

ETSI ES 202 718 V1.4.1 (2020-05)



**Speech and multimedia Transmission Quality (STQ);
Transmission Requirements for IP-based Narrowband and
Wideband Home and Network Media Gateways from
a QoS Perspective as Perceived by the User**

Reference

RES/STQ-278

Keywords

QoS, speech

ETSI

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

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

Siret N° 348 623 562 00017 - NAF 742 C
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° 7803/88

Important notice

The present document can be downloaded from:

<http://www.etsi.org/standards-search>

The present document may be made available in electronic versions and/or in print. The content of any electronic and/or print versions of the present document shall not be modified without the prior written authorization of ETSI. In case of any existing or perceived difference in contents between such versions and/or in print, the prevailing version of an ETSI deliverable is the one made publicly available in PDF format at www.etsi.org/deliver.

Users of the present document should be aware that the document may be subject to revision or change of status.

Information on the current status of this and other ETSI documents is available at

<https://portal.etsi.org/TB/ETSIDeliverableStatus.aspx>

If you find errors in the present document, please send your comment to one of the following services:

<https://portal.etsi.org/People/CommiteeSupportStaff.aspx>

Copyright Notification

No part may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm except as authorized by written permission of ETSI.

The content of the PDF version shall not be modified without the written authorization of ETSI.

The copyright and the foregoing restriction extend to reproduction in all media.

© ETSI 2020.

All rights reserved.

DECT™, **PLUGTESTS™**, **UMTS™** and the ETSI logo are trademarks of ETSI registered for the benefit of its Members.

3GPP™ and **LTE™** are trademarks of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners.

oneM2M™ logo is a trademark of ETSI registered for the benefit of its Members and of the oneM2M Partners.

GSM® and the GSM logo are trademarks registered and owned by the GSM Association.

Contents

Intellectual Property Rights	5
Foreword.....	5
Modal verbs terminology.....	5
Introduction	5
1 Scope	6
2 References	6
2.1 Normative references	6
2.2 Informative references.....	8
3 Definition of terms, symbols and abbreviations.....	8
3.1 Terms.....	8
3.2 Symbols.....	9
3.3 Abbreviations	9
4 General considerations	10
4.1 Default Coding Algorithm.....	10
4.2 End-to-end considerations	11
4.3 Parameters to be investigated	11
4.3.1 Applicability of parameters to different MGWs	11
5 Test equipment	13
5.1 IP half channel measurement adaptor.....	13
5.2 Environmental conditions for tests	13
5.3 Accuracy of measurements and test signal generation	13
5.4 Network impairment simulation.....	13
6 Requirements and associated Measurement Methodologies	14
6.0 General	14
6.1 Test setup.....	14
6.1.0 General.....	14
6.1.1 Setup for Media Gateways with 4-wire interface	15
6.1.2 Setup for Media Gateways with 2-wire interface	16
6.1.3 Setup for Media Gateways with Wireless Access.....	16
6.1.4 Setup for IP-to-IP Media Gateways	17
6.1.5 Test Signal Levels.....	17
6.1.6 Background noise simulation.....	17
6.1.7 Echo paths.....	18
6.2 Coding independent parameters	19
6.2.1 Send Frequency response.....	19
6.2.2 Circuit Loudness Rating in Send	20
6.2.3 Linearity Range for CLR (SND).....	20
6.2.4 Send Distortion	21
6.2.5 Spurious Out-of-Band Signals in Send direction	22
6.2.6 Send Noise	22
6.2.7 Receive Frequency Response	23
6.2.8 Circuit Loudness Rating in Receive	24
6.2.9 Linearity Range for CLR (RCV)	24
6.2.10 Receive Distortion	26
6.2.11 Out-of-Band Signals Wideband to Narrowband Transcoding	26
6.2.12 Spurious Out-of-band Signals Narrowband to Wideband Transcoding	27
6.2.13 Minimum Activation Level and Sensitivity in Receive Direction	28
6.2.14 Receive Noise	28
6.2.15 Double Talk Performance.....	28
6.2.15.0 General	28
6.2.15.1 Attenuation Range in Send Direction during Double Talk $A_{H,S,dt}$	29
6.2.15.2 Attenuation Range in Receive Direction during Double Talk $A_{H,R,dt}$	30

6.2.15.3	Detection of Echo Components during Double Talk	31
6.2.15.4	Minimum Activation Level and Sensitivity of Double Talk Detection	33
6.2.16	Switching Characteristics	33
6.2.16.0	General	33
6.2.16.1	Activation in Send Direction	33
6.2.16.2	Activation in Receive Direction	34
6.2.16.3	Silence Suppression and Comfort Noise Generation	35
6.2.17	Background Noise Performance	35
6.2.17.0	General	35
6.2.17.1	Performance in Send Direction in the Presence of Background Noise	35
6.2.17.2	Quality of Speech with Background Noise	36
6.2.17.3	Quality of Background Noise Transmission (with Far End Speech)	37
6.2.17.4	Quality of Background Noise Transmission (with Near End Speech)	37
6.2.18	Quality of Echo Cancellation	38
6.2.18.0	General	38
6.2.18.1	Echo Performance According to Recommendation ITU-T G.168	38
6.2.18.2	Terminal Coupling Loss (TCLw, NB)	38
6.2.18.3	Terminal Coupling Loss (TCL, WB)	39
6.2.18.4	Temporal Echo Effects	39
6.2.18.5	Spectral Echo Attenuation	40
6.2.18.6	Occurrence of Artefacts	41
6.2.18.7	Variable Echo Paths	41
6.2.19	Variant Impairments	41
6.2.19.1	Clock Accuracy Send	41
6.2.19.2	Clock Accuracy Receive	41
6.2.19.3	Send Packet Delay Variation	42
6.2.20	Immunity to DTMF False Detection in Send Direction	42
6.2.21	Send and receive delay - round trip delay	43
6.3	Codec Specific Requirements	44
6.3.1	Objective Listening Speech Quality MOS-LQO in Send direction	44
6.3.2	Objective Listening Quality MOS-LQO in Receive direction	45
6.3.3	Quality of Jitter buffer adjustment	47
Annex A (informative):	Impulse Response of a Narrowband and Wideband DECT PP	49
Annex B (normative):	Test signal for immunity to DTMF false detection in send direction	51
Annex C (informative):	Example IP Delay variation for jitterbuffer quality measurements	52
Annex D (informative):	Bibliography	53
History		54

Intellectual Property Rights

Essential patents

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

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

Trademarks

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

Foreword

This ETSI Standard (ES) has been produced by ETSI Technical Committee Speech and multimedia Transmission Quality (STQ).

Modal verbs terminology

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

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

Introduction

Traditionally, analogue and digital telephones were interfacing switched-circuit 64 kbit/s PCM networks. With the fast growth of IP networks, packet-switched networks (VoIP) interfacing PSTN networks and mobile networks, as well as different types of IP-terminals, are being rapidly introduced. Different types of gateways are used to interconnect to such IP networks. Since the IP networks will be in many cases interworking with the traditional PSTN and private networks, many of the basic transmission requirements have to be harmonized between these different types of network from an end-to-end perspective, including specifications for the edge points.

The present document covers IP-based narrowband and wideband home and network media gateways. It aims to enhance the interoperability and end-to-end quality.

In contrast to other standards which define minimum performance requirements, it is the intention of the present document to specify gateway equipment requirements which enable manufacturers and service providers to enable end-to-end speech performance as perceived by the user. These requirements are absolutely necessary to ensure a good quality, but they are not sufficient. They have to be combined with requirements (and associated relevant measurement methods) for other elements in the transmission chain (core IP network, PSTN, terminals), as well as for the whole mouth-to-ear transmission path.

1 Scope

The present document provides speech transmission performance requirements for narrowband and wideband media gateways from a QoS perspective as perceived by the user. Media gateways can be network or home based, they may include a transcoding function. The present document covers the following types of IP-based media gateways:

- ATA (Analogue Terminal Adapter), home gateway IP to POTS
- ITA (ISDN Terminal Adapter), home gateway IP to ISDN
- IAD (Integrated Access device), home router including ATA or ITA
- Network based ATA and ITA
- Carrier grade media gateway, network gateway IP to TDM
- IP-to-IP media gateway, network gateway with transcoding and/or other media processing
- New Generation DECT Fixed part with IP interface (only parameters not covered by New Generation DECT)

Interfaces of media gateways used together with terminals as a system (i.e. connected via Ethernet or with a proprietary interface) are excluded in the present document and should be measured according to the relevant terminal standard.

If a media gateway includes more than one interface type (e.g. POTS and ISDN), each interface has to be dealt with differently.

The requirements available in the present document will ensure a high compatibility with IP- and TDM-based fixed and wireless terminals and networks, including DECT and mobile terminals.

It is the aim to optimize interoperability, the listening and talking quality and the conversational performance. Related requirements and test methods are defined in the present document.

The present document does not apply to media gateways with 4-wire analogue interfaces.

The requirements for MGWs with respect to voiceband data (VBD) are out of scope in the present document. These requirements are covered in ETSI TS 102 929 [i.4].

2 References

2.1 Normative references

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

Referenced documents which are not found to be publicly available in the expected location might be found at <https://docbox.etsi.org/Reference/>.

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

The following referenced documents are necessary for the application of the present document.

- [1] ETSI EN 300 726: "Digital cellular telecommunications system (Phase 2+) (GSM); Enhanced Full Rate (EFR) speech transcoding (GSM 06.60)".
- [2] ETSI TS 126 171 (V6.0.0): "Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); AMR speech codec, wideband; General description (3GPP TS 26.171 version 6.0.0 Release 6)".

- [3] Recommendation ITU-T G.107: "The E-model, a computational model for use in transmission planning".
- [4] Recommendation ITU-T G.108, including amendments 1 and 2: "Application of the E-model: A planning guide".
- [5] Recommendation ITU-T G.109: "Definition of categories of speech transmission quality".
- [6] Recommendation ITU-T G.100.1: "The use of the decibel and of relative levels in speechband telecommunications".
- [7] Recommendation ITU-T G.111: "Loudness Ratings (LRs) in an international connection".
- [8] Recommendation ITU-T G.122: "Influence of national systems on stability and talker echo in international connections".
- [9] Recommendation ITU-T G.711: "Pulse code modulation (PCM) of voice frequencies".
- [10] Recommendation ITU-T G.723.1: "Dual rate speech coder for multimedia communications transmitting at 5.3 and 6.3 kbit/s".
- [11] Recommendation ITU-T G.726: "40, 32, 24, 16 kbit/s Adaptive Differential Pulse Code Modulation (ADPCM)".
- [12] Recommendation ITU-T G.729: "Coding of speech at 8 kbit/s using conjugate-structure algebraic-code-excited linear prediction (CS-ACELP)".
- [13] Recommendation ITU-T G.729.1: "G.729-based embedded variable bit-rate coder: An 8-32 kbit/s scalable wideband coder bitstream interoperable with G.729".
- [14] Recommendation ITU-T P.863.1: "Application guide for Recommendation ITU-T P.863".
- [15] Recommendation ITU-T P.340: "Transmission characteristics and speech quality parameters of hands-free terminals".
- [16] Recommendation ITU-T P.501: "Test signals for use in telephony".
- [17] Recommendation ITU-T P.502: "Objective test methods for speech communication systems using complex test signals".
- [18] Recommendation ITU-T P.56: "Objective measurement of active speech level".
- [19] IEC 61260-1: "Electroacoustics - Octave-band and fractional-octave-band filters - Part 1: Specification".
- [20] Recommendation ITU-T P.800.1: "Mean Opinion Score (MOS) terminology".
- [21] ETSI TS 102 971: "Access and Terminals (AT); Public Switched Telephone Network (PSTN); Harmonized specification of physical and electrical characteristics of a 2-wire analogue interface for short line interface".
- [22] ETSI ES 201 970: "Access and Terminals (AT); Public Switched Telephone Network (PSTN); Harmonized specification of physical and electrical characteristics at a 2-wire analogue presented Network Termination Point (NTP)".
- [23] Recommendation ITU-T G.168: "Digital network echo cancellers".
- [24] Recommendation ITU-T P.863: "Perceptual objective listening quality prediction".
- [25] Recommendation ITU-T G.722: "7 kHz audio-coding within 64 kbit/s".
- [26] Recommendation ITU-T G.722.1: "Low-complexity coding at 24 and 32 kbit/s for hands-free operation in systems with low frame loss".
- [27] Recommendation ITU-T G.722.2: "Wideband coding of speech at around 16 kbit/s using Adaptive Multi-Rate Wideband (AMR-WB)".

- [28] Recommendation ITU-T P.1010: "Fundamental voice transmission objectives for VoIP terminals and gateways".
- [29] IETF RFC 3550: "RTP: A Transport Protocol for Real-Time Applications".
- [30] Void.
- [31] TIA-920.130-B: "Telecommunications Communications Products Transmission Requirements for Digital Interface Communications Devices with Headsets".
- [32] ETSI ES 202 737: "Speech and multimedia Transmission Quality (STQ); Transmission requirements for narrowband VoIP terminals (handset and headset) from a QoS perspective as perceived by the user".
- [33] ETSI TS 103 224: "Speech and multimedia Transmission Quality (STQ); A sound field reproduction method for terminal testing including a background noise database".

2.2 Informative references

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

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

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

- [i.1] ETSI EG 202 425: "Speech Processing, Transmission and Quality Aspects (STQ); Definition and implementation of VoIP reference point".
 - [i.2] IETF RFC 2833: "RTP Payload for DTMF Digits, Telephony Tones and Telephony Signals".
 - [i.3] IETF RFC 4733: "RTP Payload for DTMF Digits, Telephony Tones, and Telephony Signals".
 - [i.4] ETSI TS 102 929: "Speech and multimedia Transmission Quality (STQ); Procedures for the identification and selection of common modes of de-jitter buffers and echo cancellers".
 - [i.5] Void.
 - [i.6] Netem™.
- NOTE: Information available at <https://wiki.linuxfoundation.org/networking/netem>.
- [i.7] IETF RFC 4737: "Packet Reordering Metrics".
 - [i.8] ETSI EG 202 396-3: "Speech and multimedia Transmission Quality (STQ); Speech Quality performance in the presence of background noise Part 3: Background noise transmission - Objective test methods".

3 Definition of terms, symbols and abbreviations

3.1 Terms

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

0dB_r point: reference point always located at the digital side of the gateway, for IP-to-IP gateways located at the input of the MGW under test

NOTE: See Recommendation ITU-T G.100.1 [6].

2-wire interface: in the context of the present document, telephony analogue interface over 2-wires used in the local loop

4-wire interface: 4-wire digital interface with separate channels for both directions, irrespective of the physical transmission technology

codec: combination of an analogue-to-digital encoder and a digital-to-analogue decoder operating in opposite directions of transmission in the same equipment

Composite Source Signal (CSS): signal composed in time by various signal elements

MGW with 2-wire interface: MGW with an analogue 2-wire interface (ATA)

MGW with 4-wire interface: MGW with only 4-wire interfaces

EXAMPLE: ITA, IP-to-IP and wireless access points.

receive direction: direction from packet switched interfaces towards a synchronous interface (e.g. ISDN, analogue) or between two packet switched interfaces (for media gateways with packet switched transport on only one side)

NOTE: For media gateways with packet switched transport on both sides (IP-to-IP-MGW), the requirements of the receive direction have to be applied in both directions.

receive interface: interface in the measurement setup, where a receive signal is injected and/or a send signal is measured

reordering: packet order changes during transfer over the network [i.7], packets arrive out of order at the receiver (i.e. RTP packets)

send direction: direction from a synchronous interface (e.g. ISDN, analogue) towards a packet switched interface (for media gateways with packet switched interface on only one side)

NOTE: For media gateways with packet switched interfaces on both sides the requirements of the send direction are not relevant.

send interface: interface in the measurement setup, where a send signal is injected and/or a receive signal is measured

wireless home MGW: home MGW with wireless interface to the phone

EXAMPLE: Wifi or DECT.

