (DTS) representing the codec equivalent of an analogue sinusoidal signal with an RMS value of 3,14 dB below the maximum full‑load capacity of the codec shall generate 0 dBm across a 600 ohm load;

A/D converter - a 0 dBm signal generated from a 600 ohm source shall give the digital test sequence (DTS) representing the codec equivalent of an analogue sinusoidal signal with an RMS value of 3,14 dB below the maximum full‑load capacity of the codec.

------------------------- END OF CHANGE 6 -------------------------

Editors' Note:

At SA4#113-e, changes #1 to #6 (clauses 1 to 5 of TS 26.132) are preliminary agreed (stable since several meetings)?

For sake of clarity and traceability, the following changes are organized in a different order than in 3GPP TS 26.132. Consecutive changes are now grouped by the same type of measurements (instead of being grouping by NB/WB/SWB/FB). This should make it easier to track identical/similar changes to related measurement descriptions. For the approval of the final CR, it is intended to reorder the changes according to the actual clause numbers.

In several clauses, no change was needed for electrical interface UE:

- 9.11 Echo control characteristics (SWB): the descriptions point to clause 8.11 (WB)

- 10.11 Echo control characteristics (FB): the descriptions point to clause 9.11 (SWB)

------------------------- START OF CHANGE 7 -------------------------

## 7.2 Overall loss/loudness ratings

### 7.2.1 General

The SLR and RLR values for GSM, 3G, LTE, NR or WLAN networks apply up to the POI. However, the main determining factors are the characteristics of the UE, including the analogue to digital conversion (ADC) and digital to analogue conversion (DAC). In practice, it is convenient to specify loudness ratings to the Air Interface. For the normal case, where the GSM, 3G, LTE, NR or WLAN network introduce no additional loss between the Air Interface and the POI, the loudness ratings to the PSTN boundary (POI) will be the same as the loudness ratings measured at the Air Interface.

### 7.2.2 Connections with handset UE

[...]

### 7.2.3 Connections with desktop and vehicle-mounted hands-free UE

[...]

### 7.2.4 Connections with hand-held hands-free UE

[...]

### 7.2.5 Connections with headset UE

Same as for handset.

### 7.2.6 Connections with electrical interface UE

#### 7.2.6.1 Sending junction loudness rating (SJLR)

a) The test signal to be used for the measurements shall be the British-English single talk sequence described in ITU-T Recommendation P.501 [22]. For electrical interface UE, the active speech level of the signal shall be calibrated to -60 dBV for analogue and to -16 dBm0 for digital connections. The test signal level is calculated over the complete test signal sequence.

NOTE: The specified electrical levels correpond to an acoustic level of -4.7 dBPa at MRP.

b) The reference signal to be used for the calculation shall be the same as the test signal and is calibrated to ‑16 dBm0 (independent of analogue or digital connection).

c) The terminal is setup as described in clause 5.1.6 and the test signal is transmitted in sending direction. For the calculation, the averaged measured level at each frequency band is referred to the averaged reference signal level measured in each frequency band.

d) The sensitivity is expressed in dB. The sending junction loudness rating (SJLR) is calculated according to equation A-23d of ITU‑T Recommendation P.79 [16], bands 4-17, m = 0.175 and the weighting factors for JLR according to Table A.2 of ITU‑T Recommendation P.79 [16]. For the calculation, the average measured level at the output of system simulator for each frequency band is referred to the reference signal.

#### 7.2.6.2 Receving junction loudness rating (RJLR)

a) The test signal to be used for the measurements shall be the British-English single talk sequence described in ITU-T Recommendation P.501 [22]. The test signal level shall be -16 dBm0 measured at the digital reference point or the equivalent analogue point. The test signal level is calculated over the complete test signal sequence.

b) The reference signal to be used for the calculation shall be the same as the test signal and is calibrated to ‑39 dBV for analogue and to ‑16 dBm0 for digital connections.

c) The terminal is setup as described in clause 5.1.6 and the test signal is transmitted in receiving direction. For the calculation, the averaged measured level at each frequency band is referred to the averaged reference signal level measured in each frequency band.

d) The sensitivity is expressed in dB. The receiving junction loudness rating (RJLR) is calculated according to equation A-23d of ITU‑T Recommendation P.79 [16], bands 4-17, m = 0.175 and the weighting factors for JLR according to Table A.2 of ITU‑T Recommendation P.79 [16]. For the calculation, the average measured level at the output of the electrical interface UE for each frequency band is referred to the reference signal.

------------------------- END OF CHANGE 7 -------------------------

------------------------- START OF CHANGE 8 -------------------------

## 8.2 Overall loss/loudness ratings

### 8.2.1 General

The SLR and RLR values for GSM, 3G, LTE, NR or WLAN networks apply up to the POI. However, the main determining factors are the characteristics of the UE, including the analogue to digital conversion (ADC) and digital to analogue conversion (DAC). In practice, it is convenient to specify loudness ratings to the Air Interface. For the normal case, where the GSM, 3G, LTE, NR or WLAN network introduce no additional loss between the Air Interface and the POI, the loudness ratings to the PSTN boundary (POI) will be the same as the loudness ratings measured at the Air Interface.

### 8.2.2 Connections with handset UE

[...]

### 8.2.3 Connections with desktop and vehicle-mounted hands-free UE

[...]

### 8.2.4 Connections with hand-held hands-free UE

[...]

### 8.2.5 Connections with headset UE

Same as for handset.

### 8.2.6 Connections with electrical interface UE

#### 8.2.6.1 Sending junction loudness rating (SJLR)

a) The test signal to be used for the measurements shall be the British-English single talk sequence described in ITU-T Recommendation P.501 [22]. For electrical interface UE, the active speech level of the signal shall be calibrated to -60 dBV for analogue and to -16 dBm0 for digital connections. The test signal level is calculated over the complete test signal sequence.

NOTE: The specified electrical levels correpond to an acoustic level of -4.7 dBPa at MRP.

b) The reference signal to be used for the calculation shall be the same as the test signal and is calibrated to ‑16 dBm0 (independent of analogue or digital connection).

c) The terminal is setup as described in clause 5.1.6 and the test signal is transmitted in sending direction. For the calculation, the averaged measured level at each frequency band is referred to the averaged reference signal level measured in each frequency band.

d) The sensitivity is expressed in dB. The sending junction loudness rating (SJLR) is calculated according to equation A-23d of ITU T Recommendation P.79 [16], bands 1-20, m = 0.175 and the weighting factors for JLR according to Table A.2 of ITU T Recommendation P.79 [16]. For the calculation, the average measured level at the output of system simulator for each frequency band is referred to the reference signal.

#### 8.2.6.2 Receving junction loudness rating (RJLR)

a) The test signal to be used for the measurements shall be the British-English single talk sequence described in ITU-T Recommendation P.501 [22]. The test signal level shall be -16 dBm0 measured at the digital reference point or the equivalent analogue point. The test signal level is calculated over the complete test signal sequence.

b) The reference signal to be used for the calculation shall be the same as the test signal and is calibrated to ‑39 dBV for analogue and to ‑16 dBm0 for digital connections.

c) The terminal is setup as described in clause 5.1.6 and the test signal is transmitted in receiving direction. For the calculation, the averaged measured level at each frequency band is referred to the averaged reference signal level measured in each frequency band.

d) The sensitivity is expressed in dB. The receiving junction loudness rating (RJLR) is calculated according to equation A-23d of ITU T Recommendation P.79 [16], bands 1-20, m = 0.175 and the weighting factors for JLR according to Table A.2 of ITU T Recommendation P.79 [16]. For the calculation, the average measured level at the output of the electrical interface UE for each frequency band is referred to the reference signal.

------------------------- END OF CHANGE 8 -------------------------

------------------------- START OF CHANGE 9 -------------------------

## 9.2 Overall loss/loudness ratings

### 9.2.1 General

[...]

### 9.2.2 Connections with handset UE

[...]

### 9.2.3 Connections with desktop and vehicle-mounted hands-free UE

[...]

### 9.2.4 Connections with hand-held hands-free UE

[...]

### 9.2.5 Connections with headset UE

[...]

### 9.2.6 Connections with electrical interface UE

#### 9.2.6.1 Sending junction loudness rating (SJLR)

The description is the same as for wideband (see sub-clause 8.2.6.1).

#### 9.2.6.2 Receving junction loudness rating (RJLR)

The description is the same as for wideband (see sub-clause 8.2.6.2).

------------------------- END OF CHANGE 9 -------------------------

------------------------- START OF CHANGE 10 -------------------------

## 10.2 Overall loss/loudness ratings

### 10.2.1 General

[...]

### 10.2.2 Connections with handset UE

[...]

### 10.2.3 Connections with desktop and vehicle-mounted hands-free UE

[...]

### 10.2.4 Connections with hand-held hands-free UE

[...]

### 10.2.5 Connections with headset UE

[...]

### 10.2.6 Connections with electrical interface UE

#### 10.2.6.1 Sending junction loudness rating (SJLR)

The description is the same as for wideband (see sub-clause 8.2.6.1).

#### 10.2.6.2 Receving junction loudness rating (RJLR)

The description is the same as for wideband (see sub-clause 8.2.6.2).

------------------------- END OF CHANGE 10 -------------------------

------------------------- START OF CHANGE 11 -------------------------

## 7.3 Idle channel noise (handset, headset and electrical interface UE)

### 7.3.0 Overview

For idle noise measurements in sending and receiving directions, care should be taken that only the noise is windowed out by the analysis and the result is not impaired by any remaining reverberation or by noise and/or interference from various other sources. Some examples are air-conducted or vibration-conducted noise from sources inside or outside the test chamber, disturbances from lights and regulators, mains supply induced noise including grounding issues, test system and system simulator inherent noise as well as radio interference from the UE to test equipment such as ear simulators, microphone amplifiers, etc.

The following steps shall be followed in advance to both measurement directions:

a) test

b) The terminal should be configured to the test equipment as described in subclause 5.1.

c) A test signal may have to be intermittently applied to prevent ‘silent mode’ operation of the MS. This is for further study.

d) An optional activation sequence may be used, to e.g., override a voice activity detection. In this case, the additional test signal shall be suitable regarding level and bandwidth, like e.g., the composite source signals described in clause 7.10.

### 7.3.1 Sending

In advance to the measurement, the general steps listed in clause 7.3.0 shall be followed.

a) In advance to the noise level measurement, an optional activation sequence may be used.

b) The noise level at the output of the SS is measured with psophometric weighting. The psophometric weighting filter is described in ITU-T Recommendation O.41 [23].

c) The measured part of the noise shall be 170,667 ms (which equals 8192 samples in a 48 kHz sample rate test system). The spectral distribution of the noise is analyzed with an 8k FFT using windowing with ≤ 0,1 dB leakage for non bin-centered signals. This can be achieved with a window function commonly known as a "flat top window". Within the specified frequency range, the FFT bin that has the highest level is searched for; the level of this bin is the maximum level of a single frequency disturbance.

d)

The total noise powers obtained from such repeats shall be averaged. The total result shall be 10 \* log10 of this average in dB.

e) The single frequency maximum powers obtained from such repeats shall be averaged. The total result shall be 10\*log10 of this average in dB.

### 7.3.2 Receiving

In advance to the measurement, the general steps listed in clause 7.3.0 shall be followed.

a) In advance to the noise level measurement, an optional activation sequence may be used.

b) The noise level shall be measured with A‑weighting at the DRP with diffuse-field correction. The A-weighting filter is described in IEC 60651 [12].

c) The measured part of the noise shall be 170,667 ms (which equals 8192 samples in a 48 kHz sample rate test system). The spectral distribution of the noise is analyzed with an 8k FFT using windowing with ≤ 0.1 dB leakage for non bin-centred signals. This can be achieved with a window function commonly known as a "flat top window". Within the specified frequency range, the FFT bin that has the highest level is searched for; the level of this bin is the maximum level of a single frequency disturbance.

d) The total noise powers obtained from such repeats shall be averaged. The total result shall be 10\*log10 of this average in dB.

e) The single frequency maximum powers obtained from such repeats shall be averaged. The total result shall be 10\*log10 of this average in dB.

### 7.3.3 Sending (electrical interface UE)

Same method as in clause 7.3.1.

### 7.3.4 Receiving (electrical interface UE)

Same method as in clause 7.3.1, except that the idle noise signal is captured at the receive output of the electrical reference interface.

------------------------- END OF CHANGE 11 -------------------------

------------------------- START OF CHANGE 12 -------------------------

## 8.3 Idle channel noise (handset, headset and electrical interface UE)

### 8.3.0 Overview

For idle noise measurements in sending and receiving directions, care should be taken that only the noise is windowed out by the analysis and the result is not impaired by any remaining reverberation or by noise and/or interference from various other sources. Some examples are air-conducted or vibration-conducted noise from sources inside or outside the test chamber, disturbances from lights and regulators, mains supply induced noise including grounding issues, test system and system simulator inherent noise as well as radio interference from the UE to test equipment such as ear simulators, microphone amplifiers, etc.

The following steps shall be followed in advance to both measurement directions:

a) The test environment shall comply with the conditions described in subclause 6.1.

b) The terminal should be configured to the test equipment as described in subclause 5.1.

c) A test signal may have to be intermittently applied to prevent ‘silent mode’ operation of the MS. This is for further study.

d) An optional activation sequence may be used, to e.g., override a voice activity detection. In this case, the additional test signal shall be suitable regarding level and bandwidth, like e.g., the composite source signals described in clause 8.10.

To improve repeatability, the test sequence (optional activation followed by the noise level measurement) may be contiguously repeated one or more times.

### 8.3.1 Sending

In advance to the measurement, the general steps listed in clause 8.3.0 shall be followed.

a) In advance to the noise level measurement, an optional activation sequence may be used.

b) The noise level at the output of the SS is measured with A‑weighting. The A-weighting filter is described in IEC 60651 [12].

c) The measured part of the noise shall be 170,667 ms (which equals 8192 samples in a 48 kHz sample rate test system). The spectral distribution of the noise is analyzed with an 8k FFT using windowing with ≤ 0,1 dB leakage for non bin-centered signals. This can be achieved with a window function commonly known as a "flat top window". Within the specified frequency range, the FFT bin that has the highest level is searched for; the level of this bin is the maximum level of a single frequency disturbance.

d) The total noise powers obtained from such repeats shall be averaged. The total result shall be 10 \* log10 of this average in dB.

e) The single frequency maximum powers obtained from such repeats shall be averaged. The total result shall be 10 \* log10 of this average in dB.

### 8.3.2 Receiving

In advance to the measurement, the general steps listed in clause 8.3.0 shall be followed.

a) In advance to the noise level measurement, an optional activation sequence may be used.

b) The noise shall be measured with A‑weighting at the DRP with diffuse-field correction. The A-weighting filter is described in IEC 60651 [12].

c) The measured part of the noise shall be 170,667 ms (which equals 8192 samples in a 48 kHz sample rate test system). The spectral distribution of the noise is analyzed with an 8k FFT using windowing with ≤ 0,1 dB leakage for non bin-centered signals. This can be achieved with a window function commonly known as a "flat top window". Within the specified frequency range the FFT bin that has the highest level is searched for; the level of this bin is the maximum level of a single frequency disturbance.

d) The total noise powers obtained from such repeats shall be averaged. The total result shall be 10 \* log10 of this average in dB.

e) The single frequency maximum powers obtained from such repeats shall be averaged. The total result shall be 10 \* log10 of this average in dB.

### 8.3.3 Sending (electrical interface UE)

Same method as in clause 8.3.1.

### 8.3.4 Receiving (electrical interface UE)

Same method as in clause 8.3.1, except that the idle noise signal is captured at the receive output of the electrical reference interface.

------------------------- END OF CHANGE 12 -------------------------

------------------------- START OF CHANGE 13 -------------------------

## 9.3 Idle channel noise (handset, headset and electrical interface UE)

### 9.3.0 Overview

For idle noise measurements in sending and receiving directions, care should be taken that only the noise is windowed out by the analysis and the result is not impaired by any remaining reverberation or by noise and/or interference from various other sources. Some examples are air-conducted or vibration-conducted noise from sources inside or outside the test chamber, disturbances from lights and regulators, mains supply induced noise including grounding issues, test system and system simulator inherent noise as well as radio interference from the UE to test equipment such as ear simulators, microphone amplifiers, etc.

The following steps shall be followed in advance to both measurement directions:

a) The test environment shall comply with the conditions described in subclause 6.1.

b) The terminal should be configured to the test equipment as described in subclause 5.1.

c) A test signal may have to be intermittently applied to prevent ‘silent mode’ operation of the MS. This is for further study.

d) An optional activation sequence may be used, to e.g., override a voice activity detection. In this case, the additional test signal shall be suitable regarding level and bandwidth, like e.g., the composite source signals described in clause 9.10.