3.2 Symbols

Void.

3.3 Abbreviations

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

AM-FM	Amplitude Modulation - Frequency Modulation
AMR	Adaptive Multi Rate codec
ATA	Analogue Terminal Adapter
CLR	Circuit Loudness Rating
CS	Composite Source
CSS	Composite Source Signal
DECT	Digital Enhanced Cordless Telecommunications
DSL	Digital Subscriber Line
DSLAM	Digital Subscriber Line Access Multiplexer
DTMF	Dual Tone Multi Frequency
EC	Echo Cancellor
EFR	Enhance Full Rate codec
EL	Echo Loss
FFS	For Further Study

FFT	Fast Fourier Transformation
FP	DECT Fixed Part
G-MOS-LQOn	Overall transmission quality narrowband
G-MOS-LQOw	Overall transmission quality wideband
GSM	Global System for Mobile communication
GW	GateWay
HATS	Head And Torso Simulator
IAD	Integrated Access Device
IP	Internet Protocol
IPDV	IP Packet Delay Variation
IRS	Intermediate Reference System
ISDN	Integrated Service Digital Network
ITA	ISDN Terminal Adapter
MGW	Media GateWay
MM	Mandatory for both interfaces of the MGW
MOS-LQOy	Mean Opinion Score - Listening Quality Objective

NOTE: y being N for narrowband, M for mixed, S for super-wideband and F for fullband. See Recommendation ITU-T P.800.1 [20].

NA	Not Applicable
NB	NarrowBand
NLP	Non Linear Processor
N-MOS-LQOn	Transmission quality of the background noise narrowband
N-MOS-LQOw	Transmission quality of the background noise wideband
PBX	Private Branch eXchange
PC	Personal Computer
PCM	Pulse Code Modulation
PLC	Packet Loss Concealment
POI	Point Of Interconnect
POTS	Plain Old Telephone Service
PP	DECT Portable Part
PSTN	Public Switched Telephone Network
QoS	Quality of Service
RCV	Receiving Direction
RLR	Receive Loudness Rating
RMS	Root Mean Square
RTP	Real-Time Transport Protocol
SIP	Session Initiation Protocol
SLR	Send Loudness Rating
S-MOS-LQOn	Transmission quality of the speech narrowband
S-MOS-LQOw	Transmission quality of the speech wideband
SND	Sending Direction
TCL	Terminal Coupling Loss
TDM	Time Division Multiplexing
VAD	Voice Activity Detection
VBD	Voice Band Data
VoIP	Voice over Internet Protocol
WB	WideBand

4 General considerations

4.1 Default Coding Algorithm

Narrowband VoIP gateways shall support the coding algorithm according to Recommendation ITU-T G.711 [9] (both μ -law and A-law). VoIP gateways may support other coding algorithms.

Wideband VoIP gateways shall support the coding algorithm according to Recommendation ITU-T G.722 [25]. VoIP gateways may support other coding algorithms.

NOTE: Associated Packet Loss Concealment (PLC) e.g. as defined in Recommendation ITU-T G.711 [9] appendix I should be used.

4.2 End-to-end considerations

In order to achieve a desired end-to-end speech transmission performance (mouth-to-ear) it is recommended that the general rules of transmission planning are carried out with the E-model of Recommendation ITU-T G.107 [3]; this includes the a-priori determination of the desired category of speech transmission quality as defined in Recommendation ITU-T G.109 [5].

While, in general, the transmission characteristics of single circuit-oriented network elements, such as switches or terminals can be assumed to have a single input value for the planning tasks of Recommendation ITU-T G.108 [4] with its amendments, this approach is not applicable in packet based systems and thus there is a need for the transmission planner's specific attention.

In particular the decision as to which delay measured according to the present document is acceptable or representative for the specific configuration is the responsibility of the individual transmission planner.

Recommendation ITU-T G.108 [4] with its amendments provides further guidance on this important issue.

The following optimum parameters from a user's perspective need to be considered:

- Minimized delay in send and receive direction.
- Optimum Circuit Loudness Rating (CLR).
- Compensation for network delay variation.
- Packet loss recovery performance.
- Maximized echo loss.
- Immunity to false detection of DTMF in speech signal.

4.3 Parameters to be investigated

4.3.1 Applicability of parameters to different MGWs

Table 1: Parameter applicability

Clauses in the present document	2-wire home and network MGW	4-wire MGW (excl. IP-to-IP MGW)	4-wire MGW (IP-to-IP-MGW)	wireless home MGW (DECT FP)
6.2 Coding independent parameters				
6.2.1 Send frequency response	M	M	NA	M
6.2.2 Circuit Loudness Rating in Send	M	M	NA	M
6.2.3 Linearity Range for CLR (SND)	M	M	NA	M
6.2.4 Send Distortion	M	M	NA	M
6.2.5 Spurious Out-of-Band Signals in Send direction	M	NA	NA	NA
6.2.6 Send Noise	M	M	NA	M
6.2.7 Receive frequency response	M	M	MM	M
6.2.8 Circuit Loudness Rating in Receive	M	M	MM	M
6.2.9 Linearity Range for CLR (RCV)	M	M	MM	M
6.2.10 Receive Distortion	M	M	MM	M
6.2.11 Out-of-Band Signals in Wideband to Narrowband Transcoding	NA	M	M	M
6.2.12 Spurious Out-of-band Signals Narrowband to Wideband Transcoding	NA	M	M	M
6.2.13 Minimum activation level and sensitivity in Receive direction	FFS	FFS	FFS	FFS
6.2.14 Receive Noise	M	M	MM	M

Clauses in the present document	2-wire home and network MGW	4-wire MGW (excl. IP-to-IP MGW)	4-wire MGW (IP-to-IP-MGW)	wireless home MGW (DECT FP)
6.2.15 Double Talk Performance				
6.2.15.1 Attenuation Range in Send Direction during Double Talk	M (note 1)	M (note 1)	M	M (note 1)
6.2.15.2 Attenuation Range in Receive Direction during Double Talk	M (note 1)	M (note 1)	M	M (note 1)
6.2.15.3 Detection of Echo Components during Double Talk	M (note 1)	M (note 1)	M	M (note 2)
6.2.15.4 Minimum activation level and sensitivity of double talk detection	FFS	FFS	FFS	FFS
6.2.16 Switching characteristics				
6.2.16.1 Activation in Send Direction	M	M	NA	M
6.2.16.2 Activation in Receive Direction	M	M	M	M
6.2.16.3 Silence Suppression and Comfort Noise Generation	FFS	FFS	FFS	FFS
6.2.17 Background Noise Performance				
6.2.17.1 Performance in send direction in the presence of background noise	M	M	MM	M
6.2.17.2 Quality of Speech with Background Noise	M	M	MM	M
6.2.17.3 Quality of Background Noise Transmission (with Far End Speech)	M (note 1)	M (note 1)	MM	M (note 1)
6.2.17.4 Quality of Background Noise Transmission (with Near End Speech)	M	M	MM	M
6.2.18 Quality of echo cancellation				
6.2.18.1 Echo Performance according to Recommendation ITU-T G.168	NA	M (note 1)	NA	NA
6.2.18.2 TCLw (NB)	M (note 1)	M (note 1) (NB)	NA	M (note 2) (NB)
6.2.18.3 TCL (WB)	NA	M (note 1) (WB)	NA	M (note 2) (WB)
6.2.18.4 Temporal echo effects	M (note 1)	M (note 1)	NA	M (note 2)
6.2.18.5 Spectral Echo Attenuation	M (note 1)	M (note 1)	NA	M (note 2)
6.2.18.6 Occurrence of Artefacts	FFS	FFS	NA	FFS
6.2.19 Variant Impairments; Network dependant				
6.2.19.1 Clock accuracy send	M	M	MM	M
6.2.19.2 Clock accuracy receive	M	M	MM	NA
6.2.19.3 Send delay variation	M	M	MM	M
6.2.20 Immunity to DTMF false detection in send direction	M	M	MM	M
6.2.21 Roundtrip Delay	M	M	M	NA
6.3 Codec Specific Requirements				
6.3.1 Objective Listening Speech Quality MOS-LQO in Send direction	M	M	M	M
6.3.2 Objective Listening Speech Quality MOS-LQO in Receive direction	M	M	M	M
6.3.3 Quality of Jitter buffer adjustment	M	M	M	M (note 3)
M:	Mandatory			
MM:	Mandatory for both interfaces of the MGW			
NA:	Not Applicable			
FFS:	For Further Study			
NOTE 1:	Measurement to be done with different echopaths (see clause 6.1.7).			
NOTE 2:	Measurement to be done with Ref PP settings "34/42dB TCLw" only. Echopath set accordingly (see clause 6.1.7).			
NOTE 3:	Measurement mandatory, if PP does not support PLC.			

5 Test equipment

5.1 IP half channel measurement adaptor

The IP half channel measurement adaptor is described in ETSI EG 202 425 [i.1]. Such an apparatus is required to code and insert audio signals into IP packets send to the IP receive interface of the gateway under test, as well as to capture and decode audio signals constituting the payload of IP packets received from the IP sending interface of the gateway under test.

5.2 Environmental conditions for tests

The following conditions shall apply for the testing environment:

- a) Ambient temperature: 15 °C to 35 °C;
- b) Relative humidity: 5 % to 85 %;
- c) Air pressure: 86 kPa to 106 kPa (860 mbar to 1 060 mbar).

5.3 Accuracy of measurements and test signal generation

Unless specified otherwise, the accuracy of measurements made by test equipment shall be equal to or better than:

Table 2: Measurement Accuracy

Item	Accuracy
Electrical signal level	±0,2 dB for levels ≥ -50 dBV ±0,4 dB for levels < -50 dBV
Frequency	±0,2 %
Time	±0,2 %

Unless specified otherwise, the accuracy of the signals generated by the test equipment shall be better than:

Table 3: Accuracy of test signal generation

Quantity	Accuracy
Electrical excitation levels	±0,4 dB across the whole frequency range
Frequency generation	±2 % (see note)
Time	±0,2 %
Specified component values	±1 %
NOTE:	This tolerance may be used to avoid measurements at critical frequencies, e.g. those due to sampling operations within the terminal under test.

If the equipment is powered by other means and those means are not supplied as part of the apparatus, all tests shall be carried out within the power supply limit declared by the supplier. If the power supply is a.c. the test shall be conducted within ±4 % of the rated frequency.

5.4 Network impairment simulation

At least one set of requirements is based on the assumption of an error free packet network, and at least one other set of requirements is based on a defined simulated malperformance of the packet network.

An appropriate network simulator has to be used, for example Netem™ [i.6].

The key points of Netem™ can be summarized as follows:

- Netem™ is part of the networking function of Linux™. With Netem™, there can be generated loss, duplication, delay jitter and reordering (and the jitter distribution can be chosen during runtime). Netem™ can be run on a Linux™-PC running as a bridge or a router.
- It is not advised to define specific distortion patterns for testing in standards, because it will be easy to adapt devices to these patterns (as it is already done for test signals). But if a pattern is unknown to a manufacturer, the same pattern can be used by a test lab for different devices and gives comparable results.

NOTE: Netem™ and Linux™ are examples of suitable products available commercially. This information is given for the convenience of users of the present document and does not constitute an endorsement by ETSI of these products.

Requirements for the network impairment simulation can be found in Annex D of ETSI ES 202 737 [32].

6 Requirements and associated Measurement Methodologies

6.0 General

Differences between different media gateway types are dealt with in the respective requirements.

In the case of IP-to-IP MGW packet based interfaces are provided at both sides of the gateway. Therefore the receive requirements apply, for both interfaces.

In the case of wireless home MGW (DECT FP) the applicable parameters in Table 1 should be measured according to the 4-wire MGW requirements and measurement methods (wideband and narrowband) if nothing else is stated.

NOTE 1: In general the test methods as described in the present document apply. If alternative methods exist they may be used if they have been proven to give the same result as the method described in the present document.

NOTE 2: Due to the time variant nature of IP connections delay variation may impair the measurements. In such cases the measurement has to be repeated until a valid measurement result is achieved.

6.1 Test setup

6.1.0 General

The preferred way of testing a gateway is to connect its interfaces to network simulators with exact defined settings and access points. The test sequences are fed in electrically, using a reference codec or using the direct signal processing approach.

When VoIP runs on the gateway under test only in conjunction with a registration by an application server (e.g. SIP proxy), the network simulator may need to provide also the registration functionality.

Alternatively, if for the IP-interfaces another technology than Ethernet is used (for instance DSL access, it may be necessary to add additional equipment in the test setup for connecting the measurement equipment (e.g. a DSLAM, if the IP-interface works over DSL). There should be no speech signal processing in this additional equipment (the media payload has to be passed transparent through this equipment, while e.g. header manipulation is allowed). The influence of this additional equipment (delay and eventually delay variation) has to be taken into account for the measurements.

NOTE 1: It is up to the testlab to identify potential time invariances or non linearities in the network used for interconnection and to take those effects into account properly.

NOTE 2: It is up to the testlab to identify potential additional delays caused by low bandwidth connections (e.g. MGW with DSL connection and a low bitrate profile) and to take those effects into account properly.

With this setup it is possible to measure the parameters listed in the present document over a whole network, if the behaviour of the network is known.

In the present document, the terms "send" and "receive" can be found in the pictures of the relevant test setup.

When a coder with variable bit rate is used for testing the MGW parameters, the bit rate recognized giving the best characteristics and/or the ones commonly used should be selected, e.g.:

- AMR-NB (ETSI TS 126 171 [2]): 12,2 kbit/s.
- AMR-WB (G.722.2 [27]): 12,65 kbit/s and/or 23,85 kbit/s.
- Recommendation ITU-T G.729.1 [13]: 32 kbit/s.

NOTE 3: Although packet capturing and network simulation in figures 1 to 4 are shown in one box they may be separate devices.