To improve repeatability, the test sequence (optional activation followed by the noise level measurement) may be contiguously repeated one or more times.

### 9.3.1 Sending

In advance to the measurement, the general steps listed in clause 9.3.0 shall be followed.a) In advance to the noise level measurement, an optional activation sequence may be used.

b) The noise level at the output of the SS is measured from 100 Hz to 16 kHz with A‑weighting. The A-weighting filter is described in IEC 60651.

c) The measured part of the noise shall be 170,667 ms (which equals 8192 samples in a 48 kHz sample rate test system). The spectral distribution of the noise is analyzed with an 8k FFT using windowing with ≤ 0,1 dB leakage for non bin-centered signals. This can be achieved with a window function commonly known as a "flat top window". Within the specified frequency range, the FFT bin that has the highest level is searched for; the level of this bin is the maximum level of a single frequency disturbance.

d) The total noise powers obtained from such repeats shall be averaged. The total result shall be 10 \* log10 of this average in dB.

e) The single frequency maximum powers obtained from such repeats shall be averaged. The total result shall be 10 \* log10 of this average in dB.

### 9.3.2 Receiving

In advance to the measurement, the general steps listed in clause 9.3.0 shall be followed.

a) In advance to the noise level measurement, an optional activation sequence may be used.

b) The noise shall be measured from 100 Hz to 20 kHz with A‑weighting at the DRP with diffuse-field correction. The A-weighting filter is described in IEC 60651.

c) The measured part of the noise shall be 170,667 ms (which equals 8192 samples in a 48 kHz sample rate test system). The spectral distribution of the noise is analyzed with an 8k FFT using windowing with ≤ 0,1 dB leakage for non bin-centered signals. This can be achieved with a window function commonly known as a "flat top window". Within the specified frequency range the FFT bin that has the highest level is searched for; the level of this bin is the maximum level of a single frequency disturbance.

d) The total noise powers obtained from such repeats shall be averaged. The total result shall be 10 \* log10 of this average in dB.

e) The single frequency maximum powers obtained from such repeats shall be averaged. The total result shall be 10 \* log10 of this average in dB.

### 9.3.3 Sending (electrical interface UE)

Same method as in clause 9.3.1.

### 9.3.4 Receiving (electrical interface UE)

Same method as in clause 9.3.1, except that the idle noise signal is captured at the receive output of the electrical reference interface.

------------------------- END OF CHANGE 13 -------------------------

------------------------- START OF CHANGE 14 -------------------------

## 10.3 Idle channel noise (handset, headset and electrical interface UE)

### 10.3.0 Overview

The test method is the same as for super-wideband (see sub-clause 9.3).

### 10.3.1 Sending

The test method is the same as for super-wideband (see sub-clause 9.3.1), except that the noise level is measured in the frequency range from 100 Hz to 20 kHz.

### 10.3.2 Receiving

The test method is the same as for super-wideband (see sub-clause 9.3.2), except that the noise level is measured in the frequency range from 100 Hz to 20 kHz.

### 10.3.3 Sending (electrical interface UE)

Same method as in clause 10.3.1.

### 10.3.4 Receiving (electrical interface UE)

Same method as in clause 10.3.1, except that the idle noise signal is captured at the receive output of the electrical reference interface.

------------------------- END OF CHANGE 14 -------------------------

------------------------- START OF CHANGE 15 -------------------------

## 7.4 Sensitivity/frequency characteristics

### 7.4.0 General

For checking the sensitivity/frequency characteristics against performance requirements (as in e.g., 3GPP TS 26.131 [1]), any given tolerance mask shall be defined for each center frequency of the fractional octave bands, which is used in the respective test method. If necessary, the tolerance mask is interpolated linearly for a certain center frequency between the two closest neighbouring data points on a log-frequency scale and the magnitude in dB.

### 7.4.1 Handset and headset UE sending

[...]

### 7.4.2 Handset and headset UE receiving

[...]

### 7.4.3 Desktop and vehicle-mounted hands-free UE sending

[...]

### 7.4.4 Desktop and vehicle-mounted hands-free UE receiving

[...]

### 7.4.5 Hand-held hands-free UE sending

[...]

### 7.4.6 Hand-held hands-free UE receiving

[...]

### 7.4.7 Electrical interface UE sending

a) The test signal to be used for the measurements shall be the British-English single talk sequence described in ITU-T Recommendation P.501 [22]. The active speech level of the signal shall be calibrated to -60 dBV for analogue and to -16 dBm0 for digital connections. The test signal level is calculated over the complete test signal sequence.

b) The reference signal to be used for the calculation shall be the same as the test signal and is calibrated to ‑4.7 dBPa (independent of analogue or digital connection).

c) The electrical interface is setup as described in clause 5.1.6. Measurements shall be made at 1/12-octave intervals as given by the R.40 series of preferred numbers in ISO 3 [54] for frequencies from 100 Hz to 4 kHz inclusive. For the calculation, the averaged measured level at the electrical reference point for each frequency band is referred to the averaged reference signal level measured in each frequency band.

d) The sensitivity is expressed in terms of dB.

### 7.4.8 Electrical interface UE receiving

a) The test signal to be used for the measurements shall be the British-English single talk sequence described in ITU-T Recommendation P.501 [22]. The test signal level shall be -16 dBm0 measured at the digital reference point or the equivalent analogue point. The test signal level is calculated over the complete test signal sequence.

b) The reference signal to be used for the calculation shall be the same as the test signal and is calibrated to ‑39 dBV for analogue and to ‑16 dBm0 for digital connections.

c) The handset terminal is setup as described in clause 5. Measurements shall be made at 1/12-octave intervals as given by the R.40 series of preferred numbers in ISO 3 [54] for frequencies from 100 Hz to 4 kHz inclusive. For the calculation, the average measured level at the output of the electrical interface UE for each frequency band is referred to the reference signal.

d) The sensitivity is expressed in terms of dB.

------------------------- END OF CHANGE 15 -------------------------

------------------------- START OF CHANGE 16 -------------------------

## 8.4 Sensitivity/frequency characteristics

### 8.4.0 General

For checking the sensitivity/frequency characteristics against performance requirements (as in e.g., 3GPP TS 26.131 [1]), any given tolerance mask shall be defined for each center frequency of the fractional octave bands, which is used in the respective test method. If necessary, the tolerance mask is interpolated linearly for a certain center frequency between the two closest neighbouring data points on a log-frequency scale and the magnitude in dB.

### 8.4.1 Handset and headset UE sending

[...]

### 8.4.2 Handset and headset UE receiving

[...]

### 8.4.3 Desktop and vehicle-mounted hands-free UE sending

[...]

### 8.4.4 Desktop and vehicle-mounted hands-free UE receiving

[...]

### 8.4.5 Hand-held hands-free UE sending

[...]

### 8.4.6 Hand-held hands-free UE receiving

[...]

### 8.4.7 Electrical interface UE sending

a) The test signal to be used for the measurements shall be the British-English single talk sequence described in ITU-T Recommendation P.501 [22]. The active speech level of the signal shall be calibrated to -60 dBV for analogue and to -16 dBm0 for digital connections. The test signal level is calculated over the complete test signal sequence.

b) The reference signal to be used for the calculation shall be the same as the test signal and is calibrated to ‑4.7 dBPa (independent of analogue or digital connection).

c) The electrical interface is setup as described in clause 5.1.6. Measurements shall be made at 1/12-octave intervals as given by the R.40 series of preferred numbers in ISO 3 [54] for frequencies from 100 Hz to 8 kHz inclusive. For the calculation, the averaged measured level at the electrical reference point for each frequency band is referred to the averaged reference signal level measured in each frequency band.

d) The sensitivity is expressed in terms of dB.

### 8.4.8 Electrical interface UE receiving

a) The test signal to be used for the measurements shall be the British-English single talk sequence described in ITU-T Recommendation P.501 [22]. The test signal level shall be -16 dBm0 measured at the digital reference point or the equivalent analogue point. The test signal level is calculated over the complete test signal sequence.

b) The reference signal to be used for the calculation shall be the same as the test signal and is calibrated to ‑39 dBV for analogue and to ‑16 dBm0 for digital connections.

c) The handset terminal is setup as described in clause 5. Measurements shall be made at 1/12-octave intervals as given by the R.40 series of preferred numbers in ISO 3 [54] for frequencies from 100 Hz to 8 kHz inclusive. For the calculation, the average measured level at the output of the electrical interface UE for each frequency band is referred to the reference signal.

d) The sensitivity is expressed in terms of dB.

------------------------- END OF CHANGE 16 -------------------------

------------------------- START OF CHANGE 17 -------------------------

## 9.4 Sensitivity/frequency characteristics

### 9.4.0 General

For checking the sensitivity/frequency characteristics against performance requirements (as in e.g., 3GPP TS 26.131 [1]), any given tolerance mask shall be defined for each center frequency of the fractional octave bands, which is used in the respective test method. If necessary, the tolerance mask is interpolated linearly for a certain center frequency between the two closest neighbouring data points on a log-frequency scale and the magnitude in dB.

.

### 9.4.1 Handset and headset UE sending

[...]

### 9.4.2 Handset and headset UE receiving

[...]

### 9.4.3 Desktop and vehicle-mounted hands-free UE sending

[...]

### 9.4.4 Desktop and vehicle-mounted hands-free UE receiving

[...]

### 9.4.5 Hand-held hands-free UE sending

[...]

### 9.4.6 Hand-held hands-free UE receiving

[...]

### 9.4.7 Electrical interface UE sending

a) The test signal to be used for the measurements shall be the British-English single talk sequence described in ITU-T Recommendation P.501 [22]. The active speech level of the signal shall be calibrated to -60 dBV for analogue and to -16 dBm0 for digital connections. The test signal level is calculated over the complete test signal sequence.

b) The reference signal to be used for the calculation shall be the same as the test signal and is calibrated to ‑4.7 dBPa (independent of analogue or digital connection).

c) The electrical interface is setup as described in clause 5.1.6. Measurements shall be made at both 1/3-octave and 1/12-octave intervals as given by the R.10 and R.40 series of preferred numbers in ISO 3 [54] for frequencies from 100 Hz to 16 kHz inclusive. For the calculation, the averaged measured level at the electrical reference point for each frequency band is referred to the averaged reference signal level measured in each frequency band.

d) The sensitivity is expressed in terms of dB.

### 9.4.8 Electrical interface UE receiving

a) The test signal to be used for the measurements shall be the British-English single talk sequence described in ITU-T Recommendation P.501 [22]. The test signal level shall be -16 dBm0 measured at the digital reference point or the equivalent analogue point. The test signal level is calculated over the complete test signal sequence.

b) The reference signal to be used for the calculation shall be the same as the test signal and is calibrated to ‑39 dBV for analogue and to ‑16 dBm0 for digital connections.

c) The handset terminal is setup as described in clause 5. Measurements shall be made at both 1/3-octave and 1/12-octave intervals as given by the R.10 and R.40 series of preferred numbers in ISO 3 [54] for frequencies from 100 Hz to 16 kHz inclusive. For the calculation, the average measured level at the output of the electrical interface UE for each frequency band is referred to the reference signal.

d) The sensitivity is expressed in terms of dB.

------------------------- END OF CHANGE 17 -------------------------

------------------------- START OF CHANGE 18 -------------------------

## 10.4 Sensitivity/frequency characteristics

### 10.4.0 General

For checking the sensitivity/frequency characteristics against performance requirements (as in e.g., 3GPP TS 26.131 [1]), any given tolerance mask shall be defined for each center frequency of the fractional octave bands, which is used in the respective test method. If necessary, the tolerance mask is interpolated linearly for a certain center frequency between the two closest neighbouring data points on a log-frequency scale and the magnitude in dB.

### 10.4.1 Handset and headset UE sending

[...]

### 10.4.2 Handset and headset UE receiving

[...]

### 10.4.3 Desktop and vehicle-mounted hands-free UE sending

[...]

### 10.4.4 Desktop and vehicle-mounted hands-free UE receiving

[...]

### 10.4.5 Hand-held hands-free UE sending

[...]

### 10.4.6 Hand-held hands-free UE receiving

[...]

### 10.4.7 Electrical interface UE sending

The test method is the same as for super-wideband (see sub-clause 9.4.7).

### 10.4.8 Electrical interface UE receiving

The test method is the same as for super-wideband (see sub-clause 9.4.8, observing the signal properties for fullband described in sub-clause 5.4).

------------------------- END OF CHANGE 18 -------------------------

------------------------- START OF CHANGE 19 -------------------------

## 7.5 Sidetone characteristics

### 7.5.1 Connections with handset UE

#### 7.5.1.0 General

The test signal to be used for the measurements shall be the British-English single talk sequence described in ITU-T Recommendation P.501 [22]. The spectrum of the acoustic signal produced by the artificial mouth is calibrated under free-field conditions at the MRP. The test signal level shall be -4,7 dBPa measured at the MRP. The test signal level is calculated over the complete test signal sequence.

#### 7.5.1.1 void

#### 7.5.1.2 Connections with handset UE – HATS method

a) The handset UE is setup as described in clause 5. The application force shall be 13 N on the Type 3.3 artificial ear.

b) Where a user operated volume control is provided, the measurements shall be carried out at the nominal setting of the volume control. In addition the measurement is repeated at the maximum volume control setting. It is expected that for other positions of the volume control setting a consistent behaviour to that of the nominal and maximum settings should be observed. Additional measurements for these positions are not required.

c) Measurements shall be made at 1/12-octave intervals as given by the R.40 series of preferred numbers in ISO 3 [54] for frequencies from 100 Hz to 8 kHz inclusive. The averaged measured level at DRP in each frequency band is referred to the averaged test signal level measured in each frequency band.

d) The sidetone path loss (LmeST), as expressed in dB, shall be calculated from each 1/3-octave band (ITU-T Recommendation P.79 [16], table B.1, bands 4 to 17). The Sidetone Masking Rating (STMR), expressed in dB, shall be calculated from formula B-4 of ITU-T Recommendation P.79 [16], using m = 0.225 and the weighting factors in table B.2 (unsealed condition) of ITU-T Recommendation P.79 [16]. No leakage correction (LE) shall be applied. DRP-ERP correction is used.

e) In case the STMR is below the limit, the measurement shall be repeated with the electrical sidetone path disabled and both sets of results shall be reported. In case the STMR is below the limit also with the electrical sidetone path disabled, the result shall not be regarded as a failure. Disconnecting the call is normally disabling the electrical sidetone path; otherwise the UE can be switched off to enter the wanted state.

### 7.5.2 Headset UE

a) The test signal to be used for the measurements shall be the British-English single talk sequence described in ITU-T Recommendation P.501 [22]. The spectrum of the acoustic signal produced by the artificial mouth is calibrated under free-field conditions at the MRP. The test signal level shall be -4,7 dBPa measured at the MRP. The test signal level is calculated over the complete test signal sequence.

b) Where a user operated volume control is provided, the measurements shall be carried out at the nominal setting of the volume control. In addition the measurement is repeated at the maximum volume control setting. It is expected that for other positions of the volume control setting a consistent behaviour to that of the nominal and maximum settings should be observed. Additional measurements for these positions are not required.

c) Measurements shall be made at 1/12-octave intervals as given by the R.40 series of preferred numbers in ISO 3 [54] for frequencies from 100 Hz to 8 kHz inclusive. The averaged measured level at DRP in each frequency band is referred to the averaged test signal level measured in each frequency band.

d) The sidetone path loss (LmeST), as expressed in dB, shall be calculated from each 1/3-octave band (ITU-T Recommendation P.79 [16], table B.1, bands 4 to 17). The STMR (in dB) shall be calculated from formula B-4 of ITU-T Recommendation P.79 [16], using m = 0.225 and the weighting factors in table B.2 (unsealed condition) of ITU-T Recommendation P.79 [16]. No leakage correction (LE) shall be applied. DRP-ERP correction is used.

e) In case the STMR is below the limit, the measurement shall be repeated with the electrical sidetone path disabled and both sets of results shall be reported. In case the STMR is below the limit also with the electrical sidetone path disabled, the result shall not be regarded as a failure. Disconnecting the call is normally disabling the electrical sidetone path; otherwise the UE can be switched off to enter the wanted state.

### 7.5.3 Hands-free UE (all categories)

[...]