6.1.1 Setup for Media Gateways with 4-wire interface

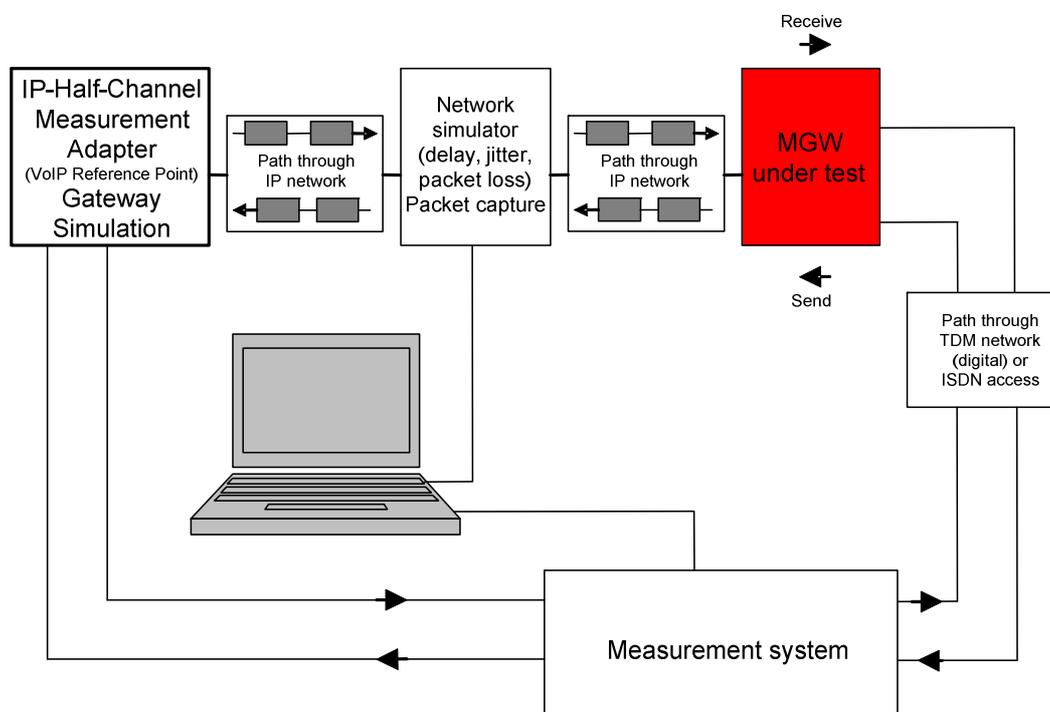


Figure 1: Half channel measurement for MGW with 4-wire interface

6.1.2 Setup for Media Gateways with 2-wire interface

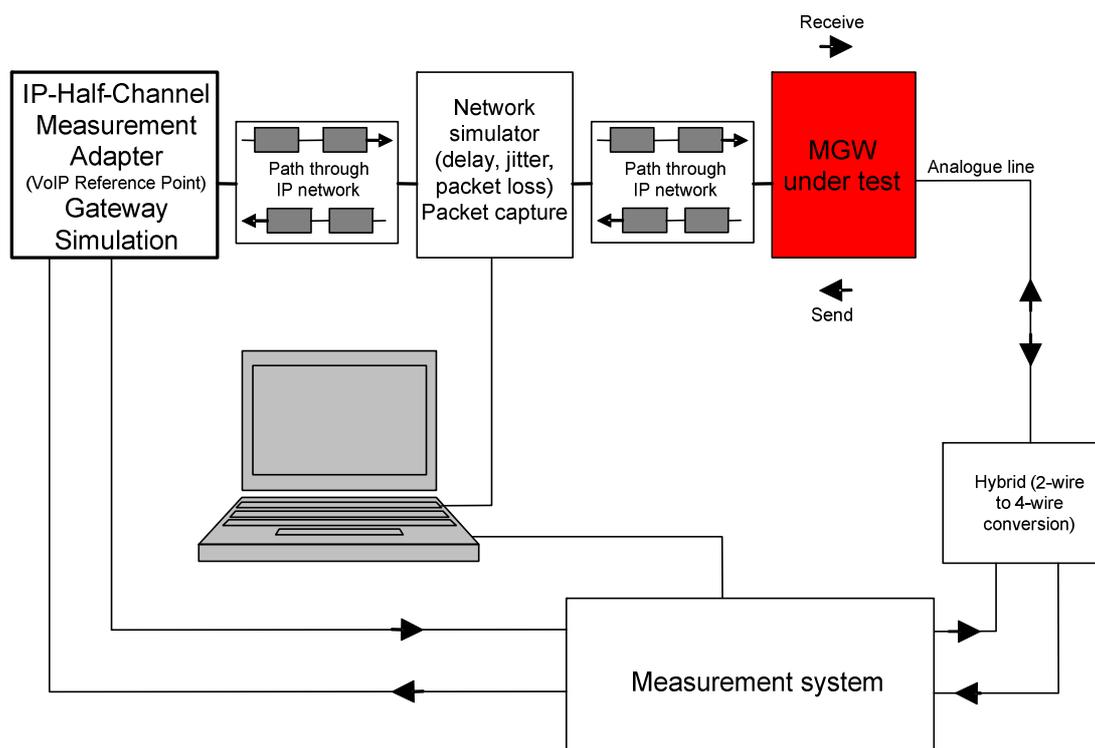


Figure 2: Half channel measurement for MGW with 2-wire interface

6.1.3 Setup for Media Gateways with Wireless Access

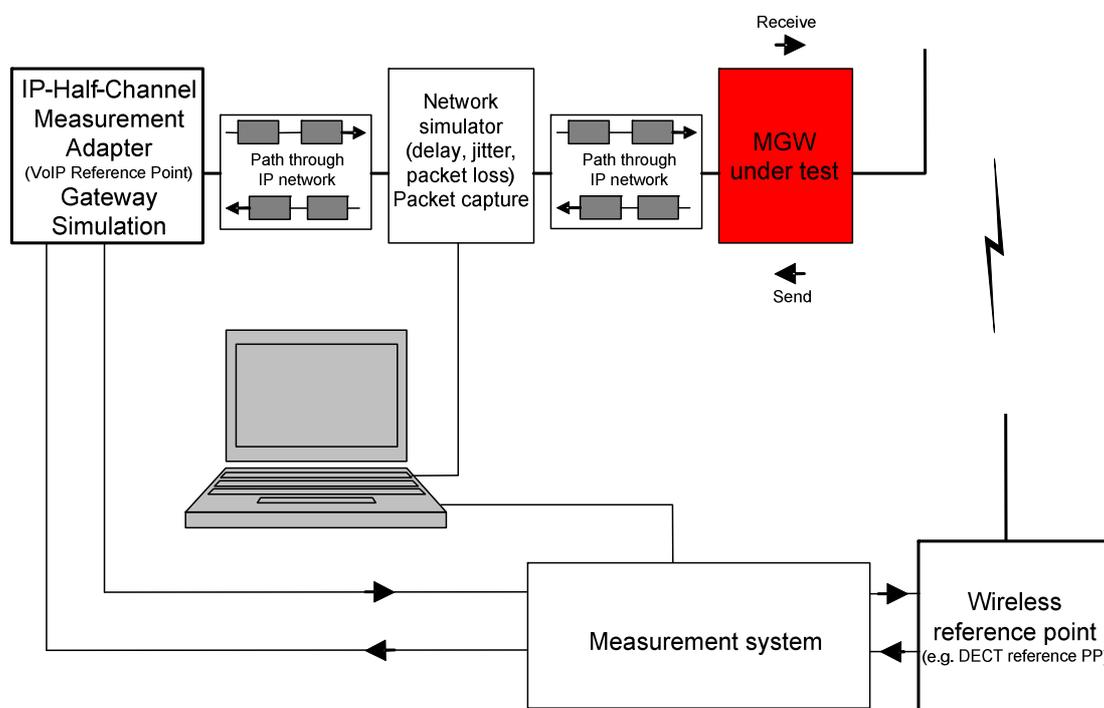
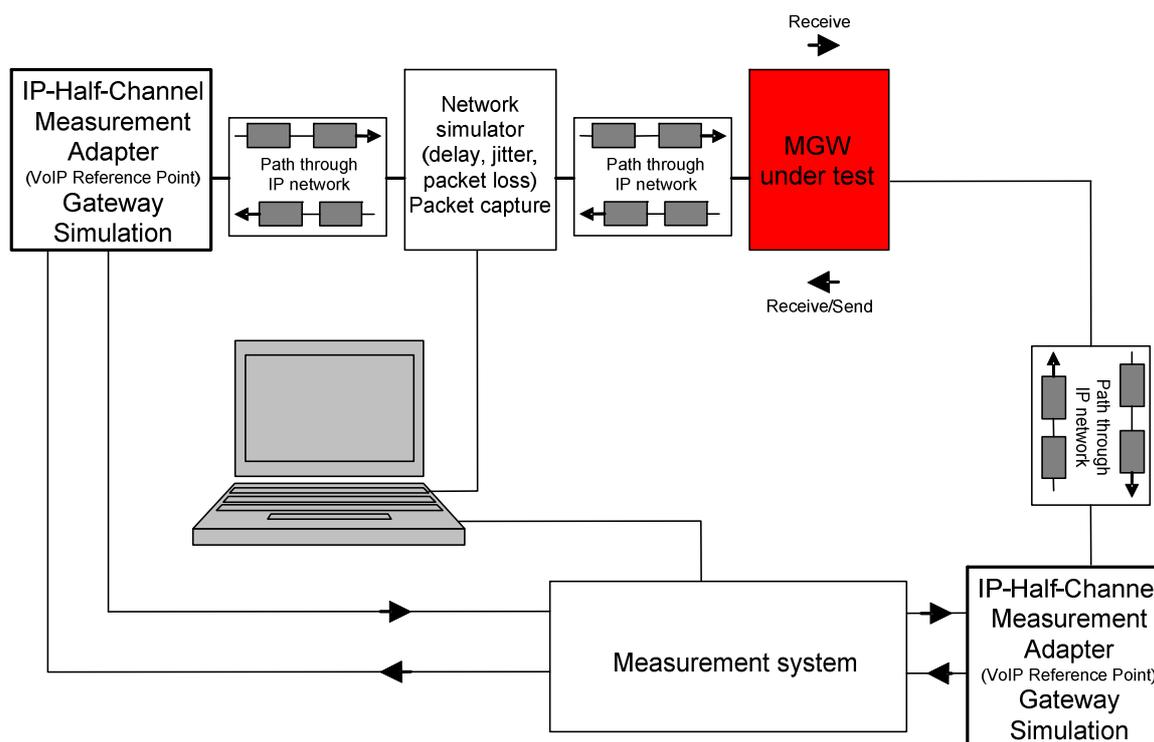


Figure 3: Half channel measurement for MGW with wireless access

6.1.4 Setup for IP-to-IP Media Gateways



NOTE 1: For measuring both directions of an IP-to-IP MGW, the network simulator can be moved to the other side of the MGW.

NOTE 2: If the network simulator is moved to the other side of the IP-to-IP MGW the measurements to be conducted in receive direction are to be conducted in this scenario on the other channel (send in the previous scenario).

Figure 4: Half channel measurement for IP-to-IP MGW

6.1.5 Test Signal Levels

Unless specified otherwise, the test signal level at the digital inputs shall be -16 dBm0. For analogue inputs (2-wire) the test signal level should be -13 dBm for home MGWs and -16 dBm for network MGWs at the media gateway interface.

NOTE: For analogue inputs of home MGWs, the test signal has to be 3 dB higher than for network MGWs, to take into account the direct connection of a phone to the home MGW (without access line as for the network MGW).

6.1.6 Background noise simulation

Background noise signals used for testing should be recorded according to the description in ETSI TS 103 224 [33].

ETSI TS 103 224 [33] contains a description of the recording arrangement for realistic background noises, a description of the setup for a loudspeaker arrangement suitable to simulate a background noise field in a lab-type environment and a database of realistic background noises, which can be used for testing the terminal performance with a variety of different background noises.

In order to create a representative electrical test signal for the MGW containing speech at a nominal level mixed with the amount of background noise picked up by a terminal, the setup in ETSI TS 103 224 [33] is used. The terminal is connected to a reference interface providing nominal properties for the electrical interface as used by the terminal. The signal (speech plus noise) is recorded at this interface and inserted through the appropriate reference interfaces as described in clauses 6.1.1 to 6.1.4 in such a way that the signal level and spectral content delivered to the MGW under test is equivalent to the one it would have seen if the terminal was connected directly. Either terminals considered to be representative for the type of terminal attached to the MGW are used or individual terminals are used.

NOTE: Due to terminal geometry, microphone technique and signal processing used in the phone the signals acquired at the electrical interface may highly vary. As a consequence the performance of the MGW may highly depend on the terminal connected.

The following noises of ETSI TS 103 224 [33] shall be used.

Table 4: Background noises

Pub Noise (Pub)	HATS and microphone array in a pub	30 s	1: 77,2 dB 2: 76,6 dB 3: 75,7 dB 4: 76,0 dB 5: 76,0 dB 6: 76,3 dB 7: 76,0 dB 8: 76,4 dB
Sales Counter (SalesCounter)	HATS and microphone array in a supermarket	30 s	1: 66,6 dB 2: 66,1 dB 3: 65,7 dB 4: 66,5 dB 5: 66,3 dB 6: 66,8 dB 7: 66,6 dB 8: 67,1 dB
Callcenter 2 (Callcenter)	HATS and microphone array in business office	30 s	1: 60,2 dB 2: 60,0 dB 3: 60,1 dB 4: 60,8 dB 5: 60,2 dB 6: 60,6 dB 7: 60,2 dB 8: 60,7 dB

6.1.7 Echo paths

For some measurements (e.g. double talk performance) the echo path has an influence on the measurement results. This means that such measurements have to be done with some typical different echo paths. The echo paths should reflect some typical terminals and network conditions, including worst case scenarios. Nonetheless it has to be ensured that the used echo attenuations are within the specified limits of G.168 [23].

For narrowband 2-wire MGW the following echo path simulations are used:

- Different DECT phone echo paths
- Different analogue phone echo paths
- Different analogue hands-free echo paths

For narrowband 4-wire MGW the following echo path simulations are used:

- Different ISDN phone echo paths
- Different ISDN phone hands-free echo paths
- Different DECT phone echo paths
- Infinite echo loss
- Different analogue terminal echo paths
- Different analogue terminal hands-free echo paths

If the narrowband 4-wire mediagateway is connected to a TDM network the additional delay introduced by the TDM network has to be included in the echopath.

For wireless home MGW (DECT FP) the following echo path simulations are used:

- Different DECT PP echo paths, but not below a TCLw of 34 dB (NB), 42 dB (WB)

No explicit echo paths are given in the present document. The types of echo paths used are subject to be agreed upon with the network operator or the test house.

The same set of selected echo paths should be used for all measurements which need different echo paths.

NOTE 1: With the exception of the measurements in clause 6.2.18.1 the echo canceller has to be adapted to each echo path before starting the measurements. This is done by setting the selected echo path and sending at least two sentences (one male, one female) of the British-English single talk sequence described in clause 7.3.2 of Recommendation ITU-T P.501 [16] in the send direction.

NOTE 2: A DECT phone echo path impulse response is provided in annex A and can be used in combination with appropriate hybrid echo simulation or artificial echo loss simulation in case of an ISDN fixed part.

6.2 Coding independent parameters

6.2.1 Send Frequency response

Requirement

The frequency response for 4-wire MGW shall be according to Table 5 and Table 6.

Table 5: Send frequency response for 4-wire MGW

Frequency	Upper Limit	Lower Limit
100 Hz	2 dB	
300 Hz	2 dB	-2 dB
3 400 Hz	2 dB	-2 dB
4 000 Hz	2 dB	

Table 6: Send frequency response for wideband 4-wire MGW

Frequency	Upper Limit	Lower Limit
100 Hz	2 dB	
200 Hz	2 dB	-2 dB
7 000 Hz	2 dB	-2 dB
8 000 Hz	2 dB	

NOTE: The frequency response characteristics requirements apply to codecs having flat response characteristics. If a codec with non-flat characteristics is used the requirement has to be corrected by the ideal response characteristics of this codec.

The frequency response for 2-wire MGW shall be according to Table 7 (see ETSI TS 102 971 [21]) for both home and network MGWs.

Table 7: Send frequency response for narrowband 2-wire MGW

Frequency	Upper Limit	Lower Limit
0 Hz	2 dB	
300 Hz	2 dB	-2 dB
3 400 Hz	2 dB	-2 dB
4 000 Hz	2 dB	

Measurement Method

The test signal to be used for the measurements shall be the British-English single talk sequence described in clause 7.3.2 of Recommendation ITU-T P.501 [16]. The test signal level is averaged over the complete test signal sequence.

Measurements shall be made at one twelfth-octave intervals as given by the R.40 series of preferred numbers in IEC 61260-1 [19] for frequencies from 100 Hz to 4 kHz (8 kHz for wideband 4-wire MGWs) inclusive. For the calculation the averaged measured level at each frequency band is referred to the averaged test signal level measured in each frequency band.

The sensitivity is expressed in terms of dB.