### 7.5.3a Electrical interface UE

a) The test signal to be used for the measurements shall be the British-English single talk sequence described in ITU-T Recommendation P.501 [22]. The active speech level of the signal shall be calibrated to -60 dBV for analogue and to -16 dBm0 for digital connections. The test signal level is calculated over the complete test signal sequence.b) The reference signal to be used for the calculation shall be the same as the test signal and is calibrated to ‑4.7 dBPa (independent of analogue or digital connection).

c) Where a user operated volume control is provided, the measurements shall be carried out at the nominal setting of the volume control. In addition, the measurement is repeated at the maximum volume control setting. It is expected that for other positions of the volume control setting a consistent behaviour to that of the nominal and maximum settings should be observed. Additional measurements for these positions are not required.

d) Measurements shall be made at 1/12-octave intervals as given by the R.40 series of preferred numbers in ISO 3 [54] for frequencies from 100 Hz to 8 kHz inclusive. The averaged measured level at the electrical receiving reference interface in each frequency band is referred to the averaged reference signal level measured in each frequency band.

e) The measured sidetone sensitivity is corrected by a default sensitivity of 22.9 dBPa/V for analogue and 2.1 dBPa/V for digital connections (corresponding both to a binaural narrowband RLR of 8 dB). This correction transfers the measured electrical sensitivity via an ideal headset (assuming a flat transfer function regarding ERP) to the acoustical domain.

NOTE: The difference in dB between nominal receiving levels of analogue (-39 dBV) and digital (-16 dBm0 = ‑18.2 dBV) connection equals 20.8 dB. This offset is taken into account for the default sensitivity of the analogue connection (22.9 dBPa/V - 20.8 dB = 2.1 dBPa/V).

f) The sidetone path loss and the STMR (in dB) shall be calculated from formula 5-1 of ITU-T P.79 [16], using m=0.225 and the weighting factors in Table 3 of ITU-T P.79 [16]. Leakage correction shall not be applied.

### 7.5.4 Sidetone delay for handset, headset or electrical interface UE

a) The handset or headset terminal is setup as described in clause 5.

b) The test signal is a CS-signal complying with ITU-T Recommendation P.501 using a PN-sequence with a length, T, of 4 096 points (for a 48 kHz sample rate test system). The duration of the complete test signal is as specified in ITU-T Recommendation P.501. The level of the signal shall be ‑4,7 dBPa at the MRP for handset or headset UE. For electrical interface UE, the level of the signal shall be -60 dBV for analogue and to -16 dBm0 for digital connections.

c) The cross-correlation function Φxy(τ) between the input signal Sx(t) generated by the test system in send direction and the output signal Sy(t) measured at the artificial ear (for handset/headset UE) or at the electrical reference interface (for electrical interface UE) is calculated in the time domain:

 

d) The measurement window, *T*, shall be identical to the test signal period, T, with the measurement window synchronized to the PN-sequence of the test signal.

e) The sidetone delay is calculated from the envelope E(τ) of the cross-correlation function Φxy(τ).The envelope E(τ) is calculated by the Hilbert transformation H{xy(τ)} of the cross-correlation:

 

 

f) For handset/headset UE:

For electrical interface UE:
Since there is no direct sound produced by the artifical mouth and captured by the artifical ear, the maximum of the envelope function directly corresponds to the sidetone delay.

NOTE: It is assumed that the measured sidetone delay is less than T/2.

------------------------- END OF CHANGE 19 -------------------------

------------------------- START OF CHANGE 20 -------------------------

## 8.5 Sidetone characteristics

### 8.5.1 Connections with handset UE

The test signal to be used for the measurements shall be the British-English single talk sequence described in ITU-T Recommendation P.501 [22]. The spectrum of the acoustic signal shall be produced by the HATS. The test signal level shall be ‑4,7 dBPa measured at the MRP. The test signal level is calculated over the complete test signal sequence.

a) The handset UE is set up as described in clause 5. The application force shall be 13 N on the Type 3.3 artificial ear.

b) Where a user operated volume control is provided, the measurements shall be carried out at the nominal setting of the volume control. In addition the measurement is repeated at the maximum volume control setting. It is expected that for other positions of the volume control setting a consistent behaviour to that of the nominal and maximum settings should be observed. Additional measurements for these positions are not required,

c) Measurements shall be made at 1/12-octave intervals as given by the R.40 series of preferred numbers in ISO 3 [54] for frequencies from 100 Hz to 8 kHz inclusive. The averaged measured level at DRP in each frequency band is referred to the averaged test signal level measured in each frequency band.

d) The sidetone path loss (LmeST), as expressed in dB, shall be calculated from each 1/3-octave band (ITU-T Recommendation P.79 [16], table B.1, bands 1 to 20). The Sidetone Masking Rating (STMR), expressed in dB, shall be calculated from formula B-4 of ITU-T Recommendation P.79 [16], using m = 0.225 and the weighting factors in table B2 (unsealed condition) of ITU-T Recommendation P.79 [16]. No leakage correction (LE) shall be applied. DRP-ERP correction is used.

e) In case the STMR is below the limit, the measurement shall be repeated with the electrical sidetone path disabled and both sets of results shall be reported. In case the STMR is below the limit also with the electrical sidetone path disabled, the result shall not be regarded as a failure. Disconnecting the call is normally disabling the electrical sidetone path; otherwise the UE can be switched off to enter the wanted state.

### 8.5.2 Headset UE

The test signal to be used for the measurements shall be the British-English single talk sequence described in ITU-T Recommendation P.501 [22]. The spectrum of the acoustic signal produced by the artificial mouth is calibrated under free-field conditions at the MRP. The test signal level shall be –4,7 dBPa measured at the MRP. The test signal level is calculated over the complete test signal sequence.

a) The headset UE is set up as described in clause 5.

b) Where a user operated volume control is provided, the measurements shall be carried out at the nominal setting of the volume control. In addition the measurement is repeated at the maximum volume control setting. It is expected that for other positions of the volume control setting a consistent behaviour to that of the nominal and maximum settings should be observed. Additional measurements for these positions are not required,

c) Measurements shall be made at 1/12-octave intervals as given by the R.40 series of preferred numbers in ISO 3 [54] for frequencies from 100 Hz to 8 kHz inclusive. The averaged measured level at DRP in each frequency band is referred to the averaged test signal level measured in each frequency band.

d) The sidetone path loss (LmeST), as expressed in dB, shall be calculated from each 1/3-octave band (ITU-T Recommendation P.79 [16], table B.1, bands 1 to 20). The STMR (in dB) shall be calculated from formula B-4 of ITU‑T Recommendation P.79 [16], using m = 0.225 and the weighting factors in table B.2 (unsealed condition) of ITU-T Recommendation P.79 [16]. No leakage correction (LE) shall be applied. DRP-ERP correction is used.

e) In case the STMR is below the limit, the measurement shall be repeated with the electrical sidetone path disabled and both sets of results shall be reported. In case the STMR is below the limit also with the electrical sidetone path disabled, the result shall not be regarded as a failure. Disconnecting the call is normally disabling the electrical sidetone path; otherwise the UE can be switched off to enter the wanted state.

### 8.5.3 Hands-free UE (all categories)

No requirement other than echo control.

### 8.5.3a Electrical interface UE

a) The test signal to be used for the measurements shall be the British-English single talk sequence described in ITU-T Recommendation P.501 [22]. The active speech level of the signal shall be calibrated to -60 dBV for analogue and to -16 dBm0 for digital connections. The test signal level is calculated over the complete test signal sequence.b) The reference signal to be used for the calculation shall be the same as the test signal and is calibrated to ‑4.7 dBPa (independent of analogue or digital connection).

c) Where a user operated volume control is provided, the measurements shall be carried out at the nominal setting of the volume control. In addition, the measurement is repeated at the maximum volume control setting. It is expected that for other positions of the volume control setting a consistent behaviour to that of the nominal and maximum settings should be observed. Additional measurements for these positions are not required.

d) Measurements shall be made at 1/12-octave intervals as given by the R.40 series of preferred numbers in ISO 3 [54] for frequencies from 100 Hz to 8 kHz inclusive. The averaged measured level at the electrical receiving reference interface in each frequency band is referred to the averaged reference signal level measured in each frequency band.

e) The measured sidetone sensitivity is corrected by a default sensitivity of 21.7 dBPa/V for analogue and 0.9 dBPa/V for digital connections (corresponding both to a binaural wideband RLR of 8 dB). This correction transfers the measured electrical sensitivity via an ideal headset (assuming a flat transfer function regarding ERP) to the acoustical domain.

NOTE: The difference in dB between nominal receiving levels of analogue (-39 dBV) and digital (-16 dBm0 = -18.2 dBV) connection equals 20.8 dB. This offset is taken into account for the default sensitivity of the analogue connection (21.7 dBPa/V - 20.8 dB = 0.9 dBPa/V).

f) The sidetone path loss and the STMR (in dB) shall be calculated from formula 5-1 of ITU-T P.79 [16], using m=0.225 and the weighting factors in Table 3 of ITU-T P.79 [16]. Leakage correction shall not be applied.

### 8.5.4 Sidetone delay for handset, headset or electrical interface UE

a) The handset or headset terminal is setup as described in clause 5.

b) The test signal is a CS-signal complying with ITU-T Recommendation P.501 using a PN-sequence with a length, T, of 4 096 points (for a 48 kHz sample rate test system). The duration of the complete test signal is as specified in ITU-T Recommendation P.501. The level of the signal shall be ‑4,7 dBPa at the MRP for handset or headset UE. For electrical interface UE, the level of the signal shall be -60 dBV for analogue and to -16 dBm0 for digital connections.

c) The cross-correlation function Φxy(τ) between the input signal Sx(t) generated by the test system in send direction and the output signal Sy(t) measured at the artificial ear (for handset/headset UE) or at the electrical reference interface (for electrical interface UE) is calculated in the time domain:

  (1)

d) The measurement window, *T*, shall be identical to the test signal period, T, with the measurement window synchronized to the PN-sequence of the test signal.

e) The sidetone delay is calculated from the envelope E(τ) of the cross-correlation function Φxy(τ).The envelope E(τ) is calculated by the Hilbert transformation H {xy(τ)} of the cross-correlation:

  (2)

  (3)

f) For handset/headset UE:

For electrical interface UE:
Since there is no direct sound produced by the artifical mouth and captured by the artifical ear, the maximum of the envelope function directly corresponds to the sidetone delay.

NOTE: It is assumed that the measured sidetone delay is less than T/2.

------------------------- END OF CHANGE 20 -------------------------

------------------------- START OF CHANGE 21 -------------------------

## 9.5 Sidetone characteristics

### 9.5.1 Connections with handset UE

The test method is the same as for wideband (see sub-clause 8.5.1).

### 9.5.2 Headset UE

The test method is the same as for wideband (see sub-clause 8.5.2).

### 9.5.3 Hands-free UE (all categories)

No requirement other than echo control.

### 9.5.3a Electrical interface UE

The test method is the same as for wideband (see sub-clause 8.5.3a).

### 9.5.4 Sidetone delay for handset, headset or electrical interface UE

The test method is the same as for wideband (see sub-clause 8.5.4).

------------------------- END OF CHANGE 21 -------------------------

------------------------- START OF CHANGE 22 -------------------------

## 10.5 Sidetone characteristics

### 10.5.1 Connections with handset UE

The test method is the same as for super-wideband (see sub-clause 9.5.1).

### 10.5.2 Headset UE

The test method is the same as for super-wideband (see sub-clause 9.5.2).

### 10.5.3 Hands-free UE (all categories)

No requirement other than echo control.

### 10.5.3a Electrical interface UE

The test method is the same as for super-wideband (see sub-clause 9.5.3a).

### 10.5.4 Sidetone delay for handset, headset or electrical interface UE

The test method is the same as for super-wideband (see sub-clause 9.5.4).

------------------------- END OF CHANGE 22 -------------------------

------------------------- START OF CHANGE 23 -------------------------

## 7.7 Acoustic echo control

### 7.7.1 General

The echo loss (EL) presented by the GSM, 3G, LTE or NR networks at the POI should be at least 46 dB during single talk. This value takes into account the fact that UE is likely to be used in a wide range of noise environments.

The calculation of weighted terminal coupling loss (TCLw) is based on the attenuation from reference point input to reference point output versus frequency bands. The following common measurement steps are applicable for all types of UE described below:

a)

b)

c) The analysis shall be conducted in 1/3-octave band intervals between 300 to 3400 Hz as given by the R.10 series of preferred numbers in ISO 3 [54].

d)

### 7.7.2 Acoustic echo control in a hands-free UE

The hands-free UE is setup in a room with acoustic properties similar to a typical "office-type" room; a vehicle-mounted hands-free UE should be tested in a vehicle or vehicle simulator, as specified by the UE manufacturer (see also 3GPP TS 03.58 [11]). The ambient noise level ≤ 70 dBPa(A).

The TCLw is measured and calculated according to clause 7.7.1.

### 7.7.3 Acoustic echo control in handset UE

The handset UE is set up according to clause 5. The ambient noise level shall be ≤ ‑64 dBPa(A).

The TCLw is measured and calculated according to clause 7.7.1.

### 7.7.4 Acoustic echo control in a headset UE

The headset UE is set up according to clause 5. The ambient noise level shall be ≤ ‑64 dBPa(A).

The TCLw is calculated according to clause 7.7.1.

### 7.7.5 Acoustic echo control in a electrical interface UE

The electrical interface UE is setup according to clause 5.1.6. In order to simulate an acoustic echo, the electrical reference interface shall introduce an echo loss of 30 dB.

The TCLw is measured and calculated according to clause 7.7.1.

------------------------- END OF CHANGE 23 -------------------------

------------------------- START OF CHANGE 24 -------------------------

## 8.7 Acoustic echo control

### 8.7.1 General

The echo loss (EL) presented by the GSM, 3G, LTE, NR or WLAN networks at the POI should be at least 46 dB during single talk. This value takes into account the fact that UE is likely to be used in a wide range of noise environments.

The calculation of terminal coupling loss (TCL) is based on the attenuation from reference point input to reference point output versus frequency bands. The following common measurement steps are applicable for all types of UE described below:

a) The attenuation from reference point input to reference point output shall be measured using the compressed real speech signal described in clause 7.3.3 of ITU-T P.501 Amendment 1 [33]. The test signal level shall be ‑10 dBm0.

b) The first 17,0 s of the test signal (6 sentences) are discarded from the analysis to allow for convergence of the acoustic echo canceller. The analysis is performed over the remaining length of the test sequence (last 6 sentences).

c) The analysis shall be conducted in 1/3-octave band intervals as given by the R.10 series of preferred numbers in ISO 3 [54]. For the calculation, the averaged measured echo level at each frequency band is referred to the averaged test signal level measured in each frequency band.

d) The TCL is calculated according to ITU-T Recommendation G.122 [8], annex B, clause B.4 (trapezoidal rule), but using the frequency range between 300 to 6700 Hz (instead of 300 Hz to 3400 Hz).

### 8.7.2 Acoustic echo control in a hands-free UE

The hands-free UE is setup in a room with acoustic properties similar to a typical "office-type" room; a vehicle-mounted hands-free UE should be tested in a vehicle or vehicle simulator, as specified by the UE manufacturer (see also 3GPP TS 03.58 [11]). The ambient noise level shall be ≤ -70 dBPa(A).

The TCL is measured and calculated according to clause 8.7.1.

### 8.7.3 Acoustic echo control in a handset UE

The handset UE is set up according to clause 5. The ambient noise level shall be ≤ -64 dBPa(A).

The TCL is measured and calculated according to clause 8.7.1.

### 8.7.4 Acoustic echo control in a headset UE

The headset is set up according to clause 5. The ambient noise level shall be ≤ -64 dBPa(A).

The TCL is measured and calculated according to clause 8.7.1.

### 8.7.5 Acoustic echo control in a electrical interface UE

The electrical interface UE is setup according to clause 5.1.6. In order to simulate an acoustic echo, the electrical reference interface shall introduce an echo loss of 30 dB.

The TCL is measured and calculated according to clause 8.7.1.

------------------------- END OF CHANGE 24 -------------------------

------------------------- START OF CHANGE 25 -------------------------

## 9.7 Acoustic echo control

### 9.7.1 General

The echo loss (EL) presented by the GSM, 3G, LTE, NR or WLAN networks at the POI should be at least 46 dB during single talk. This value takes into account the fact that UE is likely to be used in a wide range of noise environments.

NOTE: A test method fully adapted to super-wideband acoustic echo control is for further study.