6.2.2 Circuit Loudness Rating in Send

Requirement

The nominal value of Circuit Loudness Rating in Send (CLR) for MGWs with 4-wire interface shall be:

- $CLR(SND) = 0 \text{ dB} \pm 1 \text{ dB}$

The nominal value of Circuit Loudness Rating in Send (CLR) for MGWs with 2-wire interface shall be (in accordance with ETSI TS 102 971 [21] and ETSI ES 201 970 [22]):

- $CLR(SND) = 3 \text{ dB} \pm 1 \text{ dB}$ (Home MGW)
- $CLR(SND) = 0 \text{ dB} \pm 1 \text{ dB}$ (Network MGW)

Measurement Method

The test signal to be used for the measurements shall be the British-English single talk sequence described in clause 7.3.2 of Recommendation ITU-T P.501 [16]. The test signal level is averaged over the complete test signal sequence.

For narrowband MGWs, the send sensitivity shall be calculated as average loss in the frequency range from 300 Hz to 3 400 Hz (Recommendation ITU-T G.111, annex A [7]).

For wideband MGWs, the send sensitivity shall be calculated as average loss in the frequency range from 100 Hz to 7 000 Hz.

The sensitivity is expressed in terms of dB.

6.2.3 Linearity Range for CLR (SND)

Requirement

For MGW with 4-wire interface, the send sensitivity determined with input levels between -36 dBm0 and -11 dBm0 shall not differ by more than $\pm 0,5 \text{ dB}$ from the send sensitivity determined with an input level of -16 dBm0. For the input level of -6 dBm0 a limit of $+2/-1 \text{ dB}$ applies.

Table 8: Linearity range for CLR (SND) for MGW with 4-wire interface

Linearity range of CLR: $\Delta CLR = CLR - CLR@-16 \text{ dBm0}$			
Input Level	Target ΔCLR	Upper limit	Lower limit
-36 dBm0	0	0,5 dB	-0,5 dB
-31 dBm0	0	0,5 dB	-0,5 dB
-26 dBm0	0	0,5 dB	-0,5 dB
-21 dBm0	0	0,5 dB	-0,5 dB
-16 dBm0	0	0 dB	0 dB
-11 dBm0	0	0,5 dB	-0,5 dB
-6 dBm0	0	2 dB	-1 dB

NOTE 1: With -6 dBm0 clipping of the measurement signal occurs already in the digital coding. This behaviour is reflected in the enlarged limits for this input level.

For Home-MGW with 2-wire interface, the send sensitivity determined with input levels between -33 dBm and -8 dBm shall not differ by more than $\pm 0,5 \text{ dB}$ from the send sensitivity determined with an input level of -16 dBm. For the input level of -3 dBm a limit of $+2/-1 \text{ dB}$ applies.

Table 9: Linearity range for CLR (SND) for Home-MGW with 2-wire interface

Linearity range of CLR: $\Delta\text{CLR} = \text{CLR} - \text{CLR}@-16 \text{ dBm}_0$			
Input Level	Target ΔCLR	Upper limit	Lower limit
-33 dBm	0	0,5 dB	-0,5 dB
-28 dBm	0	0,5 dB	-0,5 dB
-23 dBm	0	0,5 dB	-0,5 dB
-18 dBm	0	0,5 dB	-0,5 dB
-13 dBm	0	0 dB	0 dB
-8 dBm	0	0,5 dB	-0,5 dB
-3 dBm	0	2 dB	-1 dB

For Network-MGW with 2-wire interface, the send sensitivity determined with input levels between -40 dBm and -6 dBm shall not differ by more than $\pm 0,5$ dB from the send sensitivity determined with an input level of -20 dBm. For the input level of -6 dBm a limit of +2/-1 dB applies.

Table 10: Linearity range for CLR (SND) for Network-MGW with 2-wire interface

Linearity range of CLR: $\Delta\text{CLR} = \text{CLR} - \text{CLR}@-16 \text{ dBm}_0$			
Input Level	Target ΔCLR	Upper limit	Lower limit
-40 dBm	0	0,5 dB	-0,5 dB
-35 dBm	0	0,5 dB	-0,5 dB
-30 dBm	0	0,5 dB	-0,5 dB
-25 dBm	0	0,5 dB	-0,5 dB
-20 dBm	0	0,5 dB	-0,5 dB
-15 dBm	0	0 dB	0 dB
-10 dBm	0	0,5 dB	-0,5 dB
-6 dBm	0	2 dB	-1 dB

NOTE 2: It is assumed that the variation of gain is mostly codec independent. In case codec specific requirements are needed this is found in the codec specific clause.

NOTE 3: The broader tolerance range at the highest input level takes into account that at this level clipping will occur.

Measurement Method

The test signal to be used for the measurements shall be the British-English single talk sequence described in clause 7.3.2 of Recommendation ITU-T P.501 [16]. The test signal level is averaged over the complete test signal sequence.

For narrowband MGWs, the send sensitivity shall be calculated as average loss in the frequency range from 300 Hz to 3 400 Hz (Recommendation ITU-T G.111, annex A [7]).

For wideband MGWs, the send sensitivity shall be calculated as average loss in the frequency range from 100 Hz to 7 000 Hz.

6.2.4 Send Distortion

Requirement

The ratio of signal to harmonic distortion shall be above the following mask.

Table 11: Send distortion for narrowband MGW

Frequency	Signal to distortion ratio limit, receive
315 Hz	26 dB
400 Hz	30 dB
500 Hz	30 dB
800 Hz	30 dB
1 kHz	30 dB

Frequency	Signal to distortion ratio limit, receive
NOTE:	Limits at intermediate frequencies lie on a straight line drawn between the given values on a linear (dB ratio) - logarithmic (frequency) scale.

Table 12: Send distortion for wideband MGW

Frequency	Signal to distortion ratio limit, receive
315 Hz	26 dB
400 Hz	30 dB
500 Hz	30 dB
800 Hz	30 dB
1 kHz	30 dB
2 kHz	30 dB
NOTE:	Limits at intermediate frequencies lie on a straight line drawn between the given values on a linear (dB ratio) - logarithmic (frequency) scale.

Measurement Method

The signal used is an activation signal followed by a sine-wave signal with a frequency at 315 Hz, 400 Hz, 500 Hz, 630 Hz, 800 Hz and 1 000 Hz for narrowband, additionally 2 000 Hz for wideband.

The female speaker signal of the short conditioning sequence described in clause 7.3.7 of Recommendation ITU-T P.501 [16] shall be used for activation.

Measurements are made at 315 Hz, 400 Hz, 500 Hz, 630 Hz, 800 Hz, 1 000 Hz for narrowband, additionally 2 000 Hz for wideband.

The signal to harmonic distortion ratio is measured selectively up to 3,15 kHz for narrowband MGW, up to 6,3 kHz for wideband MGW.

NOTE: Depending on the type of codec the test signal used may need to be adapted.

6.2.5 Spurious Out-of-Band Signals in Send direction

Requirement

For MGW with 2-wire interface:

- The level of any image frequency produced in the time slot corresponding to the input connection should be at least 25 dB below the level of the test signal.

Measurement Method

A sine-wave signal in the range from 4,6 kHz to 72 kHz is applied to the 2-wire interface of an input connection at a level of -25 dBm₀. The level of any image frequency produced in the time slot corresponding to the input connection is measured and referred to the level of the test signal.

6.2.6 Send Noise

Requirement

For MGW with 4-wire interface:

- The maximum noise level produced by the MGW under silent conditions in the send direction shall not exceed -64 dBm_{0p} for narrowband MGWs and -68 dBm₀ (A) for wideband MGWs.
- No peaks in the frequency domain higher than 10 dB above the average noise spectrum shall occur.

For MGW with 2-wire interface:

- The maximum noise level produced by the MGW under silent conditions in the send direction shall not exceed -67 dBm_{0p}, see ETSI TS 102 971 [21].
- No peaks in the frequency domain higher than 10 dB above the average noise spectrum shall occur.

Measurement Method

For the actual measurement no test signal is used. In order to reliably activate the terminal an activation signal is introduced before the actual measurement. The female speaker signal of the short conditioning sequence described in clause 7.3.7 of Recommendation ITU-T P.501 [16] shall be used for activation. The activation signal level shall be -16 dBm₀. The activation signal level is averaged over the complete activation signal sequence.

The send noise is measured at the receive interface in the frequency range from 100 Hz to 4 kHz for narrowband MGWs, from 100 Hz to 8 kHz for wideband MGWs. The analysis window is applied directly after stopping the activation signal but taking into account the influence of all acoustical components (reverberations). The averaging time is 1 s. It has to be ensured (e.g. by monitoring the time signal) that during the test the terminal remains in activated condition. If the terminal is deactivated during the measurement, the measurement time has to be reduced to the period where the terminal remains in activated condition.

The noise level is measured in dBm_{0p} for narrowband MGW, in dBm₀ (A) for wideband MGW.

Spectral peaks are measured in the frequency domain in the frequency range from 100 Hz to 3,4 kHz (narrowband MGW) or 100 Hz to 6,3 kHz (wideband MGW). The frequency spectrum of the idle channel noise is measured by a spectral analysis having a noise bandwidth of 8,79 Hz (determined using FFT 8 k samples/48 kHz sampling rate with Hanning window or equivalent). The idle channel noise spectrum is stated in dB. A smoothed average idle channel noise spectrum is calculated by a moving average (arithmetic mean) 1/3rd octave wide across the idle noise channel spectrum stated in dB (linear average in dB of all FFT bins in the range from $2^{(-1/6)f}$ to $2^{(+1/6)f}$). Peaks in the idle channel noise spectrum are compared against a smoothed average idle channel noise spectrum.

6.2.7 Receive Frequency Response

The frequency response for 4-wire MGW shall be according to Table 13 and Table 14.

Table 13: Receive frequency response for 4-wire MGW

Frequency	Upper Limit	Lower Limit
100 Hz	2 dB	
300 Hz	2 dB	-2 dB
3 400 Hz	2 dB	-2 dB
4 000 Hz	2 dB	

Table 14: Receive frequency response for wideband 4-wire MGW

Frequency	Upper Limit	Lower Limit
100 Hz	2 dB	
200 Hz	2 dB	-2 dB
7 000 Hz	2 dB	-2 dB
8 000 Hz	2 dB	

The frequency response for 2-wire MGW shall be according to Table 15 (see ETSI TS 102 971 [21]) for both home and network MGWs.

Table 15: Receive frequency response for narrowband 2-wire MGW

Frequency	Upper Limit	Lower Limit
0 Hz	2 dB	
300 Hz	2 dB	-2 dB
3 400 Hz	2 dB	-2 dB
4 000 Hz	1 dB	

The frequency response for 2-wire MGW shall be according to ETSI TS 102 971 [21] for both home and network MGWs.

Measurement Method

The test signal to be used for the measurements shall be the British-English single talk sequence described in clause 7.3.2 of Recommendation ITU-T P.501 [16]. The test signal level is averaged over the complete test signal sequence.

Measurements shall be made at one twelfth-octave intervals as given by the R.40 series of preferred numbers in IEC 61260-1 [19] for frequencies from 100 Hz to 4 kHz (8 kHz for wideband 4-wire MGWs) inclusive. For the calculation the averaged measured level at each frequency band is referred to the averaged test signal level measured in each frequency band.

The sensitivity is expressed in terms of dB.

6.2.8 Circuit Loudness Rating in Receive

The nominal value of Circuit Loudness Rating in Receive (CLR) for MGWs with 4-wire interface shall be:

$$\text{CLR(RCV)} = 0 \text{ dB} \pm 1 \text{ dB}$$

The nominal value of Circuit Loudness Rating in Receive (CLR) for MGWs with 2-wire interface shall be (in line with ETSI TS 102 971 [21] and ETSI ES 201 970 [22]):

- CLR(RCV) = 10 dB \pm 1 dB (Home MGW)
- CLR(RCV) = 7 dB \pm 1 dB (Network MGW)

NOTE: CLR (RCV) for MGW with 2-wire interface can differ from the above recommended value due to national transmission plans.

Measurement Method

The test signal to be used for the measurements shall be the British-English single talk sequence described in clause 7.3.2 of Recommendation ITU-T P.501 [16]. The test signal level is averaged over the complete test signal sequence.

For narrowband MGWs, the receive sensitivity shall be calculated as average loss in the frequency range from 300 Hz to 3 400 Hz (Recommendation ITU-T G.111 [7], annex A).

For wideband MGWs, the receive sensitivity shall be calculated as average loss in the frequency range from 100 Hz to 7 000 Hz.

The sensitivity is expressed in terms of dB.

6.2.9 Linearity Range for CLR (RCV)

Requirement

For MGW with 4-wire interface, the receive sensitivity determined with input levels between -36 dBm0 and -11 dBm0 shall not differ by more than $\pm 0,5$ dB from the receive sensitivity determined with an input level of -16 dBm0. For the input level of -6 dBm0 a limit of +2/-1 dB applies.

Table 16: Linearity range for CLR (RCV) for MGW with 4-wire interface

Linearity range of CLR: $\Delta\text{CLR} = \text{CLR} - \text{CLR}@-16 \text{ dBm0}$			
Input Level	Target ΔCLR	Upper limit	Lower limit
-36 dBm0	0	0,5 dB	-0,5 dB
-31 dBm0	0	0,5 dB	-0,5 dB
-26 dBm0	0	0,5 dB	-0,5 dB
-21 dBm0	0	0,5 dB	-0,5 dB
-16 dBm0	0	0 dB	0 dB
-11 dBm0	0	0,5 dB	-0,5 dB
-6 dBm0	0	2 dB	-1 dB

NOTE 1: With -6 dBm0 clipping of the measurement signal occurs already in the digital coding. This behaviour is reflected in the enlarged limits for this input level.

For Home-MGW with 2-wire interface, the receive sensitivity determined with input levels between -36 dBm and -16 dBm shall not differ by more than $\pm 0,5$ dB from the send sensitivity determined with an input level of -16 dBm0. For the input level of -3 dBm a limit of $\pm 2/-1$ dB applies.

Table 17: Linearity range for CLR (RCV) for Home-MGW with 2-wire interface

Linearity range of CLR: $\Delta\text{CLR} = \text{CLR} - \text{CLR}@-16 \text{ dBm0}$			
Input Level	Target ΔCLR	Upper limit	Lower limit
-36 dBm	0	0,5 dB	-0,5 dB
-31 dBm	0	0,5 dB	-0,5 dB
-26 dBm	0	0,5 dB	-0,5 dB
-21 dBm	0	0,5 dB	-0,5 dB
-16 dBm	0	0 dB	0 dB
-11 dBm	0	0,5 dB	-0,5 dB
-6 dBm	0	2 dB	-1 dB

For Network-MGW with 2-wire interface, the receive sensitivity determined with input levels between -40 dBm and -6 dBm shall not differ by more than $\pm 0,5$ dB from the send sensitivity determined with an input level of -20 dBm. For the input level of -6 dBm a limit of $\pm 2/-1$ dB applies.

Table 18: Linearity range for CLR (RCV) for Network-MGW with 2-wire interface

Linearity range of CLR: $\Delta\text{CLR} = \text{CLR} - \text{CLR}@-16 \text{ dBm0}$			
Input Level	Target ΔCLR	Upper limit	Lower limit
-40 dBm	0	0,5 dB	-0,5 dB
-35 dBm	0	0,5 dB	-0,5 dB
-30 dBm	0	0,5 dB	-0,5 dB
-25 dBm	0	0,5 dB	-0,5 dB
-20 dBm	0	0,5 dB	-0,5 dB
-15 dBm	0	0 dB	0 dB
-10 dBm	0	0,5 dB	-0,5 dB
-6 dBm	0	2 dB	-1 dB

NOTE 2: It is assumed that the variation of gain is mostly codec independent. In case codec specific requirements are needed this is found in the codec specific clause.

NOTE 3: The broader tolerance range at the highest input level takes into account that at this level clipping will occur.

Measurement Method

The test signal to be used for the measurements shall be the British-English single talk sequence described in clause 7.3.2 of Recommendation ITU-T P.501 [16]. The test signal level is averaged over the complete test signal sequence.