The calculation of terminal coupling loss (TCL) is based on the attenuation from reference point input to reference point output versus frequency bands. The following common measurement steps are applicable for all types of UE described below:

a) The attenuation from reference point input to reference point output shall be measured using the compressed real speech signal described in clause 7.3.3 of ITU-T P.501 Amendment 1 [33]. The test signal level shall be ‑10 dBm0.

b) The first 17,0 s of the test signal (6 sentences) are discarded from the analysis to allow for convergence of the acoustic echo canceller. The analysis is performed over the remaining length of the test sequence (last 6 sentences).

c) The analysis shall be conducted in 1/3-octave band intervals as given by the R.10 series of preferred numbers in ISO 3 [54]. For the calculation, the averaged measured echo level at each frequency band is referred to the averaged test signal level measured in each frequency band.

d) The TCL is calculated according to ITU-T Recommendation G.122 [8], annex B, clause B.4 (trapezoidal rule), but using the frequency range between 300 to 6700 Hz (instead of 300 Hz to 3400 Hz).

### 9.7.2 Acoustic echo control in a hands-free UE

The hands-free UE is setup in a room with acoustic properties similar to a typical "office-type" room; a vehicle-mounted hands-free UE should be tested in a vehicle or vehicle simulator, as specified by the UE manufacturer (see also 3GPP TS 03.58 [11]). The ambient noise level shall be ≤ -70 dBPa(A).

The TCL is measured and calculated according to clause 9.7.1.

### 9.7.3 Acoustic echo control in a handset UE

The handset UE is set up according to clause 5. The ambient noise level shall be ≤ -64 dBPa(A).

The TCL is measured and calculated according to clause 9.7.1.

### 9.7.4 Acoustic echo control in a headset UE

The headset is set up according to clause 5. The ambient noise level shall be ≤ -64 dBPa(A).

The TCL is measured and calculated according to clause 9.7.1.

### 9.7.5 Acoustic echo control in a electrical interface UE

The electrical interface UE is setup according to clause 5.1.6. In order to simulate an acoustic echo, the electrical reference interface shall introduce an echo loss of 30 dB.

The TCL is measured and calculated according to clause 9.7.1.

------------------------- END OF CHANGE 25 -------------------------

------------------------- START OF CHANGE 26 -------------------------

## 10.7 Acoustic echo control

### 10.7.1 General

The description is the same as for super-wideband (see sub-clause 9.7.1).

### 10.7.2 Acoustic echo control in a hands-free UE

The test method is the same as for super-wideband (see sub-clause 9.7.2, observing the signal properties for fullband described in sub-clause 5.4).

### 10.7.3 Acoustic echo control in a handset UE

The test method is the same as for super-wideband (see sub-clause 9.7.3, observing the signal properties for fullband described in sub-clause 5.4).

### 10.7.4 Acoustic echo control in a headset UE

The test method is the same as for super-wideband (see sub-clause 9.7.4, observing the signal properties for fullband described in sub-clause 5.4).

### 10.7.5 Acoustic echo control in a electrical interface UE

The test method is the same as for super-wideband (see sub-clause 9.7.5, observing the signal properties for fullband described in sub-clause 5.4).

------------------------- END OF CHANGE 26 -------------------------

------------------------- START OF CHANGE 27 -------------------------

## 7.8 Distortion

### 7.8.1 Sending distortion

The handset, headset, or hands-free UE is setup as described in clause 5. The electrical interface UE is setup as described in clause 5.1.6.

The test signal used is a sine‑wave signal with a frequency of 1020 Hz. The sine-wave signal level shall be calibrated to the following RMS levels:

- For handset, headset, or hands-free UE: 5, 0, ‑4.7, ‑10, ‑15 and ‑20 dBPa at the MRP.

- For electrical interface UE with analogue connection: -50, -55, -60, -65, -70 and -75 dBV at the output of the electrical reference interface.

- For electrical interface UE with digital connection: -6, -11, -16, -21, -26 and -31 dBm0 at the output of the electrical reference interface.

The test signals have to be applied in this sequence, i.e., from high levels down to low levels.

The duration of the sine-wave signal is recommended to be 360 ms. The manufacturer shall be allowed to request tone lengths up to 1 s. The measured part of the signal for analysis are integer multiple of 85.333 ms and shall be at least 170.667 ms (which equals 2 \* 4096 samples in a 48 kHz sample rate test system). The times are selected to be relatively short in order to reduce the risk that the test tone is treated as a stationary signal.

It is recommended that an optional activation signal be presented immediately preceding each test signal to ensure that the UE is in a typical state during measurement. An appropriate speech or speech-like activation signal shall be chosen from ITU-T Recommendations P.501 [22] or P.50 [10]. A recommendation for the use of an activation signal as part of the measurement is defined in figure 16. The RMS level of the active parts of this activation signal is recommended to be equal to the subsequent test tone RMS level. In practice, certain types of processing may be impacted due to the introduction of the activation signal. The manufacturer shall be allowed to specify disabling of the activation signal. It shall be reported whether an activation signal was used or not, along with the characteristics of the activation signal, as specified by the manufacturer.

The ratio of the signal to total distortion power of the signal output of the SS shall be measured with the psophometric noise weighting (see ITU‑T Recommendations G.712 [21], O.41 [23] and O.132 [27]). The psophometric filter shall be normalized (0 dB gain) at 800 Hz as specified in ITU-T Recommendation O.41 [23]. The weighting function shall be applied to the total distortion component only (not to the signal component).

For measurement of the total distortion component an octave-wide band-stop filter shall be applied to the signal to suppress the sine-wave signal and associated coding artefacts. The filter shall have a lower passband ending at 0.7071 \* fS, and an upper passband starting at 1,4142 \* fS, where fS is the frequency of the sine-wave signal. The passband ripple of the filter shall be ≤ 0.2 dB. The attenuation of the band-stop filter at the sine-wave frequency shall be ≥ 60 dB. Alternatively, the described characteristics can be implemented by an appropriate weighting on the spectrum obtained from an FFT (transformation length 4096, 75% overlap, Hann window). The total distortion component is defined as the measured signal within the frequency range 200 Hz to 4 kHz, after applying psophometric and stop filters (hence no correction for the lost power due to the stop filter, known as "bandwidth correction", shall be applied).

To improve repeatability, considering the variability introduced by speech coding and voice processing, the test sequence (activation signal followed by the test signal) may be contiguously repeated one or more times. The single signal-to-total-distortion power ratios obtained from such repeats shall be averaged. The total result shall be 10 \* log10 of this average in dB.



Figure 16: Recommended activation sequence and test signal.

The activation signal consists of a "Bandlimited composite source signal with speech-like power density spectrum" signal according to ITU-T Recommendation P.501 [22] with 48,62 ms voiced part (1), 200 ms unvoiced part (2) and 101,38 ms pause (3), followed by the same signal but polarity inverted (4, 5, 6), followed by the voiced part only (7). The pure test tone is applied and after 50 ms settling time (8), the analysis is made over the following 170,667 ms (9).

NOTE 1: Void.

NOTE 2: In order to ensure that the correct part of the signal is analyzed, the total delay of the terminal and SS may have to be determined prior to the measurement.

NOTE 3: For hands-free terminals tested in environments defined in subclause 6.1.2, care should be taken that the reverberation in the test room, caused by the activation signal, does not affect the test results to an unacceptable degree, referring to subclause 5.3.

### 7.8.2 Receiving distortion

The handset, headset, or hands-free UE is setup as described in clause 5. The electrical interface UE is setup as described in clause 5.1.6.

The test signal used is a sine‑wave signal with frequencies 315, 408, 510, 816 and 1020 Hz. The signal level shall be ‑16 dBm0, except for the sine‑wave signal with a frequency 1020 Hz that shall be applied at the signal input of the SS at the following levels: 0, ‑3, ‑10, ‑16, ‑20, ‑30, ‑40, ‑45 dBm0. The test signals have to be applied in this sequence, i.e., from high levels down to low levels.

The duration of the sine-wave signal is recommended to be 360 ms. The manufacturer shall be allowed to request tone lengths up to 1 s. The measured part of the signal shall be 170.667 ms (which equals 2 \* 4096 samples in a 48 kHz sample rate test system). The times are selected to be relatively short in order to reduce the risk that the test tone is treated as a stationary signal.

It is recommended that an optional activation signal be presented immediately preceding each test signal to ensure that the UE is in a typical state during measurement. An appropriate speech or speech-like activation signal shall be chosen from ITU-T Recommendations P.501 [22] or P.50 [10]. A recommendation for the use of an activation signal as part of the measurement is defined in figure 17. The RMS level of the active parts of this activation signal is recommended to be equal to the subsequent test tone RMS level for low and medium test levels. To avoid saturation of the SS speech encoder, it is recommended for high test levels that the activation signal level be adjusted such that its peak level equals the peak level of the test tone. In practice, certain types of processing may be impacted due to the introduction of the activation signal. The manufacturer shall be allowed to specify disabling of the activation signal. It shall be reported whether an activation signal was used or not, along with the characteristics of the activation signal, as specified by the manufacturer.

The ratio of the signal to total distortion power shall be measured at:

- the applicable acoustic measurement point (DRP with diffuse-field correction for handset and headset modes; free field for hands-free modes) in case of handset, headset, or hands-free UE.

- the applicable electric measurement point (input to the electrical reference interface) in case of electrical interface UE.

Psophometric noise weighting (see ITU‑T Recommendations G.712 [21], O.41 [23] and O.132 [27]) shall be applied to the measured signal. The psophometric filter shall be normalized to have 0 dB gain at 800 Hz as specified in ITU-T Recommendation O.41 [23]. The weighting function shall be applied to the total distortion component only (not to the signal component).

For measurement of the total distortion component an octave-wide band-stop filter shall be applied to the signal to suppress the sine-wave signal and associated coding artefacts. The filter shall have a lower passband ending at 0,7071 \* fS, and an upper passband starting at 1,4142 \* fS, where fS is the frequency of the sine-wave signal. The passband ripple of the filter shall be ≤ 0.2 dB. The attenuation of the band-stop filter at the sine-wave frequency shall be ≥ 60 dB. Alternatively, the described characteristics can be implemented by an appropriate weighting on the spectrum obtained from an FFT (transformation length 4096, 75% overlap, Hann window). The total distortion component is defined as the measured signal within the frequency range 200 Hz to 4 kHz, after applying psophometric and stop filters (hence no correction for the lost power due to the stop filter, known as "bandwidth correction", shall be applied).

To improve repeatability, considering the variability introduced by speech coding and voice processing, the test sequence (activation signal followed by the test signal) may be contiguously repeated one or more times. The single signal-to-total-distortion power ratios obtained from such repeats shall be averaged. The total result shall be 10 \* log10 of this average in dB.



Figure 17: Recommended activation sequence and test signal.

The activation signal consists of a "Bandlimited composite source signal with speech-like power density spectrum" signal according to ITU-T Recommendation P.501 with 48,62 ms voiced part (1), 200 ms unvoiced part (2) and 101,38 ms pause (3), followed by the same signal but polarity inverted (4, 5, 6), followed by the voiced part only (7). The pure test tone is applied and after 50 ms settling time (8), the analysis is made over the following 170,667 ms (9).

NOTE 1: Void.

NOTE 2: In order to ensure that the correct part of the signal is analyzed, the total delay of the terminal and SS may have to be determined prior to the measurement.

NOTE 3: For hands-free terminals tested in environments defined in subclause 6.1.2, care should be taken that the reverberation in the test room, caused by the activation signal, does not affect the test results to an unacceptable degree, referring to subclause 5.3.

------------------------- END OF CHANGE 27 -------------------------

------------------------- START OF CHANGE 28 -------------------------

## 8.8 Distortion

### 8.8.1 Sending distortion

The handset, headset, or hands-free UE is setup as described in clause 5. The electrical interface UE is setup as described in clause 5.1.6.

The test signal used is a sine-wave signal with frequencies of 315, 408, 510, 816 and 1020 Hz. The sine-wave signal level shall be calibrated to the following RMS levels:

- For handset, headset, or hands-free UE: ‑4.7 dBPa at the MRP for all frequencies. For the sine-wave with a frequency of 1020 Hz, levels of 5, 0, ‑4.7, ‑10, ‑15, ‑20 dBPa shall be applied.

- For electrical interface UE with analogue connection: ‑60 dBV for all frequencies at the output of the electrical reference interface. For the sine-wave with a frequency of 1020 Hz, levels of -50, -55, -60, -65, -70 and -75 dBV shall be applied.

- For electrical interface UE with digital connection: ‑16 dBm0 for all frequencies at the output of the electrical reference interface. For the sine-wave with a frequency of 1020 Hz, levels of -6, -11, -16, -21, -26 and -31 dBm0 shall be applied.

The test signals have to be applied in this sequence, i.e., from high levels down to low levels.

The duration of the sine-wave signal is recommended to be 360 ms. The manufacturer shall be allowed to request tone lengths up to 1 s. The measured part of the signal for analysis are integer multiple of 85.333 ms and shall be at least 170,667 ms (which equals 2 \* 4096 samples in a 48 kHz sample rate test system). The times are selected to be relatively short in order to reduce the risk that the test tone is treated as a stationary signal.

It is recommended that an optional activation signal be presented immediately preceding each test signal to ensure that the UE is in a typical state during measurement (see Note 1.). An appropriate speech or speech-like activation signal shall be chosen from ITU-T Recommendations P.501 [22] or P.50 [10]. A recommendation for the use of an activation signal as part of the measurement is defined in figure 18. The RMS level of the active parts of this activation signal is recommended to be equal to the subsequent test tone RMS level. In practice, certain types of processing may be impacted due to the introduction of the activation signal. The manufacturer shall be allowed to specify disabling of the activation signal. It shall be reported whether an activation signal was used or not, along with the characteristics of the activation signal, as specified by the manufacturer.

The ratio of the signal to total distortion power of the signal output of the SS shall be measured with the psophometric noise weighting (see ITU‑T Recommendations G.712 [21], O.41 [23] and O.132 [27]). The psophometric filter shall be normalized (0 dB gain) at 800 Hz as specified in ITU-T Recommendation O.41 [23]. The weighting function shall be applied to the total distortion component only (not to the signal component).

For measurement of the total distortion component an octave-wide band-stop filter shall be applied to the signal to suppress the sine-wave signal and associated coding artefacts. The filter shall have a lower passband ending at 0.7071 \* fS, and an upper passband starting at 1.4142 \* fS, where fS is the frequency of the sine-wave signal. The passband ripple of the filter shall be ≤ 0,2 dB. The attenuation of the band-stop filter at the sine-wave frequency shall be ≥ 60 dB. Alternatively, the described characteristics can be implemented by an appropriate weighting on the spectrum obtained from an FFT (transformation length 4096, 75% overlap, Hann window). The total distortion component is defined as the measured signal within the frequency range 100 Hz to 6 kHz, after applying psophometric and stop filters (hence no correction for the lost power due to the stop filter, known as "bandwidth correction", shall be applied).

To improve repeatability, considering the variability introduced by speech coding and voice processing, the test sequence (activation signal followed by the test signal) may be contiguously repeated one or more times. The single signal-to-total-distortion power ratios obtained from such repeats shall be averaged. The total result shall be 10 \* log10 of this average in dB.



Figure 18: Recommended activation sequence and test signal.

The activation signal consists of a "Bandlimited composite source signal with speech-like power density spectrum" signal according to ITU-T Recommendation P.501 [22] with 48,62 ms voiced part (1), 200 ms unvoiced part (2) and 101,38 ms pause (3), followed by the same signal but polarity inverted (4, 5, 6), followed by the voiced part only (7). The pure test tone is applied and after 50 ms settling time (8), the analysis is made over the following 170,667 ms (9).

NOTE 1: Depending on the type of codec the test signal used may need to be adapted. If a sine-wave is not usable, an alternative test signal could be a band-limited noise signal centered on the above frequencies.

NOTE 2: Void.

NOTE 3: Void.

NOTE 4: In order to ensure that the correct part of the signal is analyzed, the total delay of the terminal and SS may have to be determined prior to the measurement.

NOTE 5: For hands-free terminals tested in environments defined in subclause 6.1.2, care should be taken that the reverberation in the test room, caused by the activation signal, does not affect the test results to an unacceptable degree, referring to subclause 5.3.

### 8.8.2 Receiving distortion

The handset, headset, or hands-free UE is setup as described in clause 5. The electrical interface UE is setup as described in clause 5.1.6.

The test signal used is a sine‑wave signal with frequencies 315, 408, 510, 816 and 1020 Hz. The signal level shall be ‑16 dBm0, except for the sine‑wave signal with a frequency 1020 Hz that shall be applied at the signal input of the SS at the following levels: 0, ‑3, ‑10, ‑16, ‑20, ‑30, ‑40, ‑45 dBm0. The test signals have to be applied in this sequence, i.e., from high levels down to low levels.

The duration of the sine-wave signal is recommended to be 360 ms. The manufacturer shall be allowed to request tone lengths up to 1 s. The measured part of the signal shall be 170,667 ms (which equals 2 \* 4096 samples in a 48 kHz sample rate test system). The times are selected to be relatively short in order to reduce the risk that the test tone is treated as a stationary signal.

It is recommended that an optional activation signal be presented immediately preceding each test signal to ensure that the UE is in a typical state during measurement (see Note 1.). An appropriate speech or speech-like activation signal shall be chosen from ITU-T Recommendations P.501 [22] or P.50 [10]. A recommendation for the use of an activation signal as part of the measurement is defined in figure 19. The RMS level of the active parts of this activation signal is recommended to be equal to the subsequent test tone RMS level for low and medium test levels. To avoid saturation of the SS speech encoder, it is recommended for high test levels that the activation signal level is adjusted so that its peak level equals the peak level of the test tone. In practice, certain types of processing may be impacted due to the introduction of the activation signal. The manufacturer shall be allowed to specify disabling of the activation signal. It shall be reported whether an activation signal was used or not, along with the characteristics of the activation signal, as specified by the manufacturer.

The ratio of the signal to total distortion power shall be measured at:

- the applicable acoustic measurement point (DRP with diffuse-field correction for handset and headset modes; free field for hands-free modes) in case of handset, headset, or hands-free UE.

- the applicable electric measurement point (input to the electrical reference interface) in case of electrical interface UE.

Psophometric noise weighting (see ITU‑T Recommendations G.712 [21], O.41 [23] and O.132 [27]) shall be applied to the measured signal. The psophometric filter shall be normalized to have 0 dB gain at 800 Hz as specified in ITU-T Recommendation O.41 [23]. The weighting function shall be applied to the total distortion component only (not to the signal component).

For measurement of the total distortion component an octave-wide band-stop filter shall be applied to the signal to suppress the sine-wave signal and associated coding artefacts. The filter shall have a lower passband ending at 0,7071 \* fS, and an upper passband starting at 1,4142 \* fS, where fS is the frequency of the sine-wave signal. The passband ripple of the filter shall be ≤ 0,2 dB. The attenuation of the band stop filter at the sine-wave frequency shall be ≥ 60 dB. Alternatively the described characteristics can be implemented by an appropriate weighting on the spectrum obtained from an FFT (transformation length 4096, 75% overlap, Hann window). The total distortion component is defined as the measured signal within the frequency range 100 Hz to 6 kHz, after applying psophometric and stop filters (hence no correction for the lost power due to the stop filter, known as "bandwidth correction", shall be applied).

To improve repeatability, considering the variability introduced by speech coding and voice processing, the test sequence (activation signal followed by the test signal) may be contiguously repeated one or more times. The single signal-to-total-distortion power ratios obtained from such repeats shall be averaged. The total result shall be 10 \* log10 of this average in dB.



Figure 19: Recommended activation sequence and test signal.

The activation signal consists of a "Bandlimited composite source signal with speech-like power density spectrum" signal according to ITU-T Recommendation P.501 with 48,62 ms voiced part (1), 200 ms unvoiced part (2) and 101,38 ms pause (3), followed by the same signal but polarity inverted (4, 5, 6), followed by the voiced part only (7). The pure test tone is applied and after 50 ms settling time (8), the analysis is made over the following 170,667 ms (9).

NOTE 1: Void.

NOTE 2: Void.

NOTE 3: In order to ensure that the correct part of the signal is analyzed, the total delay of the terminal and SS may have to be determined prior to the measurement.

NOTE 4: For hands-free terminals tested in environments defined in subclause 6.1.2, care should be taken that the reverberation in the test room, caused by the activation signal, does not affect the test results to an unacceptable degree, referring to subclause 5.3.

------------------------- END OF CHANGE 28 -------------------------

------------------------- START OF CHANGE 29 -------------------------

## 9.8 Distortion

### 9.8.1 Sending distortion

[…]

### 9.8.2 Receiving distortion

[…]

------------------------- END OF CHANGE 29 -------------------------

------------------------- START OF CHANGE 30 -------------------------

## 10.8 Distortion

### 10.8.1 Sending distortion

[…]

### 10.8.2 Receiving distortion

[…]

------------------------- END OF CHANGE 30 -------------------------

------------------------- START OF CHANGE 31 -------------------------

## 7.10 Delay

### 7.10.0 UE Delay Measurement Methodologies

[...]

### 7.10.1 Delay in sending direction (Handset UE)

[...]

### 7.10.1a Delay in sending direction (headset UE)

[...]

### 7.10.1b Delay in sending direction (electrical interface UE)

The UE delay TS in the sending direction is obtained by measuring the delay between output of the electrical reference interface and the electrical access point of the test equipment; delays introduced by the test equipment are subtracted from the measured value.



Figure 17b2a: Different entities when measuring the delay in sending direction through electical interface UE

The overall delay measured from output of the electrical reference interface to the electrical access point of the test equipment is TS + TTES, as illustrated in Figure 17b2a.

The test method is the same as for handset UE (clause 7.10.1), except that the source levels are as follows:

- for analogue connections, -60 dBV at electrical reference interface output.

- for digital connection, -16 dBm0 at electrical reference interface output.

### 7.10.2 Delay in receiving direction (handset UE)

[...]

### 7.10.2a Delay in receiving direction (headset UE)

[...]

### 7.10.2b Delay in receiving direction (electrical interface UE)

The UE delay TR in the receiving direction is obtained by measuring the delay between the electrical access point of the test equipment and the input of the electical reference interface; delays introduced by the test equipment are subtracted from the measured value.



Figure 17b4a: Different entities when measuring the delay in receiving direction through electical interface UE

The overall delay measured from the electrical access point of the test equipment to the input of the electrical reference interface is TR + TTER, as illustrated in Figure 17b4a.

The test method is the same as for handset UE (clause 7.10.2).

### 7.10.3 Delay in sending + receiving direction using "echo" method (handset UE)

[...]

### 7.10.3a Delay in sending + receiving direction using "echo" method (headset UE)

[...]

### 7.10.3b Delay in sending + receiving direction using "echo" method (electrical interface UE)

The UE delay is obtained by measuring the delay between the input and output of the electrical reference interface; delays introduced by the test equipment and system simulator, TSS, is subtracted from the measured value.

The test method is the same as for handset UE (clause 7.10.3), except that the source levels are as follows:

- for analogue connections, -60 dBV at electrical reference interface output.

- for digital connection, -16 dBm0 at electrical reference interface output.

### 7.10.4 Delay and speech quality in conditions with packet arrival time variations and packet loss (handset, headset, electrical interface UE)

#### 7.10.4.1 Delay in sending direction

The UE delay in the sending direction, TS, shall be measured in jitter and error free conditions according to clause 7.10.0.

#### 7.10.4.2 Delay in receiving direction

For this test it shall be ensured that the call is originated from the UE.

NOTE 1: Differences have been observed between UE-originated calls and UE-terminated calls. For better consistency, calls from the UE are used.

The test signal consists of 3 repeats of the Composite Source Signal (CSS) according to ITU-T Recommendation P.501 [22] followed by a speech signal of 160s. During the first two CSS signals the terminal can adapt its jitter buffer. The third CSS is used for measuring the delay in constant-delay condition, and the speech signal is used for delay and quality measurement in the packet impairment condition.

Constant delay Tc corresponding to the minimum delay of the profile (i.e. the compensation value for the profile) shall be added at the beginning of the different delay/loss profiles, to avoid unecessary delay jumps between the two measurement phases and realistic conditions for the second measurement test phase.

In receiving direction, the delay between the electrical access point of the test equipment and the reference point (RP), TTEAP-RP(t) = TR-jitter(t)+ TTER, is measured in two successive phases:

1) First the delay in constant-delay condition TTEAP-RP-constant is measured as described in steps 1 to 4, clause 7.10.2/7.10.2a/7.10.2b, using the third CSS signal. The constant delay Tc is subtracted from TTEAP-RP-constant to obtain TR-constant.

2) Then the delay with packet impairment TR-jitter(t) is measured continuously for a speech signal during the inclusion of packet delay and loss profiles in the receiving direction RTP voice stream.

The reference point is defined as follows:

- for handset and headset UE, the reference point is the DRP.

- for electrical interface UE, the reference point is the input of the electrical reference interface.

Packet impairments shall be applied between the reference client and system simulator eNodeB. Separate calls shall be established for each packet impairment condition.

The start of the delay profiles must be synchronized with the start of the downlink speech material reproduction (compensated by the delay between reproduction and the point of impairment insertion, i.e. the delay of the reference client) in order to ensure a repeatable application of impairments to the test speech signal. Tests shall be performed with DTX enabled in the reference client.

NOTE 2: RTP packet impairments representing packet delay variations and loss in LTE transmission scenarios are specified in Annex E. These LTE jitter/loss profiles are reused also for tests with WLAN and NR access. Care must be taken that the system simulator uses a dedicated bearer with no buffering/scheduling of packets for transmission.

For the CSS signal repeated 3 times, the pseudo random noise (pn)-part of the CSS has to be longer than the maximum expected delay. It is recommended to use a pn sequence of 32 k samples (with 48 kHz sampling rate). The test signal level is -16 dBm0 measured at the digital reference point or the equivalent analogue point.

For the speech signal, 8 English test sentences according to ITU-T P.501 Annex C.2.3, normalized to an active speech level of -16dBm0, are used (2 male, 2 female speakers). The sequences are concatenated in such a way that all sentences are centered within a 4.0s time window, which results in an overall duration of 32.0s. The sequences are repeated 5 times, resulting in a test file 160.0s long. The first 2 sentences are used for convergence of the UE jitter buffer manager and are discarded from the analysis. Equivalent implementations of the concatenation by repeating the test sentences in sequence may be used.

For the delay calculation with the speech signal, a cross-correlation with a rectangular window length of 4s, centered at each sentence of the stimulus file, is used. The process is repeated for each sample. For each cross correlation, the maximum of the envelope is obtained producing one delay value per sentence.

The UE delay in the receive direction, TR-jitter(t), is obtained by subtracting the delay introduced by the test equipment and the simulated transport network packet delay introduced by the delay and loss profile (as specified for the respective profile in Annex E) from the first electrical event at the electrical access point of the test equipment to the first bit of the corresponding speech frame at the system simulator antenna, TTER, from the measured TTEAP-RP(t).

The difference DT between maximum receiving delay obtained with at least 5 individual calls (see clause 7.10.2) and the delay TR-constant measured for the CSS signal in constant delay condition is calculated. The quantity "Call-to-Call Variability Adjustment" (CCVA) = max(0,DT) shall be added to the obtained delay for the speech signal TR-jitter(t).

For stationary packet delay variation test conditions (test condition 1 and 2), the first 2 sentences are used for convergence of the jitter buffer management and are discarded from the analysis. The CCVA-adjusted UE delay (TR-CCVA(t) = TR-jitter(t) + CCVA) in the receiving direction shall be reported as the maximum value excluding the two largest values of the remaining sequence of the 38 sentence delay values, i.e. the 95-percentile value of TR-CCVA(t). The TR-CCVA values for all 40 sentences shall be reported in the test report.

NOTE 3: The synchronization of the speech frame processing in the UE to the bits of the speech frames at the UE antenna may lead to a variability of up to 20 ms of the measured UE receive delay between different calls. This synchronization is attributed to the UE receiving delay according to the definition of the UE delay reference points The effect of this possible call-to-call variation is taken into account with the CCVA = max(0,DT) value.

#### 7.10.4.3 Speech quality loss in conditions with packet arrival time variations and packet loss

For the evaluation of speech quality loss in conditions with packet arrival time variations and packet loss, the test signal described in clause 7.10.4.2 shall be used. The first 2 sentences are used for convergence of the UE jitter buffer manager and are discarded from the analysis. Two 48 kHz recordings are used to produce the speech quality loss metric:

- A recording obtained in jitter and error free conditions with the test signal described in clause 7.10.4.2 (reference condition)

- A recording obtained during the application of packet arrival time variations and packet loss as described in clause 7.10.4.2 (test condition)

The speech quality of the signal is estimated using the measurement algorithm described in ITU-T Recommendation P.863 [44] in super-wideband mode. For narrowband speech, the method according to Appendix III of P.863 [44] shall be used. Level pre-alignment to -26 dBov of recordings shall be used – see P.863.1 clause 10.2 [45].

NOTE: For the analysis of acoustical measurements, ITU-T P.863 [44] assumes diffuse-field equalized recordings. For this reason, signals at DRP are diffuse-field corrected for testing handset and headset UE. For electrical interface UE, only the level pre-alignment is applied.

A score shall be computed for each 8s speech sentence pair and averaged to produce a mean MOS-LQO value for the reference and test conditions.

MOS-LQOREF

MOS-LQOTEST

NOTE: This evaluation of the speech quality requirement is only applicable to test conditions with a stationary statistic of the packet delay variation. Evaluation of the speech quality for a test condition with non-stationary packet delay variations is for further study.

The synchronization between stimuli and degraded condition shall be done by the test system before applying the P.863 algorithm on each sentence pair.

### 7.10.5 UE send clock accuracy

[...]

### 7.10.6 UE receiving with clock skew

For further study.

------------------------- END OF CHANGE 31 -------------------------

------------------------- START OF CHANGE 32 -------------------------

## 8.10 Delay

### 8.10.0 UE Delay Measurement Methodologies

[...]

### 8.10.1 Delay in sending direction (Handset UE)

[...]

### 8.10.1a Delay in sending direction (headset UE)

[...]

### 8.10.1b Delay in sending direction (electrical interface UE)

The UE delay TS in the sending direction is obtained by measuring the delay between output of the electrical reference interface and the electrical access point of the test equipment; delays introduced by the test equipment are subtracted from the measured value.



Figure 19b2a: Different entities when measuring the delay in sending direction through electical interface UE

The overall delay measured from output of the electrical reference interface to the electrical access point of the test equipment is TS + TTES, as illustrated in Figure 19b2a.

The test method is the same as for handset UE (clause 8.10.1), except that the source levels are as follows:

- for analogue connections, -60 dBV at electrical reference interface output.

- for digital connection, -16 dBm0 at electrical reference interface output.

### 8.10.2 Delay in receiving direction (handset UE)

[...]

### 8.10.2a Delay in receiving direction (headset UE)

[...]

### 8.10.2b Delay in receiving direction (electrical interface UE)

The UE delay TR in the receiving direction is obtained by measuring the delay between the electrical access point of the test equipment and the input of the electical reference interface; delays introduced by the test equipment are subtracted from the measured value.



Figure 19b4a: Different entities when measuring the delay in receiving direction through electical interface UE

The overall delay measured from the electrical access point of the test equipment to the input of the electrical reference interface is TR + TTER, as illustrated in Figure 19b4a.

The test method is the same as for handset UE (clause 8.10.2).

### 8.10.3 Delay in sending + receiving direction using "echo" method (handset UE)

[...]

### 8.10.3a Delay in sending + receiving direction using "echo" method (headset UE)

[...]

### 8.10.3b Delay in sending + receiving direction using "echo" method (electrical interface UE)

The UE delay is obtained by measuring the delay between the input and output of the electrical reference interface; delays introduced by the test equipment and system simulator, TSS, is subtracted from the measured value.

The test method is the same as for handset UE (clause 8.10.3), except that the source levels are as follows:

- for analogue connections, -60 dBV at electrical reference interface output.

- for digital connection, -16 dBm0 at electrical reference interface output.

### 8.10.4 Delay and speech quality in conditions with packet arrival time variations and packet loss (handset, headset, electrical interface UE)

#### 8.10.4.1 Delay in sending direction

The UE delay in the sending direction, TS, shall be measured in jitter and error free conditions according to clause 8.10.0.

#### 8.10.4.2 Delay in receiving direction

For this test it shall be ensured that the call is originated from the UE.

NOTE 1: Differences have been observed between UE-originated calls and UE-terminated calls. For better consistency, calls from the UE are used.

The test signal consists of 3 repeats of the Composite Source Signal (CSS) according to ITU-T Recommendation P.501 [22] followed by a speech signal of 160s. During the first two CSS signals the terminal can adapt its jitter buffer. The third CSS is used for measuring the delay in constant-delay condition, and the speech signal is used for delay and quality measurement in the packet impairment condition.

Constant delay Tc corresponding to the minimum delay of the profile (i.e. the compensation value for the profile) shall be added at the beginning of the different delay/loss profiles, to avoid unecessary delay jumps between the two measurement phases and realistic conditions for the second measurement test phase.

In receiving direction, the delay between the electrical access point of the test equipment and the reference point (RP), TTEAP-RP(t) = TR-jitter(t)+ TTER, is measured in two successive phases:

1) First the delay in constant-delay condition TTEAP-RP-constant is measured as described in steps 1 to 4, clause 8.10.2/8.10.2a/8.10.2b, using the third CSS signal. The constant delay Tc is subtracted from TTEAP-RP to obtain TR-constant.