For narrowband MGWs, the receive sensitivity shall be calculated as average loss in the frequency range from 300 Hz to 3 400 Hz (Recommendation ITU-T G.111 [7], annex A).

For wideband MGWs, the receive sensitivity shall be calculated as average loss in the frequency range from 100 Hz to 7 000 Hz.

6.2.10 Receive Distortion

Requirement

The ratio of signal to harmonic distortion shall be above the following mask.

Table 19: Receive distortion for narrowband MGW

Frequency	Signal to distortion ratio limit, receive
315 Hz	26 dB
400 Hz	30 dB
500 Hz	30 dB
800 Hz	30 dB
1 kHz	30 dB
NOTE:	Limits at intermediate frequencies lie on a straight line drawn between the given values on a linear (dB ratio) - logarithmic (frequency) scale.

Table 20: Receive distortion for wideband MGW

Frequency	Signal to distortion ratio limit, receive
315 Hz	26 dB
400 Hz	30 dB
500 Hz	30 dB
800 Hz	30 dB
1 kHz	30 dB
2 kHz	30 dB
NOTE:	Limits at intermediate frequencies lie on a straight line drawn between the given values on a linear (dB ratio) - logarithmic (frequency) scale.

Measurement Method

The signal used is an activation signal followed by a sine-wave signal with a frequency at 315 Hz, 400 Hz, 500 Hz, 630 Hz, 800 Hz and 1 000 Hz for narrowband, additionally 2 000 Hz for wideband.

The female speaker signal of the short conditioning sequence described in clause 7.3.7 of Recommendation ITU-T P.501 [16] shall be used for activation.

Measurements are made at 315 Hz, 400 Hz, 500 Hz, 630 Hz, 800 Hz, 1 000 Hz for narrowband, additionally 2 000 Hz for wideband.

The signal to harmonic distortion ratio is measured selectively up to 3,15 kHz for narrowband MGW, up to 6,3 kHz for wideband MGW.

NOTE: Depending on the type of codec the test signal used may need to be adapted.

6.2.11 Out-of-Band Signals Wideband to Narrowband Transcoding

For MGW with 4-wire interface, this parameter has to be measured only, if the MGW supports a mode, where one side is wideband and the other side narrowband (transcoding wideband to narrowband). This measurement is only applicable for transcoding from wideband to narrowband.

Requirement

With any signal above 4,6 kHz and up to 8 kHz applied with a level of -16 dBm0, the level of any image frequency shall be below the level obtained for the reference signal by at least the amount (in dB) specified in Table 21.

Table 21: Out-of-band signal limit, receive

Frequency	Minimum attenuation
4,6 kHz	30 dB
8 kHz	40 dB
NOTE: The limits for intermediate frequencies lie on a straight line drawn between the given values on a linear (dB) - logarithmic (kHz) scale.	

Measurement Method

For a correct activation of the system, the female speaker signal of the short conditioning sequence described in clause 7.3.7 of Recommendation ITU-T P.501 [16] shall be used for activation. Level of this activation signal shall be -16 dBm0.

For the test, an out-of-band signal shall be provided as a frequency band signal centred on 4,65 kHz, 5 kHz, 6 kHz, 6,5 kHz, 7 kHz and 7,5 kHz respectively. The level of any image frequencies at the digital interface shall be measured.

The levels of these signals shall be -16 dBm0.

The complete test signal is constituted by t1 ms of in-band signal (reference signal), t2 ms of out-of-band signal and another time t1 ms of in-band signal (reference signal).

The observation of the output signal on the first and second in-band signals permits the control if the set is correctly activated during the out-of-band measurement. This measurement shall be performed during t2 period.

A value of 250 ms is suggested for t1.

t2 depends on the integration time of the analyser, typically less than 150 ms.

NOTE: Depending on the type of codec the test signal used may need to be adapted.

6.2.12 Spurious Out-of-band Signals Narrowband to Wideband Transcoding

For MGW with 4-wire interface, this parameter has to be measured only if the MGW supports a mode, where one side is wideband and the other side narrowband (transcoding wideband to narrowband). This measurement is for transcoding from narrowband to wideband.

If the gateway provides wideband extension techniques this measurement is not applicable.

Requirement

Any spurious out-of-band image signals in the frequency range from 4,6 kHz to 8 kHz measured selectively shall be lower than the in-band level measured with a reference signal. The minimum level difference between the reference signal level and the out-of-band image signal level shall be as given in Table 22.

Table 22: Spurious out of band signal limits, receive

Frequency	Minimum attenuation
4,6 kHz	35 dB
8 kHz	45 dB
NOTE: The limits for intermediate frequencies lie on a straight line drawn between the given values on a linear (dB) - logarithmic (kHz) scale.	

Measurement Method

The test arrangement is according to clause 6.1.

The signal used is an activation signal followed by a sine-wave signal. For input signals at the frequencies 500 Hz, 1 000 Hz, 2 000 Hz and 3 150 Hz applied at the level of -16 dBm0, the level of spurious out-of-band image signals at frequencies up to 8 kHz is measured selectively at measurement point.

The female speaker signal of the short conditioning sequence described in clause 7.3.7 of Recommendation ITU-T P.501 [16] shall be used for activation. The level of this activation signal is -16 dBm0.

NOTE: Depending on the type of codec the test signal used may need to be adapted.

6.2.13 Minimum Activation Level and Sensitivity in Receive Direction

For further study.

6.2.14 Receive Noise

Requirement

For MGW with 4-wire interface: The maximum noise level produced by the MGW under silent conditions in the receive direction shall not exceed -64 dBm0p for narrowband MGWs and -68 dBm0 (A) for wideband MGWs.

No peaks in the frequency domain higher than 10 dB above the average noise spectrum shall occur.

For MGW with 2-wire interface:

The maximum noise level produced by the MGW under silent conditions in the receive direction shall not exceed -70 dBm0p, see ETSI TS 102 971 [21].

No peaks in the frequency domain higher than 10 dB above the average noise spectrum shall occur.

Measurement Method

For the actual measurement no test signal is used. In order to reliably activate the terminal an activation signal is introduced before the actual measurement. The female speaker signal of the short conditioning sequence described in clause 7.3.7 of Recommendation ITU-T P.501 [16] shall be used for activation. The activation signal level shall be -16 dBm0 for 4-wire MGW, -16 dBm for 2-wire home-GW and -19 dBm for 2-wire network MGW. The activation signal level is averaged over the complete activation signal sequence.

The receive noise is measured at the send interface in the frequency range from 100 Hz to 4 kHz for narrowband MGWs, from 100 Hz to 8 kHz for wideband MGWs. The analysis window is applied directly after stopping the activation signal but taking into account the influence of all acoustical components (reverberations). The averaging time is 1 s. It has to be ensured (e.g. by monitoring the time signal) that during the test the terminal remains in activated condition. If the terminal is deactivated during the measurement, the measurement time has to be reduced to the period where the terminal remains in activated condition.

The noise level is measured in dBm0p for narrowband MGW, in dBm0 (A) for wideband MGW.

Spectral peaks are measured in the frequency domain in the frequency range from 100 Hz to 3,4 kHz (narrowband MGW) or 100 Hz to 6,3 kHz (wideband MGW). The frequency spectrum of the idle channel noise is measured by a spectral analysis having a noise bandwidth of 8,79 Hz (determined using FFT 8 k samples/48 kHz sampling rate with Hanning window or equivalent). The idle channel noise spectrum is stated in dB. A smoothed average idle channel noise spectrum is calculated by a moving average (arithmetic mean) $1/3^{\text{rd}}$ octave wide across the idle noise channel spectrum stated in dB (linear average in dB of all FFT bins in the range from $2^{(-1/6)f}$ to $2^{(+1/6)f}$). Peaks in the idle channel noise spectrum are compared against a smoothed average idle channel noise spectrum.

6.2.15 Double Talk Performance

6.2.15.0 General

During double talk the speech is mainly determined by 2 parameters: impairment caused by echo during double talk and level variation between single and double talk (attenuation range).

In order to guarantee sufficient quality under double talk conditions the Talker Echo Loudness Rating should be high and the attenuation inserted should be as low as possible. MGWs which do not allow double talk in any case should provide a good echo attenuation which is realized by a high attenuation range in this case.

The most important parameters determining the speech quality during double talk are (see Recommendations ITU-T P.340 [15] and P.502 [17]):

- Attenuation range in send direction during double talk $A_{H,S,dt}$.
- Attenuation range in receive direction during double talk $A_{H,R,dt}$.
- Echo attenuation during double talk.

The double talk performance may be highly influenced by the performance of the echo canceller, especially by the NLP implementation. The double talk performance should be checked by using the relevant echo paths as described in clause 6.2.18.1.

6.2.15.1 Attenuation Range in Send Direction during Double Talk $A_{H,S,dt}$

Requirement

Based on the level variation in send direction during double talk $A_{H,S,dt}$ the behaviour of the MGW can be classified according to Table 23.

For media gateway Type 1 is required for all selected echo paths (see clause 6.1.7).

Table 23

Category (according to Recommendation ITU-T P.340 [15])	1	2a	2b	2c	3
	<i>Full Duplex Capability</i>	<i>Partial Duplex Capability</i>			<i>No Duplex Capability</i>
$A_{H,S,dt}$ [dB]	≤ 3	≤ 6	≤ 9	≤ 12	> 12

The terminal shall have full duplex capability (category 1).

The requirement applies for nominal signal levels in send and receive directions as well as for the level combinations +6 dB (re. nominal level) in send/-6 dB (re. nominal level) in receive and +6 dB (re. nominal level) in receive/-6 dB (re. nominal level) in send.

In general Table 23 provides a quality classification of MGWs regarding double talk performance. However, this does not mean that a MGW which is category 1 based on the double talk performance is of high quality concerning the overall quality as well.

Measurement Method

The long conditioning sequence described in clause 7.3.7 of Recommendation ITU-T P.501 [16] shall be used for conditioning the MGW, with the female speaker in the send direction. The test signal to determine the attenuation range during double talk is the double talk speech sequence as defined in clause 7.3.5 of Recommendation ITU-T P.501 [16] as shown in figure 5. The competing speaker is always inserted as the double talk sequence sdt(t) either in send or receive and is used for analysis.

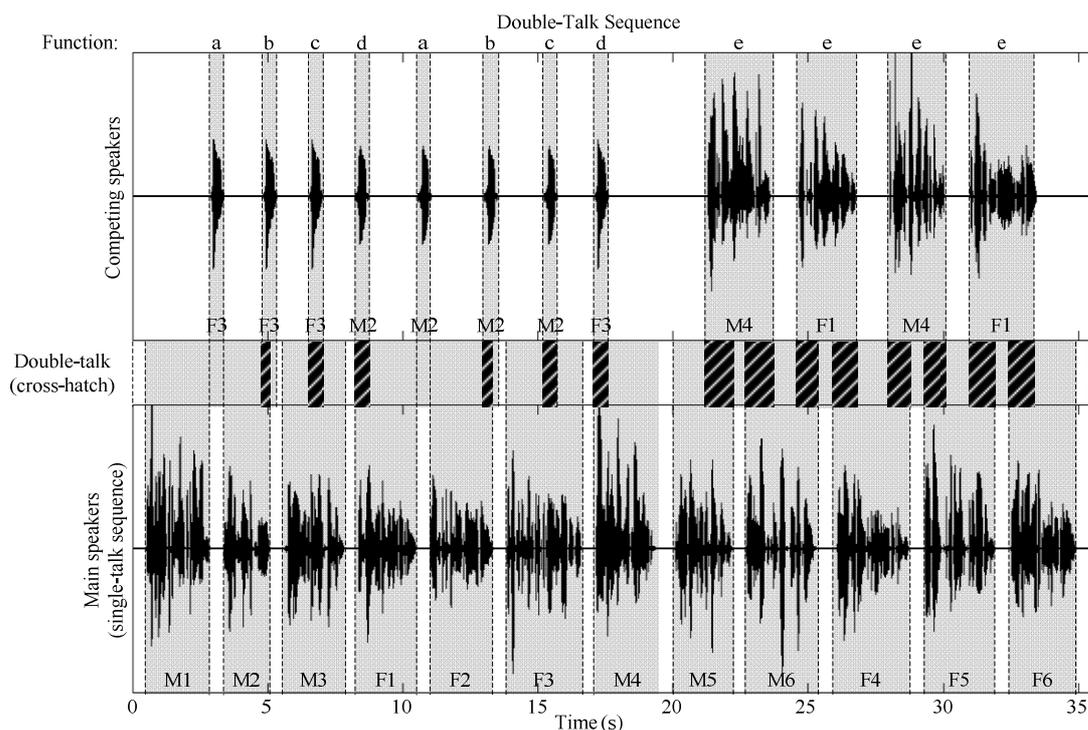


Figure 5: Double talk test sequence with overlapping speech sequences in send and receive direction

The test arrangement is according to clause 6.1.

The attenuation range during double talk is determined as described in Appendix III of Recommendation ITU-T P.502 [17]. The double talk performance is analysed for each word and sentence produced by the competing speaker. The requirement has to be met for each word and sentence produced by the competing speaker.

6.2.15.2 Attenuation Range in Receive Direction during Double Talk $A_{H,R,dt}$

Requirement

Based on the level variation in receive direction during double talk $A_{H,R,dt}$ the behaviour of the MGW can be classified according to Table 24.

For media gateway Type 1 is required for all selected echo paths (see clause 6.1.7).

Table 24

Category (according to Recommendation ITU-T P.340 [15])	1	2a	2b	2c	3
	Full Duplex Capability	Partial Duplex Capability			No Duplex Capability
$A_{H,R,dt}$ [dB]	≤ 3	≤ 5	≤ 8	≤ 10	> 10

The terminal shall have full duplex capability (category 1).

The requirement applies for nominal signal levels in send and receive directions as well as for the level combinations +6 dB (re. nominal level) in send/-6 dB (re. nominal level) in receive and +6 dB (re. nominal level) in receive/-6 dB (re. nominal level) in send.

In general Table 24 provides a quality classification of terminals regarding double talk performance. However, this does not mean that a terminal which is category 1 based on the double talk performance is of high quality concerning the overall quality as well.

Measurement Method

The long conditioning sequence described in clause 7.3.7 of Recommendation ITU-T P.501 [16] shall be used for conditioning the MGW, with the female speaker in the send direction. The test signal to determine the attenuation range during double talk is shown in figure 5. A sequence of speech signal is used which is inserted in parallel in send and receive direction. The test signals are synchronized in time at the acoustical interface. The delay of the test arrangement should be constant during the measurement.

The test arrangement is according to clause 6.1.

The attenuation range during double talk is determined as described in Appendix III of Recommendation ITU-T P.502 [17]. The double talk performance is analysed for each word and sentence produced by the competing speaker. The requirement has to be met for each word and sentence produced by the competing speaker.

6.2.15.3 Detection of Echo Components during Double Talk

Requirement

Echo Loss (EL) during double talk is the echo suppression provided by the MGW during double talk measured at the receive interface.

For media gateway Type 1 is required for all selected echo paths (see clause 6.1.7).

NOTE: The echo attenuation during double talk is based on the parameter Talker Echo Loudness Rating (TELRdt). It is assumed that the terminal at the opposite end of the connection provides nominal Loudness Rating (SLR + RLR = 10 dB).

Under these conditions the requirements given in Table 25 are applicable (more information can be found in annex A of the Recommendation ITU-T P.340 [15]).

Table 25

Category (according to Recommendation ITU-T P.340 [15])	1	2a	2b	2c	3
	<i>Full Duplex Capability</i>	<i>Partial Duplex Capability</i>			<i>No Duplex Capability</i>
Echo Loss [dB]	≥ 27	≥ 23	≥ 17	≥ 11	< 11

Measurement Method

The test arrangement is according to clause 6.1.

The double talk signal consists of a sequence of orthogonal signals which are realized by voice-like modulated sine waves spectrally shaped similar to speech. A detailed description can be found in Recommendation ITU-T P.501 [16].

The signals are fed simultaneously in send and receive direction. The level in send direction is -16 dBm0 for 4-wire MGW, -13 dBm for 2-wire home MGW and -16 dBm for 2 wire network MGW (nominal level), the level in receive direction is -16 dBm0 (nominal level). The settings for the signals are as follows.

Table 26: Parameters of the two Test Signals for Double Talk Measurement based on AM-FM modulated sine waves for narrowband MGW

Send Direction		Receive Direction	
$f_0^{(1)}$ [Hz]	$\pm\Delta f^{(1)}$ [Hz]	$f_0^{(2)}$ [Hz]	$\pm\Delta f^{(2)}$ [Hz]
250	±5	270	±5
500	±10	540	±10
750	±15	810	±15
1 000	±20	1 080	±20
1 250	±25	1 350	±25
1 500	±30	1 620	±30
1 750	±35	1 890	±35
2 000	±40	2 160	±35
2 250	±40	2 400	±35
2 500	±40	2 650	±35
2 750	±40	2 900	±35
3 000	±40	3 150	±35
3 250	±40	3 400	±35
3 500	±40	3 650	±35
3 750	±40	3 900	±35

NOTE: Parameters of the Shaping Filter:
 $f \geq 250$ Hz: Low Pass Filter, 5 dB/oct.

Table 27: Parameters of the two Test Signals for Double Talk Measurement based on AM-FM modulated sine waves for wideband MGW

Send Direction		Receive Direction	
$f_0^{(1)}$ [Hz]	$\pm\Delta f^{(1)}$ [Hz]	$f_0^{(2)}$ [Hz]	$\pm\Delta f^{(2)}$ [Hz]
125	±2,5	180	±2,5
250	±5	270	±5
500	±10	540	±10
750	±15	810	±15
1 000	±20	1 080	±20
1 250	±25	1 350	±25
1 500	±30	1 620	±30
1 750	±35	1 890	±35
2 000	±40	2 160	±35
2 250	±40	2 400	±35
2 500	±40	2 650	±35
2 750	±40	2 900	±35
3 000	±40	3 150	±35
3 250	±40	3 400	±35
3 500	±40	3 650	±35
3 750	±40	3 900	±35
4 000	±40	4 150	±35
4 250	±40	4 400	±35
4 500	±40	4 650	±35
4 750	±40	4 900	±35
5 000	±40	5 150	±35
5 250	±40	5 400	±35
5 500	±40	5 650	±35
5 750	±40	5 900	±35
6 000	±40	6 150	±35
6 250	±40	6 400	±35
6 500	±40	6 650	±35
6 750	±40	6 900	±35
7 000	±40		

NOTE: Parameters of the Shaping Filter:
 $f \geq 250$ Hz: Low Pass Filter, 5 dB/oct.

The test signal is measured at the receive interface. The measured signal consists of the double talk signal which was fed in at the send interface and the echo signal. The echo signal is filtered by comb filter using mid-frequencies and bandwidth according to the signal components of the signal in receive direction (see Recommendation ITU-T P.501 [16]). The filter will suppress frequency components of the double talk signal.

In each frequency band which is used in receive direction the echo attenuation can be measured separately. The requirement for category 1 is fulfilled if in any frequency band the echo signal is either below the signal noise or below the required limit. If echo components are detectable, the classification is based on Table 25. The echo attenuation is to be achieved for **each individual frequency band** according to the different categories.

The measurement is done with every selected echo path after training of the echo canceller (see clause 6.1.7).

6.2.15.4 Minimum Activation Level and Sensitivity of Double Talk Detection

For further study.

6.2.16 Switching Characteristics

6.2.16.0 General

Additional requirements may be needed in order to further investigate the effect of NLP implementations on the users' perception of speech quality.

6.2.16.1 Activation in Send Direction

The activation in send direction is mainly determined by the built-up time $T_{r,S,min}$ and the minimum activation level ($L_{S,min}$). The minimum activation level is the level required to remove the inserted attenuation in send direction during idle mode. The built-up time is determined for the test signal burst which is applied with the minimum activation level.

The activation level described in the following is always referred to the test signal level at the send interface of the MGW.

Requirements

- For 4-wire MGW the minimum activation level $L_{S,min}$ shall be ≤ -32 dBm0.
- For 2-wire home MGW the minimum activation level $L_{S,min}$ shall be ≤ -29 dBm.
- For 2-wire network MGW the minimum activation level $L_{S,min}$ shall be ≤ -32 dBm.

The built-up time $T_{r,S,min}$ (measured with minimum activation level) should be ≤ 15 ms.

Measurement Method

The test signal is the short words for activation sequence described in clause 7.3.4 of Recommendation ITU-T P.501 [16] with increasing level for each single word.

The settings of the test signal are as follows.

Table 28

	Single word Duration/ Pause Duration	Level of the first single word	Level Difference between two Periods of the Test Signal
Single word to determine switching characteristic in send direction	~600 ms/ ~400 ms	-35 dBm0 (4-wire MGW) -32 dBm (2-wire home MGW) -35 dBm (2-wire network MGW)	1 dB
NOTE: The signal level is determined for each utterance individually according to Recommendation ITU-T P.56 [18].			

It is assumed that the pause length of about 400 ms is longer than the hang-over time so that the test object is back to idle mode after each single word.

The test arrangement is described in clause 6.1.

The level of the transmitted signal is measured at the receive interface. The test signal is filtered by the transfer function of the test object. The measured signal level is referred to the filtered test signal level and displayed versus time. The levels are calculated from the time domain using an integration time of 5 ms.

The minimum activation level is determined from the single word which indicates the first activation of the test object. The time between the beginning of the single word and the complete activation of the test object is measured.

6.2.16.2 Activation in Receive Direction

The activation in receive direction is mainly determined by the built-up time $T_{r,S,min}$ and the minimum activation level ($L_{S,min}$). The minimum activation level is the level required to remove the inserted attenuation in receive direction during idle mode. The built-up time is determined for the test signal burst which is applied with the minimum activation level.

The activation level described in the following is always referred to the test signal level at the receive interface of the MGW.

Requirements

The minimum activation level $L_{S,min}$ shall be ≤ -32 dBm0.

The built-up time $T_{r,S,min}$ (measured with minimum activation level) should be ≤ 15 ms.

Measurement Method

The test signal is the short words for activation sequence described in clause 7.3.4 of Recommendation ITU-T P.501 [16] with increasing level for each single word.

The settings of the test signal are as follows.

Table 29

	Single word Duration/ Pause Duration	Level of the first single word	Level Difference between two Periods of the Test Signal
Single word to determine switching characteristic in Receive Direction	~600 ms/ ~400 ms	-35 dBm0	1 dB
NOTE: The signal level is determined for each utterance individually according to Recommendation ITU-T P.56 [18].			

The terminal shall have full duplex capability (category 1).

The requirement applies for nominal signal levels in send and receive directions as well as for the level combinations +6 dB (re. nominal level) in send/-6 dB (re. nominal level) in receive and +6 dB (re. nominal level) in receive/-6 dB (re. nominal level) in send.

It is assumed that the pause length of about 400 ms is longer than the hang-over time so that the test object is back to idle mode after each single word.

The test arrangement is described in clause 6.1.

The level of the transmitted signal is measured at the send interface. The test signal is filtered by the transfer function of the test object. The measured signal level is referred to the filtered test signal level and displayed versus time. The levels are calculated from the time domain using an integration time of 5 ms.

The minimum activation level is determined from the single word which indicates the first activation of the test object. The time between the beginning of the single word and the complete activation of the test object is measured.

6.2.16.3 Silence Suppression and Comfort Noise Generation

To provide optimum speech quality as perceived by end users, silence suppression should be avoided in the transmission path.

Requirement and measurement method for further study.

6.2.17 Background Noise Performance

6.2.17.0 General

To provide optimum speech quality as perceived by end users, noise cancelling should be avoided in the transmission path. The preferred location for the implementation of such feature is the terminal.

However, the requirement and measurement methods exposed below are meant to ensure good performance of the media gateway echo canceller in the presence of background noise.

6.2.17.1 Performance in Send Direction in the Presence of Background Noise

Requirement

The level of comfort noise, if implemented, shall be within a range of +2 and -5 dB compared to the original (transmitted) background noise. The noise level is calculated with psophometric weighting for narrowband MGW and with A-weighting for wideband MGW.

NOTE 1: It is advisable that the comfort noise matches the original signal as good as possible (from a perceptual point of view).

NOTE 2: Input for further specification necessary (e.g. on temporal matching).

The spectral difference between comfort noise and original (transmitted) background noise shall be within the mask given through straight lines between the breaking points on a logarithmic (frequency) - linear (dB sensitivity) scale as given in Table 30 and Table 31.

Table 30: Requirements for Spectral Adjustment of Comfort Noise (Mask) for narrowband MGW

Frequency	Upper Limit	Lower Limit
200 Hz	12 dB	-12 dB
800 Hz	12 dB	-12 dB
800 Hz	10 dB	-10 dB
2 000 Hz	10 dB	-10 dB
2 000 Hz	6 dB	-6 dB
4 000 Hz	6 dB	-6 dB

NOTE: All sensitivity values are expressed in dB on an arbitrary scale.

Table 31: Requirements for Spectral Adjustment of Comfort Noise (Mask) for wideband MGW

Frequency	Upper Limit	Lower Limit
200 Hz	12 dB	-12 dB
800 Hz	12 dB	-12 dB
800 Hz	10 dB	-10 dB
2 000 Hz	10 dB	-10 dB
2 000 Hz	6 dB	-6 dB
4 000 Hz	6 dB	-6 dB
8 000 Hz	6 dB	-6 dB

NOTE: All sensitivity values are expressed in dB on an arbitrary scale.

Measurement Method

The background noise simulation as described in clause 6.1 is used.

First the background noise transmitted in send is recorded at the POI for a period of at least 20 s.

In a second step a test signal is applied in receive direction consisting of an initial pause of 10 s and a periodical repetition of the female speaker signal of the short conditioning sequence described in clause 7.3.7 of Recommendation ITU-T P.501 [16] in receive direction (duration 10 s) with nominal level to enable comfort noise injection simultaneously with the background noise. For the measurement the background noise sequence has to be started at the same point as it was started in the previous measurement.

The transmitted signal is recorded in send direction at the POI.

The power density spectra measured in send direction without far end speech simulation averaged between 10 s and 20 s is referred to the power density spectrum measured in send direction determined during the period with far end speech in receive direction averaged between 10 s and 20 s. Level and spectral differences between both power density spectra are analysed and compared to the requirements.

6.2.17.2 Quality of Speech with Background Noise

Generally this measurement is performed in send direction. For IP-to-IP MGW the measurement is made in both directions.

Requirement

Speech Quality for narrowband and wideband systems can be tested based on ETSI EG 202 396-3 [i.8].

For narrowband MGW:

- For the background noises defined in clause 6.1 the following requirements apply:
 - N-MOS-LQOn $\geq 3,5$
 - S-MOS-LQOn $\geq 3,5$
 - G-MOS-LQOn $\geq 3,5$

For wideband MGW:

- For the background noises defined in clause 6.1 the following requirements apply:
 - N-MOS-LQOw $\geq 3,5$
 - S-MOS-LQOw $\geq 3,5$
 - G-MOS-LQOw $\geq 3,5$

NOTE: It is recommended to test the MGW performance with other types of background noises if the terminal is likely to be exposed to other noises than specified in clause 6.1.

Measurement Method

The speech plus background noise signal generated as described in clause 6.1 is used. The test arrangement is described in clause 6.1.

The background noise should be applied for at least 5 s in order to adapt noise reduction algorithms in advance to the test.

The near end speech signal consists of 8 sentences of speech (2 male and 2 female talkers, 2 sentences each). Appropriate speech samples can be found in Recommendation ITU-T P.501 [16]. The preferred language for narrowband MGW is English since the objective method was validated with English language in narrowband. The preferred language for wideband MGW is French since the objective method was validated with French language in wideband. The speech signal level is 3 dB higher than nominal.

Three signals are required for the tests:

- 1) The clean speech signal is used as the undisturbed reference (see ETSI EG 202 396-3 [i.8]).
- 2) The previously recorded speech plus undisturbed background noise signal (see clause 6.1.6) is used as the unprocessed speech plus noise reference.
- 3) The send signal, recorded at the output of the gateway send path is used as the degraded signal.

N-MOS-LQOn, S-MOS-LQOn and G-MOS-LQOn are calculated per sentence as described in ETSI EG 202 396-3 [i.8].

N-MOS-LQOw, S-MOS-LQOw and G-MOS-LQOw are calculated per sentence as described in ETSI EG 202 396-3 [i.8].

The resulting N-MOS-LQO-, S-MOS-LQO- and G-MOS-LQO-values are averaged over all 8 sentences.

6.2.17.3 Quality of Background Noise Transmission (with Far End Speech)

Requirement

The test is carried out applying a speech signal in receive direction. During and after the end of the speech signal the signal level in send direction should not vary more than 10 dB (during transition to transmission of background noise without far end speech). The measurement is conducted for all types of background noise as defined in clause 6.1 and with all selected echo paths (see clause 6.1.7).

Measurement Method

The test arrangement is according to clause 6.1.

The background noises are generated as described in clause 6.1.

First the measurement is conducted without inserting the signal at the far end. At least 10 s of noise is analysed. The background signal level versus time is calculated using a time constant of 35 ms. This is the reference signal.

In a second step the same measurement is conducted but with inserting the speech signal at the far end. The exactly identical background noise signal is applied. The background noise signal shall start at the same point in time which was used for the measurement without far end signal. The background noise should be applied for at least 5 s in order to allow adaptation of the noise reduction algorithms. After at least 5 s a series of the female speaker signal of the short conditioning sequence described in clause 7.3.7 of Recommendation ITU-T P.501 [16] is applied in receive direction with duration of at least 10 s. The test signal level is -16 dBm0 at the receive interface.

The send signal is recorded at the output of the gateway send path. The test signal level versus time is calculated using a time constant of 35 ms.

The level variation in receive direction is determined during the time interval when the speech signal is applied and after it stops. The level difference is determined from the difference of the recorded signal levels vs. time between reference signal and the signal measured with far end signal.

The measurement is done for each defined background noise type with all of the selected echo paths after training of the echo canceller (see clause 6.1.7).

6.2.17.4 Quality of Background Noise Transmission (with Near End Speech)

Requirement

The test is carried out applying a speech signal in send direction. During and after the end of the speech signal bursts the signal level in send direction should not vary more than 10 dB.

Measurement Method

The test arrangement is according to clause 6.1.

The background noises are generated as described in clause 6.1. The background noise should be applied for at least 5 s in order to allow adaptation of the noise reduction algorithms.

As near end speech signal a series of the short word for activation described in clause 7.3.4 of Recommendation ITU-T P.501 [16] is applied in send direction with duration of ≥ 4 speech bursts. The speech signal plus the background noise are pre-recorded as described in clause 6.1. The speech signal level is 3 dB higher than nominal.

The send signal is recorded at the electrical reference point. The test signal level versus time is calculated using a time constant of 35 ms.

First the measurement is conducted without inserting the signal at the near end. The signal level is analysed vs. time. In a second step the same measurement is conducted but with inserting the speech signal at the near end. The level variation is determined by the difference between the background noise signal level without inserting the speech signal and the maximum level of the noise signal during and after the speech bursts in send direction.

6.2.18 Quality of Echo Cancellation

6.2.18.0 General

Echo cancellation is not needed in every case. If not needed (e.g. terminal delivering enough echo cancellation), it should be possible to switch off the echo canceller by the means of signalling (e.g. DECT depending on terminal capabilities).