2) Then the delay with packet impairment TR-jitter(t) is measured continuously for a speech signal during the inclusion of packet delay and loss profiles in the receiving direction RTP voice stream.

The reference point is defined as follows:

- for handset and headset UE, the reference point is the DRP.

- for electrical interface UE, the reference point is the input of the electrical reference interface.

Packet impairments shall be applied between the reference client and system simulator eNodeB. Separate calls shall be established for each packet impairment condition.

The start of the delay profiles must be synchronized with the start of the downlink speech material reproduction (compensated by the delay between reproduction and the point of impairment insertion, i.e. the delay of the reference client) in order to ensure a repeatable application of impairments to the test speech signal. Tests shall be performed with DTX enabled in the reference client.

NOTE 2: RTP packet impairments representing packet delay variations and loss in LTE transmission scenarios are specified in Annex E. These LTE jitter/loss profiles are reused also for tests with WLAN and NR access. Care must be taken that the system simulator uses a dedicated bearer with no buffering/scheduling of packets for transmission.

For the CSS signal repeated 3 times, the pseudo random noise (pn)-part of the CSS has to be longer than the maximum expected delay. It is recommended to use a pn sequence of 32 k samples (with 48 kHz sampling rate). The test signal level is -16 dBm0 measured at the digital reference point or the equivalent analogue point.

For the speech signal, 8 English test sentences according to ITU-T P.501 Annex C.2.3, normalized to an active speech level of -16dBm0, are used (2 male, 2 female speakers). The sequences are concatenated in such a way that all sentences are centered within a 4.0s time window, which results in an overall duration of 32.0s. The sequences are repeated 5 times, resulting in a test file 160.0s long. The first 2 sentences are used for convergence of the UE jitter buffer manager and are discarded from the analysis. Equivalent implementations of the concatenation by repeating the test sentences in sequence may be used.

For the delay calculation with the speech signal, a cross-correlation with a rectangular window length of 4s, centered at each sentence of the stimulus file, is used. The process is repeated for each sample. For each cross correlation, the maximum of the envelope is obtained producing one delay value per sentence.

The UE delay in the receive direction, TR-jitter(t), is obtained by subtracting the delay introduced by the test equipment and the simulated transport network packet delay introduced by the delay and loss profile (as specified for the respective profile in Annex E) from the first electrical event at the electrical access point of the test equipment to the first bit of the corresponding speech frame at the system simulator antenna, TTER, from the measured TTEAP-RP(t).

The difference DT between maximum receiving delay obtained with at least 5 individual calls (see clause 8.10.2) and the delay TR-constant measured for the CSS signal in constant delay condition is calculated. The quantity "Call-to-Call Variability Adjustment" (CCVA) = max(0,DT) shall be added to the obtained delay for the speech signal TR-jitter(t).

For stationary packet delay variation test conditions (test condition 1 and 2), the first 2 sentences are used for convergence of the jitter buffer management and are discarded from the analysis. The CCVA-adjusted UE delay (TR-CCVA(t) = TR-jitter(t) + CCVA) in the receiving direction shall be reported as the maximum value excluding the two largest values of the remaining sequence of the 38 sentence delay values, i.e. the 95-percentile value of TR-CCVA(t). The TR-CCVA values for all 40 sentences shall be reported in the test report.

NOTE 3: The synchronization of the speech frame processing in the UE to the bits of the speech frames at the UE antenna may lead to a variability of up to 20 ms of the measured UE receive delay between different calls. This synchronization is attributed to the UE receiving delay according to the definition of the UE delay reference points The effect of this possible call-to-call variation is taken into account with the CCVA = max(0,DT) value.

#### 8.10.4.3 Speech quality loss in conditions with packet arrival time variations and packet loss

For the evaluation of speech quality loss in conditions with packet arrival time variations and packet loss, the test signal described in clause 8.10.4.2 shall be used. The first 2 sentences are used for convergence of the UE jitter buffer manager and are discarded from the analysis. Two 48 kHz recordings are used to produce the speech quality loss metric:

- A recording obtained in jitter and error free conditions with the test signal described in clause 8.10.4.2 (reference condition)

- A recording obtained during the application of packet arrival time variations and packet loss as described in clause 8.10.4.2 (test condition)

The speech quality of the signal is estimated using the measurement algorithm described in ITU-T Recommendation P.863 [44] in super-wideband mode. Level pre-alignment to -26 dBov of recordings shall be used – see P.863.1 clause 10.2 [45].

NOTE: for the analysis of acoustical measurements, P.863 [44] assumes diffuse-field equalized recordings. For this reason, signals at DRP are diffuse-fieldcorrected for testing handset and headset UE. For electrical interface UE, only the level pre-alignment is applied.

A score shall be computed for each 8s speech sentence pair and averaged to produce a mean MOS-LQO value for the reference and test conditions.

MOS-LQOREF

MOS-LQOTEST 

NOTE: This evaluation of the speech quality requirement is only applicable to test conditions with a stationary statistic of the packet delay variation. Evaluation of the speech quality for a test condition with non-stationary packet delay variations is for further study.

The synchronization between stimuli and degraded condition shall be done by the test system before applying the P.863 algorithm on each sentence pair.

### 8.10.5 UE send clock accuracy

[...]

### 8.10.6 UE receiving with clock skew

[...]

------------------------- END OF CHANGE 32 -------------------------

------------------------- START OF CHANGE 33 -------------------------

## 9.10 Delay

### 9.10.0 UE Delay Measurement Methodologies

[...]

### 9.10.1 Delay in sending direction (Handset UE)

[...]

### 9.10.1a Delay in sending direction (headset UE)

[...]

### 9.10.1b Delay in sending direction (electrical interface UE)

The test method is the same as in wideband (see clause 8.10.1b).

### 9.10.2 Delay in receiving direction (handset UE)

The test method is the same as in wideband (see clause 8.10.2, observing the test signal properties defined for super-wideband described in clause 5.4).

### 9.10.2a Delay in receiving direction (headset UE)

The test method is the same as in wideband (see clause 8.10.2a, observing the test signal properties for super-wideband described in clause 5.4).

### 9.10.2b Delay in receiving direction (electrical interface UE)

The test method is the same as in wideband (see clause 8.10.2b, observing the test signal properties for super-wideband described in clause 5.4).

### 9.10.3 Delay in sending + receiving direction using "echo" method (handset UE)

The test method is the same as in wideband (see clause 8.10.3, observing the test signal properties for super-wideband described in clause 5.4).

### 9.10.3a Delay in sending + receiving direction using "echo" method (headset UE)

The test method is the same as in wideband (see clause 8.10.3a, observing the test signal properties for super-wideband in clause 5.4).

### 9.10.3b Delay in sending + receiving direction using "echo" method (electrical interface UE)

The test method is the same as in wideband (see clause 8.10.3b, observing the test signal properties for super-wideband in clause 5.4).

### 9.10.4 Delay and speech quality in conditions with packet arrival time variations and packet loss (handset, headset, electrical interface UE)

#### 9.10.4.1 Delay in sending direction

The test method is the same as in wideband (see clause 8.10.4.1).

#### 9.10.4.2 Delay in receiving direction

For this test it shall be ensured that the call is originated from the UE.

NOTE 1: Differences have been observed between UE-originated calls and UE-terminated calls. For better consistency, calls from the UE are used.

The test signal consists of 3 repeats of the Composite Source Signal (CSS) according to ITU-T Recommendation P.501 [22] followed by a speech signal of 160s. During the first two CSS signals the terminal can adapt its jitter buffer. The third CSS is used for measuring the delay in constant-delay condition, and the speech signal is used for delay and quality measurement in the packet impairment condition.

Constant delay Tc corresponding to the minimum delay of the profile (i.e. the compensation value for the profile) shall be added at the beginning of the different delay/loss profiles, to avoid unecessary delay jumps between the two measurement phases and realistic conditions for the second measurement test phase.

In receiving direction, the delay between the electrical access point of the test equipment and the reference point (RP), TTEAP-RP(t) = TR-jitter(t)+ TTER, is measured in two successive phases:

1) First the delay in constant-delay condition TTEAP-RP-constant is measured as described in steps 1 to 4, clause 9.10.2/9.10.2a/9.10.2b, using the third CSS signal. The constant delay Tc is subtracted from TTEAP-RP to obtain TR-constant.

2) Then the delay with packet impairment TR-jitter(t) is measured continuously for a speech signal during the inclusion of packet delay and loss profiles in the receiving direction RTP voice stream.

The reference point is defined as follows:

- for handset and headset UE, the reference point is the DRP.

- for electrical interface UE, the reference point is the input of the electrical reference interface.

Packet impairments shall be applied between the reference client and system simulator eNodeB. Separate calls shall be established for each packet impairment condition.

The start of the delay profiles must be synchronized with the start of the downlink speech material reproduction (compensated by the delay between reproduction and the point of impairment insertion, i.e. the delay of the reference client) in order to ensure a repeatable application of impairments to the test speech signal. Tests shall be performed with DTX enabled in the reference client.

NOTE 2: RTP packet impairments representing packet delay variations and loss in LTE transmission scenarios are specified in Annex E. These LTE jitter/loss profiles are reused also for tests with WLAN and NR access. Care must be taken that the system simulator uses a dedicated bearer with no buffering/scheduling of packets for transmission.

For the CSS signal repeated 3 times, the pseudo random noise (pn)-part of the CSS has to be longer than the maximum expected delay. It is recommended to use a pn sequence of 32 k samples (with 48 kHz sampling rate). The test signal level is -16 dBm0 measured at the digital reference point or the equivalent analogue point.

For the speech signal, 8 English test sentences according to ITU-T P.501 Annex C.2.3, normalized to an active speech level of -16dBm0, are used (2 male, 2 female speakers). The sequences are concatenated in such a way that all sentences are centered within a 4.0s time window, which results in an overall duration of 32.0s. The sequences are repeated 5 times, resulting in a test file 160.0s long. The first 2 sentences are used for convergence of the UE jitter buffer manager and are discarded from the analysis. Equivalent implementations of the concatenation by repeating the test sentences in sequence may be used.

For the delay calculation with the speech signal, a cross-correlation with a rectangular window length of 4s, centered at each sentence of the stimulus file, is used. The process is repeated for each sample. For each cross correlation, the maximum of the envelope is obtained producing one delay value per sentence.

The UE delay in the receive direction, TR-jitter(t), is obtained by subtracting the delay introduced by the test equipment and the simulated transport network packet delay introduced by the delay and loss profile (as specified for the respective profile in Annex E) from the first electrical event at the electrical access point of the test equipment to the first bit of the corresponding speech frame at the system simulator antenna, TTER, from the measured TTEAP-RP(t).

The difference DT between maximum receiving delay obtained with at least 5 individual calls (see clause 9.10.2) and the delay TR-constant measured for the CSS signal in constant delay condition is calculated. The quantity "Call-to-Call Variability Adjustment" (CCVA) = max(0,DT) shall be added to the obtained delay for the speech signal TR-jitter(t).

For stationary packet delay variation test conditions (test condition 1 and 2), the first 2 sentences are used for convergence of the jitter buffer management and are discarded from the analysis. The CCVA-adjusted UE delay (TR-CCVA(t) = TR-jitter(t) + CCVA) in the receiving direction shall be reported as the maximum value excluding the two largest values of the remaining sequence of the 38 sentence delay values, i.e. the 95-percentile value of TR-CCVA(t). The TR-CCVA values for all 40 sentences shall be reported in the test report.

NOTE 3: The synchronization of the speech frame processing in the UE to the bits of the speech frames at the UE antenna may lead to a variability of up to 20 ms of the measured UE receive delay between different calls. This synchronization is attributed to the UE receiving delay according to the definition of the UE delay reference points The effect of this possible call-to-call variation is taken into account with the CCVA = max(0,DT) value.

#### 9.10.4.3 Speech quality loss in conditions with packet arrival time variations and packet loss

The test method is the same as in wideband (see clause 8.10.4.3, observing the test signal properties for super-wideband described in clause 5.4).

### 9.10.5 UE send clock accuracy

[...]

### 9.10.6 UE receiving with clock skew

[...]

------------------------- END OF CHANGE 33 -------------------------

------------------------- START OF CHANGE 34 -------------------------

## 10.10 Delay

### 10.10.0 UE Delay Measurement Methodologies

The test method is the same as in super-wideband (see clause 9.10.0).

### 10.10.1 Delay in sending direction (handset UE)

The test method is the same as in super-wideband (see clause 9.10.1)..

### 10.10.1a Delay in sending direction (headset UE)

The test method is the same as in super-wideband (see clause 9.10.1a).

### 10.10.1b Delay in sending direction (electrical interface UE)

The test method is the same as in super-wideband (see clause 9.10.1b).

### 10.10.2 Delay in receiving direction (handset UE)

The test method is the same as in super-wideband (see clause 9.10.2, observing the test signal properties for fullband described in clause 5.4).

### 10.10.2a Delay in receiving direction (headset UE)

The test method is the same as in super-wideband (see clause 9.10.2a, observing the test signal properties for fullband described in clause 5.4).

### 10.10.2b Delay in receiving direction (electrical interface UE)

The test method is the same as in super-wideband (see clause 9.10.2b, observing the test signal properties for fullband described in clause 5.4).

### 10.10.3 Delay in sending + receiving direction using "echo" method (handset UE)

The test method is the same as in super-wideband (see clause 9.10.3, observing the test signal properties for fullband described in clause 5.4).

### 10.10.3a Delay in sending + receiving direction using "echo" method (headset UE)

The test method is the same as in super-wideband (see clause 9.10.3a, observing the test signal properties for fullband described in clause 5.4).

### 10.10.3b Delay in sending + receiving direction using "echo" method (electrical interface UE)

The test method is the same as in super-wideband (see clause 9.10.3b, observing the test signal properties for fullband in clause 5.4).

### 10.10.4 Delay and speech quality in conditions with packet arrival time variations and packet loss (handset, headset, electrical interface UE)

#### 10.10.4.1 Delay in sending direction

The test method is the same as in super-wideband (see clause 9.10.4.1).

#### 10.10.4.2 Delay in receiving direction

The test method is the same as in super-wideband (see clause 9.10.4.2, observing the test signal properties for fullband described in clause 5.4).

#### 10.10.4.3 Speech quality loss in conditions with packet arrival time variations and packet loss

For further study.

NOTE: Version 2.4 of Recommendation ITU-T P.863 [44] referenced in the present document was developed and validated for applications up to super-wideband bandwidth. Version 3.0 (or later) provides support for several fullband applications and may be used in this clause.

------------------------- END OF CHANGE 34 -------------------------

------------------------- START OF CHANGE 35 -------------------------

## 7.11 Echo control characteristics

### 7.11.1 Test set-up and test signals

The device is set up according to clause 5. The ambient noise level shall be ≤ ‑64 dBPa(A).

The test shall be performed with the British-English "long" double-talk and conditioning speech sequences from ITU-T Recommendation P.501 [22], with the signals in the receiving direction band limited according to clause 5.4.

A description of the test stimuli is presented in Table 2a and Table 2b. The test sequence is composed of an initial conditioning sequence of 23,5 s and a double talk sequence of 35 s. For the analysis, the double talk sequence is divided into two segments, a first double-talk sequence with single short near-end words (0 – 20 s), and a second double-talk sequence with continuous double talk (20 – 35 s).

The sending speech during double-talk and the "near-end speech only" are recorded individually, with the "near-end speech only" sequence recorded with silence in the receiving direction. The time-alignment of the two recorded sequences is performed off-line during the analysis.

Table 2a: Test stimuli for recording of Echo Canceller operation

|  |  |  |
| --- | --- | --- |
|  | Conditioning | Single words (segment 1) and full sentence (segment 2) double talk |
| Far-end signal | FB\_female\_conditioning\_seq\_long.wav | FB\_male\_female\_single-talk\_seq.wav |
| Artificial mouth signal | FB\_male\_conditioning\_seq\_long.wav | FB\_male\_female\_double-talk\_seq.wav |

Table 2b: Test stimuli for reference "near-end speech only" recording.

|  |  |  |
| --- | --- | --- |
|  | Conditioning | Single words (segment 1) and full sentence (segment 2) double talk |
| Far-end signal | FB\_female\_conditioning\_seq\_long.wav | silence |
| Artificial mouth signal | FB\_male\_conditioning\_seq\_long.wav | FB\_male\_female\_double-talk\_seq.wav |

The level of the signal of the artificial mouth shall be -4,7 dBPa measured at the MRP. For electrical interface UE, the level of the signal shall be calibrated to -60 dBV for analogue and to -16 dBm0 for digital connections. In order to obtain a reproducible time alignment as seen by the UE, the send signal (artificial mouth, electrical reference interface output) shall be delayed by the amount of the receiving direction delay. For the purpose of this alignment, the receiving direction delay for handset and headset modes is defined from the system simulator input to the artificial ear or the electrical reference interface, respectively. For hands-free modes, the downlink delay is defined from the system simulator input to the acoustic output from the UE loudspeaker.