In IP-to-IP MGWs there should not be any echo cancellation.

6.2.18.1 Echo Performance According to Recommendation ITU-T G.168

The tests of the echo canceller cancellation performance are performed according to Recommendation ITU-T G.168 [23].

The following Tests of Recommendation ITU-T G.168 [23] apply:

- Test 2A
- Test 2C (only applicable, it is possible to inhibit the adaption of the echo canceller)
- Test 9

For all tests the NLP remains enabled, tests requiring the deactivation of the NLP are not applicable.

The performance has to be checked with all selected echo paths (see clause 6.1.7).

NOTE: For many of these measurements the H-register of the echo canceller should be reset before the actual measurement. Since this is often not feasible because e.g. the echo canceller is not accessible, it is (in conformance to Recommendation ITU-T G.168 [23]) possible to train the echo canceller with an open echo path before the actual measurement.

Requirement

See Recommendation ITU-T G.168 [23].

Measurement Method

The test arrangement is according to clause 6.1.

See Recommendation ITU-T G.168 [23].

6.2.18.2 Terminal Coupling Loss (TCL_w, NB)

Requirement

The TCL_w shall be ≥ 46 dB.

NOTE: A TCL_w ≥ 50 dB is recommended as a performance objective. Depending on the idle channel noise in the sending direction, it may not always be possible to measure an echo loss ≥ 50 dB.

Measurement Method

The test arrangement is according to clause 6.1.

The attenuation from electrical reference point input to electrical reference point output shall be measured using the compressed real speech signal described in clause 7.3.3 of Recommendation ITU-T P.501 [16]. The signal level shall be -10 dBm0.

The TCLw is calculated according to Recommendation ITU-T G.122 [8], clause B.4 (trapezoidal rule). 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. 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).

The measurement is done with every selected echo path.

6.2.18.3 Terminal Coupling Loss (TCL, WB)

Requirement

The TCL measured as unweighted Echo Loss shall be ≥ 46 dB for all selected echo paths (see clause 6.1.7).

Measurement Method

The test arrangement is according to clause 6.1.

The attenuation from electrical reference point input to electrical reference point output shall be measured using the compressed real speech signal described in clause 7.3.3 of Recommendation ITU-T P.501 [16]. The signal level shall be -10 dBm0.

TCL is calculated as unweighted echo loss from 100 Hz to 8 kHz. 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. 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). For the measurement, a time window has to be applied which is adapted to the duration of the actual test signal. The echo loss is calculated by the equations:

$$L_e = C - 10 \log_{10} \sum_{i=1}^N (A_i + A_{i-1}) (\log_{10} f_i - \log_{10} f_{i-1}) \quad (1)$$

and

$$C = 10 \log_{10} (2 (\log_{10} f_N - \log_{10} f_0)) \quad (2)$$

where

- A_0 is the output/input power ratio at frequency $f_0 = 100$ Hz;
- A_i the ratio at frequency f_i ; and
- A_N the ratio at frequency $f_N = 8\,000$ Hz.

The above equation is a generalized form of the equation defined in Recommendation ITU-T G.122 [8], clause B.4, for calculating echo loss based on tabulated data, which allows the calculation of echo loss within any frequency range between f_0 and f_N .

The measurement is done with every selected echo path.

6.2.18.4 Temporal Echo Effects

Requirement

This test is intended to verify that the system will maintain sufficient echo attenuation during single talk. The measured echo attenuation during single talk should not decrease by more than 6 dB from the maximum TCL measured during the test for all selected echo paths (see clause 6.1.7).

Measurement Method

The test arrangement is according to clause 6.1.

The test signal consists of periodically repeated Composite Source Signal according to Recommendation ITU-T P.501 [16] with an average level of -5 dBm0 as well as an average level of -25 dBm0. The echo signal is analysed during a period of at least 2,8 s which represents 8 periods of the CS signal. The integration time for the level analysis shall be 35 ms, the analysis is referred to the level analysis of the reference signal. The TCL variation is compared to the maximum TCL achieved in the test.

The measurement result is displayed as attenuation vs. time. The exact synchronization between input and output signal has to be guaranteed.

NOTE 1: In addition tests with more speech like signals as the British-English single talk sequence described in clause 7.3.2 of Recommendation ITU-T P.501 [16] should be made to see time variant behaviour of EC. However, for such tests the simple broadband attenuation based test principle as described above cannot be applied due to the time varying spectral content of the speech like signals.

NOTE 2: The analysis is conducted only during the active signal part, the pauses between the Composite Source Signals are not analysed. The analysis time is reduced by the integration time of the level analysis (35 ms).

The measurement is done with every selected echo path after training of the echo canceller (see clause 6.1.7).

6.2.18.5 Spectral Echo Attenuation

Requirement

The echo attenuation vs. frequency shall be below the tolerance mask given in Table 32 and Table 33 for all selected echo paths (see clause 6.1.7).

Table 32: Echo attenuation limits for narrowband MGW

Frequency	Limit
100 Hz	-20 dB
200 Hz	-30 dB
300 Hz	-38 dB
800 Hz	-34 dB
1 500 Hz	-33 dB
2 600 Hz	-24 dB
4 000 Hz	-24 dB

NOTE 1: All sensitivity values are expressed in dB on an arbitrary scale.
NOTE 2: The limit at intermediate frequencies lies on a straight line drawn between the given values on a log (frequency) - linear (dB) scale.

Table 33: Echo attenuation limits for wideband MGW

Frequency	Limit
100 Hz	-41 dB
1 300 Hz	-41 dB
3 450 Hz	-46 dB
5 200 Hz	-46 dB
7 500 Hz	-37 dB
8 000 Hz	-37 dB

NOTE 1: All sensitivity values are expressed in dB on an arbitrary scale.
NOTE 2: The limit at intermediate frequencies lies on a straight line drawn between the given values on a log (frequency) - linear (dB) scale.

During the measurement it should be ensured that the measured signal is really the echo signal and not the Comfort Noise which possibly may be inserted in send direction in order to mask the echo signal.

Measurement Method

The test arrangement is according to clause 6.1.

Before the actual measurement a training sequence is fed in consisting of the British-English single talk sequence described in clause 7.3.2 of Recommendation ITU-T P.501 [16] is applied. The level of the training sequence is -16 dBm0.

The test signal consists of a periodically repeated Composite Source Signal. The measurement is carried out under steady-state conditions. The average test signal level is -16 dBm0, averaged over the complete test signal. 4 CS signals including the pauses are used for the measurement which results in a test sequence length of 1,4 s. The power density spectrum of the measured echo signal is referred to the power density spectrum of the original test signal. The analysis is conducted using FFT analysis with 8 k points (48 kHz sampling rate, Hanning window).

The spectral echo attenuation is analysed in the frequency domain in dB.

The measurement including the training sequence is done with every selected echo path.

6.2.18.6 Occurrence of Artefacts

For further study.

6.2.18.7 Variable Echo Paths

For further study.

6.2.19 Variant Impairments

6.2.19.1 Clock Accuracy Send

Requirement

The clock accuracy in send direction between the MGW and the IP reference interface shall be less than 40 ppm under ideal network conditions.

NOTE: The clock accuracy does not cover all possible network configurations. Especially it is not sufficient for data transmission or distributed TDM PBX where synchronization is required.

Measurement Method

A sequence of CS signals (active signal length = 250 ms) is repeated for 120 s in order to analyse clock accuracy and any other time-variant delay. The pause length between two CS bursts is 100 ms and 1,2 s after every fourth burst in order to simulate a speech pause, which may lead to buffer adjustments. The test signal level at the MGW interface is -16 dBm0 for 4-wire MGW, -13 dBm for 2-wire home MGW and -16 dBm for 2-wire network MGW.

A cross correlation analysis versus time is carried out over the whole 120 s sequence between the received and the original test signal. The duration of the measurement (120 s) is indicated on the x-axis, the result of the cross correlation analysis (delay) is plotted on the y-axis.

The resulting clock accuracy within an analysis time range of at least 60 s is calculated as follows:

$$\text{clock accuracy [ppm]} = \frac{\text{delay drift [s]}}{\text{analysis duration [s]}} \cdot 1 \cdot 10^6 \quad (3)$$

6.2.19.2 Clock Accuracy Receive

Requirement

The clock accuracy in receive direction between the IP reference interface and the MGW shall be less than 40 ppm under ideal network conditions.

Measurement Method

A sequence of CS signals (active signal length = 250 ms) is repeated for 120 s in order to analyse clock accuracy and any other time-variant delay. The pause length between two CS bursts is 100 ms and 1,2 s after every fourth burst in order to simulate a speech pause, which may lead to buffer adjustments. The test signal level at the IP reference interface is -16 dBm0.

A cross correlation analysis versus time is carried out over the whole 120 s sequence between the received and the original test signal. The duration of the measurement (120 s) is indicated on the x-axis, the result of the cross correlation analysis (delay) is plotted on the y-axis.

The resulting clock accuracy within an analysis time range of at least 60 s is calculated as follows:

$$\text{clock accuracy [ppm]} = \frac{\text{delay drift [s]}}{\text{analysis duration [s]}} \cdot 1 \cdot 10^6 \quad (4)$$

6.2.19.3 Send Packet Delay Variation

Requirement

The measured maximum delay variation of RTP packets in send direction of the MGW under test should be less than 1 ms.

NOTE: Any delay variation of RTP packets introduced in send direction will lead to potentially increased delay due to increased de-jitter buffer at the far end terminal.

Measurement Method

The RTP data stream in send direction should be monitored with a tap or a switch providing a monitoring port, positioned at the location of the network impairment simulator (see clause 6.1). The test arrangement is according to clause 6.1.

The monitoring time should be 60 s. A signal like the British-English single talk sequence described in clause 7.3.2 of Recommendation ITU-T P.501 [16] is played back in send direction using a nominal network level of -16 dBm0 for 4-wire MGW, -13 dBm for 2-wire home MGW and -16 dBm0 for 2-wire network MGW. This speech signal is only necessary to make sure, RTP is played out, even in the case VAD is active.

The delay variation for each packet $D(i)$ is evaluated according to IETF RFC 3550 [29]:

$$\begin{aligned} d(i) &= \Delta t_{\text{eff}(i)} - \Delta t_{\text{exp}(i)} \\ D(i) &= (15 \times D(i-1) + |d(i)|) / 16 \end{aligned} \quad (5)$$

With:

- $\Delta t_{\text{exp}(i)}$ = the expected time between packet i and packet $i-1$; and
- $\Delta t_{\text{eff}(i)}$ = the effective time between packet i and packet $i-1$.

Maximum delay variation = MAX($D(i)$).

6.2.20 Immunity to DTMF False Detection in Send Direction

On the TDM interface of a gateway, the incoming audio signal may be analysed in order to detect DTMF tones to be transmitted in a separate way than speech, on the IP network (see IETF RFC 2833 [i.2], IETF RFC 4733 [i.3]). It can occur that some parts of speech signal are analysed by the gateway as DTMF and therefore processed as such. The result of it is that the far end listener will hear a tone instead of a syllable of a word. This shall be avoided.

Other aspects of DTMF transmission performance are not impacting speech quality.

Requirement

No more than 5 false detections of DTMF shall be reported on duration of 30 minutes.

Measurement Method

The test arrangement is according to clause 6.1.

The test signal to be used for the measurements shall be composed of speech with a relatively high speech activity ratio. The resulting signal, captured in the IP network, can be analysed in real time or after recording. The number of DTMF tones present (i.e. detected during this analysis, in conformance with DTMF signals specifications, in terms of frequencies, levels and durations) in this signal is counted and reported.

The test signal used is found in annex B.

6.2.21 Send and receive delay - round trip delay

The roundtrip delay of a MGW is defined as the sum of send and receive delays. In the following clauses the calculation of the requirements for send and receive delay are explained. For a telecommunication connection, only the roundtrip delay can be experienced. For this reason, also the requirement for MGWs is given only for the roundtrip delay. As long as the measured roundtrip delay fulfils the requirements, send or receive delays may be above the theoretical requirements.

Requirement

It is recognized that the end to end delay should be as small as possible in order to ensure high quality of the communication.

The roundtrip delay of the MGW T_{rtd} (sum of receive and send delay) shall be less than 100 ms. (category B in Recommendation ITU-T P-1010 [28]). From the user's perspective, a value less than 50 ms (category A in Recommendation ITU-T P-1010 [28]) is preferred.

NOTE 1: The limit for the roundtrip delay T_{rtd} of the MGW is derived from the sum of the send and receive delay limits.

NOTE 2: This requirement is based on the lowest possible delay values which can be expected under ideal network conditions. Caution should be exercised to ensure that the terminal is operated under optimum conditions in order to avoid adverse effects, e.g. network conditions, settings and memory effects of the MGW jitter buffer.

Measurement Method

The test signal to be used for the measurements shall be a Composite Source Signal (CSS) as described in Recommendation ITU-T P.501 [16]. The test signal consists of the voiced part as described in Recommendation ITU-T P.501 [16] followed by a pseudo random noise sequence with a periodicity of minimum 500 ms. The test signal level shall be -16 dBm0, measured at the electrical test point. The test signal level is averaged over the complete test signal sequence.

NOTE 3: If the expected delay is higher than 500 ms a pseudo random sequence with a higher periodicity should be used.

The delay is calculated using the cross correlation function between the signal at the output and the signal at the input. The cross correlation analysis has to be chosen in such a way that the maximum delay of 500 ms can be analysed. The measurement is corrected by the delay introduced by the test equipment.

The delay is expressed in ms, determined from the maximum of the cross correlation function.

NOTE 4: Delay may be time variant. Therefore constant monitoring of the actual delay may be required when evaluating the range of delay which can be observed in a given connection. The test setup should take into account either real network conditions or the tools needed to simulate typical causes for time variant delay (e.g. packet loss) during the measurement period.

The roundtrip delay of the MGW is the sum of send and receive delays minus the roundtrip delay of the measurement equipment and (if applicable) the network.

6.3 Codec Specific Requirements

6.3.1 Objective Listening Speech Quality MOS-LQO in Send direction

The listening speech quality tests are conducted under clean network conditions.

Requirements

For narrowband MGW the requirements for the listening speech quality are as follows.

Table 34

Speech coder	MOS-LQOF
Recommendation ITU-T G.711 [9]	> 3,4
Recommendation ITU-T G.729 [12]	> 2,9
Recommendation ITU-T G.723.1 [10]	> 2,7
Recommendation ITU-T G.726 @ 32 kbit/s [11]	> 3,1
GSM EFR [1] and AMR @ 12,2 kbit/s [2]	> 3,2
Recommendation ITU-T G.729.1 @ 8 kbit/s [13]	> 2,9

NOTE 1: The use of the codecs G.723.1 [10], G.729 [12] and G.729.1 [13] is not recommended due to low quality.

For wideband MGW the requirements for the listening speech quality are as follows.

Table 35

Speech coder	MOS-LQOF
Recommendation ITU-T G.722 [25]	> 4,0
G.729.1 @ 32 kbit/s [13]	> 4,2
G.722.1 [26]	> 4,0
L16-256 [31]	> 4,3
AMR-WB [27]	> 4,1

Measurement Method

For narrowband and wideband media gateways MOS-LQOF is measured using Recommendation ITU-T P.863.1 [14] in fullband mode.

The test signal to be used for the measurements shall be four sentence pairs (male/female) fulfilling the requirements of Recommendation ITU-T P.863.1 [14]. The 4 sentence pairs are taken from Recommendation ITU-T P.501 [16]. The test signal level is averaged over all sentence pairs (4 sentence pairs). The measurement is done four times, every time using another pair of the speech sentences. The result of the measurement is the averaged value of all four measurements.