The level of the downlink signal shall be -16 dBm0 measured at the digital reference point or the equivalent analogue point.

For electrical interface UE, an echo loss of 30 dB as described in clause 5.1.6 shall be simulated in the electrical reference interface.

[...]

------------------------- END OF CHANGE 35 -------------------------

------------------------- START OF CHANGE 36 -------------------------

## 8.11 Echo control characteristics

### 8.11.1 Test set-up and test signals

The device is set up according to clause 5. The ambient noise level shall be ≤ ‑64 dBPa(A).

The test shall be performed with the British-English "long" double-talk and conditioning speech sequences from ITU-T Recommendation P.501 [22], with the signals in the receiving direction band limited according to clause 5.4.

A description of the test stimuli is presented in Table 2e and Table 2f. The test sequence is composed of an initial conditioning sequence of 23,5 s and a double talk sequence of 35 s. For the analysis, the double talk sequence is divided into two segments, a first double-talk sequence with single short near-end words (0 – 20 s), and a second double-talk sequence with continuous double talk (20-35 s).

The sending speech during double-talk and the "near-end speech only" are recorded individually, with the "near-end speech only" sequence recorded with silence in the receiving direction. The time-alignment of the two recorded sequences is performed off-line during the analysis.

Table 2e: Test stimuli for recording of Echo Canceller operation

|  |  |  |
| --- | --- | --- |
|  | Conditioning | Single words (segment 1) and full sentence (segment 2) double talk |
| Far-end signal | FB\_female\_conditioning\_seq\_long.wav | FB\_male\_female\_single-talk\_seq.wav |
| Artificial mouth signal | FB\_male\_conditioning\_seq\_long.wav | FB\_male\_female\_double-talk\_seq.wav |

Table 2f: Test stimuli for reference "near-end speech only" recording.

|  |  |  |
| --- | --- | --- |
|  | Conditioning | Single words (segment 1) and full sentence (segment 2) double talk |
| Far-end signal | FB\_female\_conditioning\_seq\_long.wav | silence |
| Artificial mouth signal | FB\_male\_conditioning\_seq\_long.wav | FB\_male\_female\_double-talk\_seq.wav |

The level of the signal of the artificial mouth shall be - 4.7 dBPa measured at the MRP. For electrical interface UE, the level of the signal shall be calibrated to -60 dBV for analogue and to -16 dBm0 for digital connections. In order to obtain a reproducible time alignment as seen by the UE, the send signal (artificial mouth, electrical reference interface output) shall be delayed by the amount of the receiving direction delay. For the purpose of this alignment, the receiving direction delay for handset and headset modes is defined from the system simulator input to the artificial ear or the electrical reference interface, respectively. For handsfree modes, the downlink delay is defined from the system simulator input to the acoustic output from the UE loudspeaker.

The level of the downlink signal shall be -16 dBm0 measured at the digital reference point or the equivalent analogue point.

For electrical interface UE, an echo loss of 30 dB as described in clause 5.1.6 shall be simulated in the electrical reference interface.

[...]

------------------------- END OF CHANGE 36 -------------------------

------------------------- START OF CHANGE 37 -------------------------

## 7.12 Send speech quality and noise intrusiveness in the presence of ambient noise

### 7.12.1 Handset UE

[...]

### 7.12.2 Hand-held hands-free UE

[...]

### 7.12.3 Electrical interface UE

The speech quality in sending for narrowband systems is tested based on ETSI TS 103 106 [34]. This test method leads to three MOS-LQOn quality numbers:

 N-MOS-LQOn: Transmission quality of the background noise

 S-MOS-LQOn: Transmission quality of the speech

 G-MOS-LQOn: Overall transmission quality

For the measurement of electrial interface UE, pre-recorded noisy speech signals according to Annex B of Recommendation ITU‑T P.381 [53] shall be used. These noisy test sequences are available for the eight noise types described in Table 2d and were captured at the electrical output of a representative analogue headset. The corresponding speech level at MRP was calibrated to -1.7 dBPa, as described in clause 7.12.1. All test signals also include the proper conditioning sequence described in ETSI TS 103 106 [34], which is applied to the beginning of the 16-sentence test sequence.

Annex B of Recommendation ITU‑T P.381 [53] also provides the corresponding unprocessed reference speech signals, which are necessary for the calculation of S-MOS, N-MOS and G-MOS according to [b-ETSI TS 103 106]. These signals were recorded with a omnidirectional measurement microphone close to the input microphone of the representative headset.

1) The test arrangement is given in clause 5.1.6. For analogue interfaces, the noisy test sequences according to Annex B of Recommendation ITU‑T P.381 [53] shall be calibrated in a way that -26 dBov correspond to ‑60 dBV. For digital interfaces, -26 dBov shall correspond to -16 dBm0.

2) The noisy test sequence is inserted into electrical interface UE and then recorded at the POI.

3) N-MOS-LQOn, S-MOS-LQOn and G-MOS-LQOn are calculated as described in ETSI TS 103 106 [34] (narrowband mode) on a per sentence basis and averaged over all 16 sentences. The results shall be reported as average and standard deviation. Three signals are required for the tests:

– The clean speech signal is used as the undisturbed reference (see ETSI TS 103 106 [34], ETSI EG 202 396‑3 [36]).

– The speech plus undisturbed background noise signal. For each noisy test signal, a corresponding signal is available in Annex B of Recommendation ITU‑T P.381 [53] as well.

– The send signal is recorded at the POI.

4) The measurement is repeated for each ambient noise condition described in Table 2d. For each of these noise types, a corresponding test signal is available in Annex B of Recommendation ITU‑T P.381 [53].

5) The average of the results derived from all ambient noise types is calculated.

------------------------- END OF CHANGE 37 -------------------------

------------------------- START OF CHANGE 38 -------------------------

## 8.12 Send speech quality and noise intrusiveness in the presence of ambient noise

### 8.12.1 Handset UE

[...]

### 8.12.2 Hand-held hands-free UE

[...]

### 8.12.3 Electrical interface UE

The speech quality in sending for narrowband systems is tested based on ETSI TS 103 106 [34]. This test method leads to three MOS-LQOw quality numbers:

 N-MOS-LQOw: Transmission quality of the background noise

 S-MOS-LQOw: Transmission quality of the speech

 G-MOS-LQOw: Overall transmission quality

For the measurement of electrial interface UE, pre-recorded noisy speech signals according to Annex B of Recommendation ITU‑T P.381 [53] shall be used. These noisy test sequences are available for the eight noise types described in Table 2h and were captured at the electrical output of a representative analogue headset. The corresponding speech level at MRP was calibrated to -1.7 dBPa, as described in clause 8.12.1. All test signals also include the proper conditioning sequence described in ETSI TS 103 106 [34], which is applied to the beginning of the 16-sentence test sequence.

Annex B of Recommendation ITU‑T P.381 [53] also provides the corresponding unprocessed reference speech signals, which are necessary for the calculation of S-MOS, N-MOS and G-MOS according to [b-ETSI TS 103 106]. These signals were recorded with a omnidirectional measurement microphone close to the input microphone of the representative headset.

1) The test arrangement is given in clause 5.1.6. For analogue interfaces, the noisy test sequences according to Annex B of Recommendation ITU‑T P.381 [53] shall be calibrated in a way that -26 dBov correspond to ‑60 dBV. For digital interfaces, -26 dBov shall correspond to -16 dBm0.

2) The noisy test sequence is inserted into electrical interface UE and then recorded at the POI.

3) N-MOS-LQOw, S-MOS-LQOw and G-MOS-LQOw are calculated as described in ETSI TS 103 106 [34] (wideband mode) on a per sentence basis and averaged over all 16 sentences. The results shall be reported as average and standard deviation. Three signals are required for the tests:

– The clean speech signal is used as the undisturbed reference (see ETSI TS 103 106 [34], ETSI EG 202 396‑3  [36]).

– The speech plus undisturbed background noise signal. For each noisy test signal, a corresponding signal is available in Annex B of Recommendation ITU‑T P.381 [53] as well.

– The send signal is recorded at the POI.

4) The measurement is repeated for each ambient noise condition described in Table 2h. For each of these noise types, a corresponding test signal is available in Annex B of Recommendation ITU‑T P.381 [53].

5) The average of the results derived from all ambient noise types is calculated.

------------------------- END OF CHANGE 38 -------------------------

------------------------- START OF CHANGE 39 -------------------------

## 9.12 Send speech quality and noise intrusiveness in the presence of ambient noise

### 9.12.1 Handset UE

[...]

### 9.12.2 Hand-held hands-free UE

[...]

### 9.12.3 Electrical interface UE

The speech quality in sending for super-wideband systems is tested based on ETSI TS 103 281 [50]. This test method leads to three MOS-LQOfb quality numbers:

- N-MOS-LQOfb: Transmission quality of the background noise

- S-MOS-LQOfb: Transmission quality of the speech

- G-MOS-LQOfb: Overall transmission quality

For the measurement of electrial interface UE, pre-recorded noisy speech signals according to Annex B of Recommendation ITU‑T P.381 [53] shall be used. These noisy test sequences are available for the eight noise types described in Table 2i and were captured at the electrical output of a representative analogue headset. The corresponding speech level at MRP was calibrated to -1.7 dBPa, as described in clause 9.12.1. All test signals also include the proper conditioning sequence described in ETSI TS 103 281 [50], which is applied to the beginning of the 16-sentence test sequence.

Annex B of Recommendation ITU‑T P.381 [53] also provides a recording without ambient noise and without Lombard correction (-4.7 dBPa at MRP). This silence condition is needed for the calibration procedure described in clause 9.5 of ETSI TS 103 281 [50].

1) The test arrangement is given in clause 5.1.6. For analogue interfaces, the noisy test sequences according to Annex B of Recommendation ITU‑T P.381 [53] shall be calibrated in a way that -26 dBov correspond to ‑60 dBV. For digital interfaces, -26 dBov shall correspond to -16 dBm0.

2) Before starting the measurements, the calibration procedure described in clause 9.5 of ETSI TS 103 281 [50] shall be performed with the electrical interface UE. A recording in silence as per Annex B of Recommendation ITU‑T P.381 [53] shall be used for the measurement.

3) The first noisy test sequence is inserted into electrical interface UE and then recorded at the POI. Two signals are required for the prediction model:

- The clean speech signal is used as the undisturbed reference (see ETSI TS 103 281 [50])

- The send signal is recorded at the POI.

4) N-MOS-LQOfb, S-MOS-LQOfb and G-MOS-LQOfb are calculated according to the Model A objective predictor described in ETSI TS 103 281[50] on a per sentence basis and averaged over all 16 sentences. The final results are derived as follows:

- S-MOS-LQOfb = S-MOS-LQOfb\_modelA

- N-MOS-LQOfb = 1.438\*N-MOS-LQOfb\_modelA – 1.959

- G-MOS-LQOfb = G-MOS-LQOfb\_modelA

5) The measurement is repeated for each ambient noise condition described in Table 2i. For each of these noise types, a corresponding test signal is available in Annex B of Recommendation ITU‑T P.381 [53].

6) The average of the results derived from all ambient noise types is calculated.

------------------------- END OF CHANGE 39 -------------------------

------------------------- START OF CHANGE 40 -------------------------

## 10.12 Send speech quality and noise intrusiveness in the presence of ambient noise

### 10.12.1 Handset UE

The test method is the same as in super-wideband (see sub-clause 9.12.1).

### 10.12.2 Hand-held hands-free UE

The test method is the same as in super-wideband (see sub-clause 9.12.2).

### 10.12.3 Electrical interface UE

The test method is the same as in super-wideband (see sub-clause 9.12.3).

------------------------- END OF CHANGE 40 -------------------------

------------------------- START OF CHANGE 41 -------------------------

## 7.13 Jitter buffer management behaviour (handset, headset and electrical interface UE)

### 7.13.0 General

[...]

### 7.13.1 Delay histogram

For this test it shall be ensured that the call is originated from the UE.

NOTE 1: Differences have been observed between UE-originated calls and UE-terminated calls. For better consistency, calls from the UE are used.

The test signal consists of 3 repeats of the Composite Source Signal (CSS) according to ITU-T Recommendation P.501 [22] followed by a speech signal of 160s. During the first two CSS signals the terminal can adapt its jitter buffer. The third CSS is used for measuring the delay in constant-delay condition, and the speech signal is used for delay and quality measurement in the packet impairment condition.

Constant delay Tc corresponding to the minimum delay of the profile (i.e. the compensation value for the profile) shall be added at the beginning of the different delay/loss profiles, to avoid unecessary delay jumps between the two measurement phases and realistic conditions for the second measurement test phase.

In receiving direction, the delay between the electrical access point of the test equipment and the reference point (RP), TTEAP-RP(t) = TR-jitter(t)+ TTER, is measured in two successive phases:

1) First the delay in constant-delay condition TTEAP-DRP-constant is measured as described in steps 1 to 4, clause 7.10.2/7.10.2a/7.10.2b, using the third CSS signal. The constant delay Tc is subtracted from TTEAP-RP to obtain TR-constant.

2) Then the delay with packet impairment TR-jitter(t) is measured continuously for a speech signal during the inclusion of packet delay and loss profiles in the receiving direction RTP voice stream.

The reference point is defined as follows:

- for handset and headset UE, the reference point is the DRP.

- for electrical interface UE, the reference point is the input of the electrical reference interface.

Packet impairments shall be applied between the reference client and system simulator eNodeB. Separate calls shall be established for each packet impairment condition.

The start of the delay profiles must be synchronized with the start of the downlink speech material reproduction (compensated by the delay between reproduction and the point of impairment insertion, i.e. the delay of the reference client) in order to ensure a repeatable application of impairments to the test speech signal. Tests shall be performed with DTX enabled in the reference client.

NOTE 2: RTP packet impairments representing packet delay variations and loss are specified in Annex F. Care must be taken that the system simulator uses a dedicated bearer with no buffering/scheduling of packets for transmission.

For the CSS signal repeated 3 times, the pseudo random noise (pn)-part of the CSS has to be longer than the maximum expected delay. It is recommended to use a pn sequence of 32 k samples (with 48 kHz sampling rate). The test signal level is -16 dBm0 measured at the digital reference point or the equivalent analogue point.

For the speech signal, 8 English test sentences according to ITU-T P.501 Annex C.2.3, normalized to an active speech level of -16dBm0, are used (2 male, 2 female speakers). The sequences are concatenated in such a way that all sentences are centered within a 4.0s time window, which results in an overall duration of 32.0s. The sequences are repeated 5 times, resulting in a test file 160.0s long. The first 2 sentences are used for convergence of the UE jitter buffer manager and are discarded from the analysis. Equivalent implementations of the concatenation by repeating the test sentences in sequence may be used.

For the delay calculation with the speech signal, a cross-correlation with a rectangular window length of 4s, centered at each sentence of the stimulus file, is used. The process is repeated for each sample. For each cross correlation, the maximum of the envelope is obtained producing one delay value per sentence.

The UE delay in the receive direction, TR-jitter(t), is obtained by subtracting the delay introduced by the test equipment and the simulated transport network packet delay introduced by the delay and loss profile (as specified for the respective profile in Annex F) from the first electrical event at the electrical access point of the test equipment to the first bit of the corresponding speech frame at the system simulator antenna, TTER, from the measured TTEAP-RP(t).

The difference DT between maximum receiving delay obtained with at least 5 individual calls (see clause 7.10.2) and the delay TR-constant measured for the CSS signal in constant delay condition is calculated. The quantity "Call-to-Call Variability Adjustment" (CCVA) = max(0,DT) shall be added to the obtained delay for the speech signal TR-jitter(t).

The UE delay in the receiving direction shall be reported in the form of an histogram covering the range of measured CCVA-adjusted values (TR-CCVA(t) = TR-jitter(t) + CCVA) with a step of 20 ms. The following pseudo code provides an example implementation for the histogram:

lo=min(floor(TR-CCVA(t=1...40)/20)\*20)

hi=max(ceil(TR-CCVA(t=1...40)/20)\*20)

[n,x]=hist(TR-CCVA(t=1...40),lo:20:hi)

bar(x,n)

The TR-CCVA values for all 40 sentences shall also be reported in the test report.

NOTE 3: The synchronization of the speech frame processing in the UE to the bits of the speech frames at the UE antenna may lead to a variability of up to 20 ms of the measured UE receive delay between different calls. This synchronization is attributed to the UE receiving delay according to the definition of the UE delay reference points. The effect of this possible call-to-call variation is taken into account with the CCVA = max(0,DT) value.

### 7.13.2 Speech quality loss histogram

For the evaluation of speech quality loss in conditions with packet arrival time variations and packet loss, the speech test signal described in clause 7.13.1 shall be used. Two 48 kHz recordings are used to produce the speech quality loss metric:

- A recording obtained in jitter and error free conditions with the test signal described in clause 7.13.1 (reference condition)

- A recording obtained during the application of packet arrival time variations and packet loss as described in clause 7.13.1 (test condition)

The speech quality of the signal is estimated using the measurement algorithm described in ITU-T Recommendation P.863 [44] in super-wideband mode. For narrowband speech, the method according to Appendix III of P.863 [44] shall be used. Level pre-alignment to -26 dBov of recordings shall be used – see P.863.1 clause 10.2 [45].

NOTE: For the analysis of acoustical measurements, ITU-T P.863 [44] assumes diffuse-field equalized recordings. For this reason, signals at DRP are diffuse-field corrected for testing handset and headset UE. For electrical interface UE, only the level pre-alignment is applied.

A score shall be computed for each 8s speech sentence pair. The MOS-LQO values for the reference and test conditions shall be reported in the form of an histogram covering the range of measured values with a step of 0.1 and the values for all 20 sentences pairs shall also be reported in the test report. The following pseudo code provides an example implementation for the histogram:

lo=min(floor(MOS-LQO*test condition*(i=1...20)/0.1)\*0.1)

hi=max(ceil(MOS-LQO*test condition*(i=1...20)/0.1)\*0.1)

[n,x]=hist(MOS-LQO*test condition*(i=1...20),lo:0.1:hi)

bar(x,n)

The synchronization between stimuli and degraded condition shall be done by the test system before applying the P.863 algorithm on each sentence pair.

------------------------- END OF CHANGE 41 -------------------------

------------------------- START OF CHANGE 42 -------------------------

## 8.13 Jitter buffer management behaviour (handset, headset and electrical interface UE)

### 8.13.0 General

[...]

### [8.13.1 Delay histogram]

For this test it shall be ensured that the call is originated from the UE.

NOTE 1: Differences have been observed between UE-originated calls and UE-terminated calls. For better consistency, calls from the UE are used.

The test signal consists of 3 repeats of the Composite Source Signal (CSS) according to ITU-T Recommendation P.501 [22] followed by a speech signal of 160s. During the first two CSS signals the terminal can adapt its jitter buffer. The third CSS is used for measuring the delay in constant-delay condition, and the speech signal is used for delay and quality measurement in the packet impairment condition.

Constant delay Tc corresponding to the minimum delay of the profile (i.e. the compensation value for the profile) shall be added at the beginning of the different delay/loss profiles, to avoid unecessary delay jumps between the two measurement phases and realistic conditions for the second measurement test phase.

In receiving direction, the delay between the electrical access point of the test equipment and the reference point (RP), TTEAP-RP(t) = TR-jitter(t)+ TTER, is measured in two successive phases:

1) First the delay in constant-delay condition TTEAP-DRP-constant is measured as described in steps 1 to 4, clause 8.10.2, 8.10.2a/8.10.2b using the third CSS signal. The constant delay Tc is subtracted from TTEAP-RP to obtain TR-constant.

2) Then the delay with packet impairment TR-jitter(t) is measured continuously for a speech signal during the inclusion of packet delay and loss profiles in the receiving direction RTP voice stream.

The reference point is defined as follows:

- for handset and headset UE, the reference point is the DRP.

- for electrical interface UE, the reference point is the input of the electrical reference interface.

Packet impairments shall be applied between the reference client and system simulator eNodeB. Separate calls shall be established for each packet impairment condition.

The start of the delay profiles must be synchronized with the start of the downlink speech material reproduction (compensated by the delay between reproduction and the point of impairment insertion, i.e. the delay of the reference client) in order to ensure a repeatable application of impairments to the test speech signal. Tests shall be performed with DTX enabled in the reference client.

NOTE 2: RTP packet impairments representing packet delay variations and loss are specified in Annex F. Care must be taken that the system simulator uses a dedicated bearer with no buffering/scheduling of packets for transmission.

For the CSS signal repeated 3 times, the pseudo random noise (pn)-part of the CSS has to be longer than the maximum expected delay. It is recommended to use a pn sequence of 32 k samples (with 48 kHz sampling rate). The test signal level is -16 dBm0 measured at the digital reference point or the equivalent analogue point.

For the speech signal, 8 English test sentences according to ITU-T P.501 Annex C.2.3, normalized to an active speech level of -16dBm0, are used (2 male, 2 female speakers). The sequences are concatenated in such a way that all sentences are centered within a 4.0s time window, which results in an overall duration of 32.0s. The sequences are repeated 5 times, resulting in a test file 160.0s long. The first 2 sentences are used for convergence of the UE jitter buffer manager and are discarded from the analysis. Equivalent implementations of the concatenation by repeating the test sentences in sequence may be used.

For the delay calculation with the speech signal, a cross-correlation with a rectangular window length of 4s, centered at each sentence of the stimulus file, is used. The process is repeated for each sample. For each cross correlation, the maximum of the envelope is obtained producing one delay value per sentence.

The UE delay in the receive direction, TR-jitter(t), is obtained by subtracting the delay introduced by the test equipment and the simulated transport network packet delay introduced by the delay and loss profile (as specified for the respective profile in Annex F) from the first electrical event at the electrical access point of the test equipment to the first bit of the corresponding speech frame at the system simulator antenna, TTER, from the measured TTEAP-RP(t).

The difference DT between maximum receiving delay obtained with at least 5 individual calls (see clause 8.10.2) and the delay TR-constant measured for the CSS signal in constant delay condition is calculated. The quantity "Call-to-Call Variability Adjustment" (CCVA) = max(0,DT) shall be added to the obtained delay for the speech signal TR-jitter(t).

The UE delay in the receiving direction shall be reported in the form of an histogram covering the range of measured CCVA-adjusted values (TR-CCVA(t) = TR-jitter(t) + CCVA) with a step of 20 ms. The following pseudo code provides an example implementation for the histogram:

lo=min(floor(TR-CCVA(t=1...40)/20)\*20)

hi=max(ceil(TR-CCVA(t=1...40)/20)\*20)

[n,x]=hist(TR-CCVA(t=1...40),lo:20:hi)

bar(x,n)

The TR-CCVA values for all 40 sentences shall also be reported in the test report.

NOTE 3: The synchronization of the speech frame processing in the UE to the bits of the speech frames at the UE antenna may lead to a variability of up to 20 ms of the measured UE receive delay between different calls. This synchronization is attributed to the UE receiving delay according to the definition of the UE delay reference points. The effect of this possible call-to-call variation is taken into account with the CCVA = max(0,DT) value.

### 8.13.2 Speech quality loss histogram

For the evaluation of speech quality loss in conditions with packet arrival time variations and packet loss, the speech test signal described in clause 8.13.1 shall be used. Two 48 kHz recordings are used to produce the speech quality loss metric:

- A recording obtained in jitter and error free conditions with the test signal described in clause 8.13.1 (reference condition)

- A recording obtained during the application of packet arrival time variations and packet loss as described in clause 8.13.1 (test condition)

The speech quality of the signal is estimated using the measurement algorithm described in ITU-T Recommendation P.863 [44] in super-wideband mode. Level pre-alignment to -26 dBov of recordings shall be used – see P.863.1 clause 10.2 [45].

NOTE: For the analysis of acoustical measurements, ITU-T P.863 [44] assumes diffuse-field equalized recordings. For this reason, signals at DRP are diffuse-field corrected for testing handset and headset UE. For electrical interface UE, only the level pre-alignment is applied.

A score shall be computed for each 8s speech sentence pair. The MOS-LQO values for the reference and test conditions shall be reported in the form of an histogram covering the range of measured values with a step of 0.1 and the values for all 20 sentences pairs shall also be reported in the test report. The following pseudo code provides an example implementation for the histogram:

lo=min(floor(MOS-LQO*test condition*(i=1...20)/0.1)\*0.1)

hi=max(ceil(MOS-LQO*test condition*(i=1...20)/0.1)\*0.1)

[n,x]=hist(MOS-LQO*test condition*(i=1...20),lo:0.1:hi)

bar(x,n)

The synchronization between stimuli and degraded condition shall be done by the test system before applying the P.863 algorithm on each sentence pair.

------------------------- END OF CHANGE 42 -------------------------

------------------------- START OF CHANGE 43 -------------------------

## 9.13 Jitter buffer management behaviour (handset, headset and electrical interface UE)

### 9.13.0 General

[...]

### 9.13.1 Delay histogram

For this test it shall be ensured that the call is originated from the UE.

NOTE 1: Differences have been observed between UE-originated calls and UE-terminated calls. For better consistency, calls from the UE are used.

The test signal consists of 3 repeats of the Composite Source Signal (CSS) according to ITU-T Recommendation P.501 [22] followed by a speech signal of 160s. During the first two CSS signals the terminal can adapt its jitter buffer. The third CSS is used for measuring the delay in constant-delay condition, and the speech signal is used for delay and quality measurement in the packet impairment condition.

Constant delay Tc corresponding to the minimum delay of the profile (i.e. the compensation value for the profile) shall be added at the beginning of the different delay/loss profiles, to avoid unnecessary delay jumps between the two measurement phases and realistic conditions for the second measurement test phase.

In receiving direction, the delay between the electrical access point of the test equipment and the reference point (RP), TTEAP-RP(t) = TR-jitter(t)+ TTER, is measured in two successive phases:

1) First the delay in constant-delay condition TTEAP-RP-constant is measured as described in steps 1 to 4, clause 9.10.2/9.10.2a/9.10.2b, using the third CSS signal. The constant delay Tc is subtracted from TTEAP-RP to obtain TR-constant.

2) Then the delay with packet impairment TR-jitter(t) is measured continuously for a speech signal during the inclusion of packet delay and loss profiles in the receiving direction RTP voice stream.

The reference point is defined as follows:

- for handset and headset UE, the reference point is the DRP.

- for electrical interface UE, the reference point is the input of the electrical reference interface.

Packet impairments shall be applied between the reference client and system simulator eNodeB. Separate calls shall be established for each packet impairment condition.

The start of the delay profiles must be synchronized with the start of the downlink speech material reproduction (compensated by the delay between reproduction and the point of impairment insertion, i.e. the delay of the reference client) in order to ensure a repeatable application of impairments to the test speech signal. Tests shall be performed with DTX enabled in the reference client.

NOTE 2: RTP packet impairments representing packet delay variations and loss are specified in Annex F. Care must be taken that the system simulator uses a dedicated bearer with no buffering/scheduling of packets for transmission.

For the CSS signal repeated 3 times, the pseudo random noise (pn)-part of the CSS has to be longer than the maximum expected delay. It is recommended to use a pn sequence of 32 k samples (with 48 kHz sampling rate). The test signal level is -16 dBm0 measured at the digital reference point or the equivalent analogue point.

For the speech signal, 8 English test sentences according to ITU-T P.501 Annex C.2.3, normalized to an active speech level of -16dBm0, are used (2 male, 2 female speakers). The sequences are concatenated in such a way that all sentences are centred within a 4.0s time window, which results in an overall duration of 32.0s. The sequences are repeated 5 times, resulting in a test file 160.0s long. The first 2 sentences are used for convergence of the UE jitter buffer manager and are discarded from the analysis. Equivalent implementations of the concatenation by repeating the test sentences in sequence may be used.

For the delay calculation with the speech signal, a cross-correlation with a rectangular window length of 4s, centered at each sentence of the stimulus file, is used. The process is repeated for each sample. For each cross correlation, the maximum of the envelope is obtained producing one delay value per sentence.

The UE delay in the receive direction, TR-jitter(t), is obtained by subtracting the delay introduced by the test equipment and the simulated transport network packet delay introduced by the delay and loss profile (as specified for the respective profile in Annex F) from the first electrical event at the electrical access point of the test equipment to the first bit of the corresponding speech frame at the system simulator antenna, TTER, from the measured TTEAP-RP(t).

The difference DT between maximum receiving delay obtained with at least 5 individual calls (see clause 9.10.2) and the delay TR-constant measured for the CSS signal in constant delay condition is calculated. The quantity "Call-to-Call Variability Adjustment" (CCVA) = max(0,DT) shall be added to the obtained delay for the speech signal TR-jitter(t).

The UE delay in the receiving direction shall be reported in the form of an histogram covering the range of measured CCVA-adjusted values (TR-CCVA(t) = TR-jitter(t) + CCVA) with a step of 20 ms. The following pseudo code provides an example implementation for the histogram:

lo=min(floor(TR-CCVA(t=1...40)/20)\*20)

hi=max(ceil(TR-CCVA(t=1...40)/20)\*20)

[n,x]=hist(TR-CCVA(t=1...40),lo:20:hi)

bar(x,n)

The TR-CCVA values for all 40 sentences shall also be reported in the test report.

NOTE 3: The synchronization of the speech frame processing in the UE to the bits of the speech frames at the UE antenna may lead to a variability of up to 20 ms of the measured UE receive delay between different calls. This synchronization is attributed to the UE receiving delay according to the definition of the UE delay reference points. The effect of this possible call-to-call variation is taken into account with the CCVA = max(0,DT) value.

### 9.13.2 Speech quality loss histogram

For the evaluation of speech quality loss in conditions with packet arrival time variations and packet loss, the speech test signal described in clause 9.13.1 shall be used. Two 48 kHz recordings are used to produce the speech quality loss metric:

- A recording obtained in jitter and error free conditions with the test signal described in clause 9.13.1 (reference condition)

- A recording obtained during the application of packet arrival time variations and packet loss as described in clause 9.13.1 (test condition)

The speech quality of the signal is estimated using the measurement algorithm described in ITU-T Recommendation P.863 [44] in super-wideband mode. Level pre-alignment to -26 dBov of recordings shall be used – see P.863.1 clause 10.2 [45].

NOTE: For the analysis of acoustical measurements, ITU-T P.863 [44] assumes diffuse-field equalized recordings. For this reason, signals at DRP are diffuse-field corrected for testing handset and headset UE. For electrical interface UE, only the level pre-alignment is applied.

A score shall be computed for each 8s speech sentence pair. The MOS-LQO values for the reference and test conditions shall be reported in the form of an histogram covering the range of measured values with a step of 0.1 and the values for all 20 sentences pairs shall also be reported in the test report. The following pseudo code provides an example implementation for the histogram:

lo=min(floor(MOS-LQO*test condition*(i=1...20)/0.1)\*0.1)

hi=max(ceil(MOS-LQO*test condition*(i=1...20)/0.1)\*0.1)

[n,x]=hist(MOS-LQO*test condition*(i=1...20),lo:0.1:hi)

bar(x,n)

The synchronization between stimuli and degraded condition shall be done by the test system before applying the P.863 algorithm on each sentence pair.

------------------------- END OF CHANGE 43 -------------------------

------------------------- START OF CHANGE 44 -------------------------

## 10.13 Jitter buffer management behaviour (handset, headset and electrical interface UE)

### 10.13.0 General

The same considerations as described in clause 9.13.0 apply for fullband mode. The test methods are the same as in super-wideband (see clause 9.13, observing the test signal properties for fullband described in clause 5.4).

### 10.13.1 Delay histogram

The test method are the same as in super-wideband (see clause 9.13.1, observing the test signal properties for fullband described in clause 5.4).

### 10.13.2 Speech quality loss histogram

For further study.

NOTE: Version 2.4 of Recommendation ITU-T P.863 [44] referenced in the present document was developed and validated for applications up to super-wideband bandwidth. Version 3.0 (or later) provides support for several fullband applications and may be used in this clause.

------------------------- END OF CHANGE 44 -------------------------

------------------------- START OF CHANGE 45 -------------------------

Annex D (normative):
Clock skew measurement

This Annex describes a method to measure the clock skew between the reference client and the device under test.

# D.1 Test procedure

As speech test signal, the second sentence of the first female speaker (female1.wav) of the English test sentences according to ITU-T P.501 is used. When measuring in receiving direction the signal is pre-filtered according to the used bandwith and normalized to an active speech level of -16dBm0. When measuring in sending direction the signal is calibrated to an active speech level of -4.7 dBPa at MRP. For electrical interface UE, the active speech level is calibrated to -60 dBV for analogue and to -16 dBm0 for digital connections. The sequence is centered within a 4.0s time window and repeated 40 times, resulting in a test file of 160.0s length. Alternatively CSS signals may be used with the corresponding adaptation of the step size to 333ms. The test signal used shall be reported.

[...]

------------------------- END OF CHANGE 45 -------------------------