NOTE 2: With P.863 narrowband mediagateways can be measured in narrowband mode as well as in fullband mode. If backwards comparability of results is needed (e.g. with P.862 or older subjective test results), narrowband mode should be chosen.

NOTE 3: For the use of P.863 the following applies (see Recommendation ITU-T P.863.1 [14]):

- Fullband Context (MOS-LQOF):
 - Reference Signal fullband flat filtered 20 Hz to 20 kHz.
 - Narrowband system: Test Signal Narrowband IRS Send filtered.
 - Wideband system: Test Signal Wideband flat low pass filtered 7,8 kHz.

6.3.2 Objective Listening Quality MOS-LQO in Receive direction

The listening speech quality tests are conducted under clean network conditions as well as with network impairments simulated. In addition to the listening speech quality tests the delay is measured.

Requirements

The requirements for the listening speech quality and the delay under clean network conditions are as follows.

Table 36

Speech coder	MOS-LQOF
Recommendation ITU-T G.711 [9]	> 3,4
Recommendation ITU-T G.729 [12]	> 2,9
Recommendation ITU-T G.723.1 [10]	> 2,7
Recommendation ITU-T G.726 @ 32 kbit/s [11]	> 3,1
GSM EFR [1] and AMR @ 12,2 kbit/s [2]	> 3,2
Recommendation ITU-T G.729.1 @ 8 kbit/s [13]	> 2,9

NOTE 1: The use of the codecs G.723.1 [10], G.729 [12] and G.729.1 [13] is not recommended due to low quality.

For wideband MGW the requirements for the listening speech quality are as follows.

Table 37

Speech coder	MOS-LQOF
Recommendation ITU-T G.722 [25]	> 4,0
G.729.1 @ 32 kbit/s [13]	> 4,2
G.722.1 [26]	> 4,0
L16-256 [31]	> 4,3
AMR-WB [27]	> 4,1

The requirements for the listening speech quality and the delay under impaired network conditions are as follows.

The roundtrip delay of the VoIP terminal T_{rtd} (sum of receive and send delay) shall be less than delay category B. From the users' perspective, a value less than delay category A is preferred (for this, $T_{\text{rtd}}^{\text{clean}}$ needs to fulfil the requirement for delay category A in clause 6.2.21).

Table 38: Requirements for Recommendation ITU-T G.711 [9] speech codecs

Condition	MOS-LQOF	Delay category A	Delay category B
0	> 3,3	$< T_{\text{rtd}}^{\text{clean}} + 5 \text{ ms}$	< 100 ms
1	> 3,4	$< T_{\text{rtd}}^{\text{clean}} + 5 \text{ ms}$	< 100 ms
2	> 3,1	$< T_{\text{rtd}}^{\text{clean}} + 5 \text{ ms}$	< 100 ms
3	> 3,1	$< T_{\text{rtd}}^{\text{clean}} + 5 \text{ ms}$	< 100 ms
4	> 3,1	$< T_{\text{rtd}}^{\text{clean}} + 50 \text{ ms}$	< 150 ms
5	> 3,1	$< T_{\text{rtd}}^{\text{clean}} + 50 \text{ ms}$	< 150 ms
6	> 3,1	$< T_{\text{rtd}}^{\text{clean}} + 50 \text{ ms}$	< 150 ms
7	> 3,1	$< T_{\text{rtd}}^{\text{clean}} + 50 \text{ ms}$	< 150 ms

Table 39: Requirements for Recommendation ITU-T G.729 [12] speech codecs

Condition	MOS-LQOF	Delay category A	Delay category B
1	> 2,9	< $T_{\text{rtd}}^{\text{clean}} + 5 \text{ ms}$	< 100 ms
2	> 2,8	< $T_{\text{rtd}}^{\text{clean}} + 5 \text{ ms}$	< 100 ms
3	> 2,3	< $T_{\text{rtd}}^{\text{clean}} + 5 \text{ ms}$	< 100 ms
4	> 2,8	< $T_{\text{rtd}}^{\text{clean}} + 50 \text{ ms}$	< 150 ms
5	> 2,8	< $T_{\text{rtd}}^{\text{clean}} + 50 \text{ ms}$	< 150 ms
6	> 2,8	< $T_{\text{rtd}}^{\text{clean}} + 50 \text{ ms}$	< 150 ms
7	> 2,8	< $T_{\text{rtd}}^{\text{clean}} + 50 \text{ ms}$	< 150 ms

Table 40: Requirements for Recommendation ITU-T G.723.1 [10] speech codecs

Condition	MOS-LQOF	Delay category A	Delay category B
1	> 2,7	< $T_{\text{rtd}}^{\text{clean}} + 5 \text{ ms}$	< 100 ms
2	> 2,5	< $T_{\text{rtd}}^{\text{clean}} + 5 \text{ ms}$	< 100 ms
3	> 2,5	< $T_{\text{rtd}}^{\text{clean}} + 5 \text{ ms}$	< 100 ms
4	> 2,5	< $T_{\text{rtd}}^{\text{clean}} + 50 \text{ ms}$	< 150 ms
5	> 2,4	< $T_{\text{rtd}}^{\text{clean}} + 50 \text{ ms}$	< 150 ms
6	> 2,5	< $T_{\text{rtd}}^{\text{clean}} + 50 \text{ ms}$	< 150 ms
7	> 2,4	< $T_{\text{rtd}}^{\text{clean}} + 50 \text{ ms}$	< 150 ms

Table 41: Requirements for Recommendation ITU-T G.722 [25] speech codecs

Condition	MOS-LQOF	Delay category A	Delay category B
1	> 4,0	< $T_{\text{rtd}}^{\text{clean}} + 5 \text{ ms}$	< 100 ms
2	> 3,8	< $T_{\text{rtd}}^{\text{clean}} + 5 \text{ ms}$	< 100 ms
3	> 3,6	< $T_{\text{rtd}}^{\text{clean}} + 5 \text{ ms}$	< 100 ms
4	> 3,8	< $T_{\text{rtd}}^{\text{clean}} + 50 \text{ ms}$	< 150 ms
5	> 3,8	< $T_{\text{rtd}}^{\text{clean}} + 50 \text{ ms}$	< 150 ms
6	> 3,8	< $T_{\text{rtd}}^{\text{clean}} + 50 \text{ ms}$	< 150 ms
7	> 3,8	< $T_{\text{rtd}}^{\text{clean}} + 50 \text{ ms}$	< 150 ms

Table 42: Requirements for Recommendation ITU-T G.722.1 [26] speech codecs

Condition	MOS-LQOF	Delay category A	Delay category B
1	> 4,0	< $T_{\text{rtd}}^{\text{clean}} + 5 \text{ ms}$	< 100 ms
2	> 3,8	< $T_{\text{rtd}}^{\text{clean}} + 5 \text{ ms}$	< 100 ms
3	> 3,8	< $T_{\text{rtd}}^{\text{clean}} + 5 \text{ ms}$	< 100 ms
4	> 3,8	< $T_{\text{rtd}}^{\text{clean}} + 50 \text{ ms}$	< 150 ms
5	> 3,8	< $T_{\text{rtd}}^{\text{clean}} + 50 \text{ ms}$	< 150 ms
6	> 3,8	< $T_{\text{rtd}}^{\text{clean}} + 50 \text{ ms}$	< 150 ms
7	> 3,8	< $T_{\text{rtd}}^{\text{clean}} + 50 \text{ ms}$	< 150 ms

NOTE 2: The delay requirements for conditions with network impairments are based on the measured roundtrip delay of the MGW in the absence of network impairments $T_{\text{rtd}}^{\text{clean}}$ (see clause 6.2.21). A small additional tolerance takes into account the variable behaviour of the delay.

Measurement Method

For narrowband and wideband media gateways MOS-LQOF is measured using Recommendation ITU-T P.863 [24] in fullband mode.

The test signals to be used for the measurements shall be four sentence pairs (male/female) fulfilling the requirements of Recommendation ITU-T P.863.1 [14]. The 4 sentence pairs are taken from Recommendation ITU-T P.501 [16]. The test signal level is averaged over all sentence pairs (4 sentence pairs). The measurement is done four times, every time using another pair of the speech sentences. The result of the measurement is the averaged value of all four measurement results.

NOTE 3: With P.863 narrowband mediagateways can be measured in narrowband mode as well as in fullband mode. If backwards comparability of results is needed (e.g. with P.862 or older subjective test results), narrowband mode should be chosen.

NOTE 4: For the use of P.863 the following applies (see Recommendation ITU-T P.863.1 [14]):

- Fullband Context (MOS-LQOF):
 - Reference Signal Fullband flat filtered 20 Hz to 20 kHz.
 - Narrowband system: Test Signal Narrowband IRS Send filtered.
 - Wideband system: Test Signal Wideband flat low pass filtered 7,8 kHz.

For the performance tests with network impairments the following settings are used.

Table 43: Network Conditions for Electrical-Electrical Measurements (Speech Samples)

Condition	Packet Loss (Equal)	Delay Variation	Reordering
0 (see note 2) (VAD)	0	No	No
1	0	No	No
2	1 %	No	No
3	3 %	No	No
4	0	Yes (see note 1)	No
5	1 %	Yes (see note 1)	No
6	0	Yes (see note 1)	Yes
7	1 %	Yes (see note 1)	Yes

NOTE 1: See annex D of ETSI ES 202 737 [32].
 NOTE 2: VAD on, all other conditions (1-5) tested with VAD off.
 NOTE 3: For some network emulation tools, it is necessary to introduce a constant delay to offer the possibility to generate a delay variation distribution. This delay has to be subtracted from the measured delay before interpreting the results.
 NOTE 4: The settings are derived from the ones used in the ETSI Plugtest VoIP speech quality test events.
 NOTE 5: When running tests with the conditions in row, it may be necessary to make one call per condition to avoid the influence of the order of the conditions to the results.

6.3.3 Quality of Jitter buffer adjustment

The listening speech quality and the delay is measured, but with variant network impairments.

Requirements

The speech quality during and after inserted IP delay variation should be as follows.

Table 44: Requirements for variant network impairments

Codec	MOS-LQOF
G.711	> 3,1
G.729	> 2,8
G.723.1	> 2,5
G.722	> 3,6
G.722.1	> 3,8

The delay measured 20 s after ending of the IP delay variation should be maximum 10 ms higher than the delay measured before the IP delay variation.

Test Method

The test signal consists of a CSS-signal, followed by 5 times the same speech sentence, fulfilling the requirements of Recommendation ITU-T P.863.1 [14], then again a CSS signal (20 s after the IP delay variation stops). The speech signal level is averaged over all (original) sentences (8 sentences) used. This test is redone for all of the 8 sentences.

NOTE 1: The 8 sentences used consist of the 8 single sentences taken from the 4 sentence pairs used in clauses 6.3.1 and 6.3.2.

NOTE 2: For every new measurement a new call has to be setup to start with an initial delay. Depending on the algorithm used in the variable jitter buffer (e.g. jitter buffer starting with a high fill size), it may be necessary to let some time pass under clean conditions until the measurement is started.

The first CSS signal is used to measure the delay prior to the IP impairment (in clean network conditions). The second CSS signal is used to measure the delay 20 s after the IP impairment stops. The difference of the two delays is the measurement result for the variation of the jitter buffer per measurement. The overall result is the average of all 8 measurements.

The first sentence (during which IPDV of 50 ms is applied) is used to measure the speech quality during jitter buffer adaptation (low to high). MOS-LQOF of the first sentence is measured using Recommendation ITU-T P.863 [24] in fullband mode. The overall result is the average MOS-LQOF of the 8 measurements.

The second to the fifth sentence (every 5 s a sentence) are used to measure the speech quality during jitter buffer adaptation (high to low). MOS-LQOF is measured using Recommendation ITU-T P.863 [24] in fullband mode for each of these four sentences. The minimum MOS-LQOF of these four sentences is used for the averaging over all 8 measurements. The overall result for the speech quality during jitter buffer adaptation (high to low) is the average of the minimum MOS-LQOF-value of the 8 measurements.

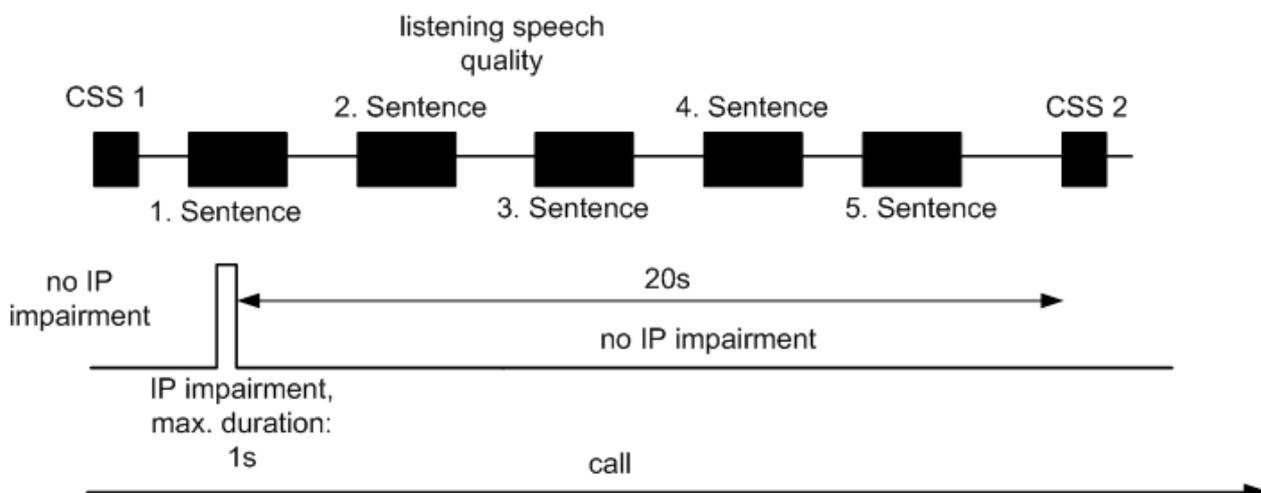


Figure 6: Test Sequence to measure quality of Jitter buffer adjustment (with 1 of 8 sentences)

The IP impairment consists of additional packet delay (IPDV) up to 50 ms, during maximum 1 s. The impairment can be in form of jitter, but also with only some single packets delayed. An example for the impairment can be found in annex C.

Annex A (informative): Impulse Response of a Narrowband and Wideband DECT PP

Table A.1 provides two impulse responses measured for a commercially available DECT PP in narrowband and wideband mode. They are used as the basis for echo path simulation of a typical DECT PP connected to an IAD. It should be noted that the impulse response only covers linear distortion; non-linear distortions as often found in portable phones today are not covered by impulse responses.

Furthermore the impulse responses have to be complemented with the impulse response of the fixed part. In case of an analogue fixed part a hybrid impulse response is used. Examples for narrowband impulse responses can be found in Recommendation ITU-T G.168 [23]. Alternatively individual responses can be used. In case of a digital fixed part the artificial echo loss of 24 dB is used.

The impulse responses contain 1 024 taps at a sampling frequency of 48 kHz (24 bit resolution). This results in an impulse response length of 21,3 ms. The impulse responses are given by the following coefficients.

Table A.1

Narrowband	Wideband
8 721	4 785
8 687	26 708
8 737	32 710
8 856	18 356
9 025	-7 482
9 228	-28 534
9 447	-30 970

The echo path simulation is separated in two parts. First the filter is scaled in such a way that 0 dB TCLw (in narrowband) and 0 dB echo attenuation (in wideband) is achieved. The required TCLw resp. echo attenuation used finally for testing is adjusted separately. For echo canceller testing a TCLw resp. echo attenuation of e.g. 35 dB is realistic and recommended.

The frequency characteristics of the echo path after adjusting the filter to 0 dB TCLw (narrowband) and 0 dB echo attenuation (wideband) is shown in figures A.1 and A.2.

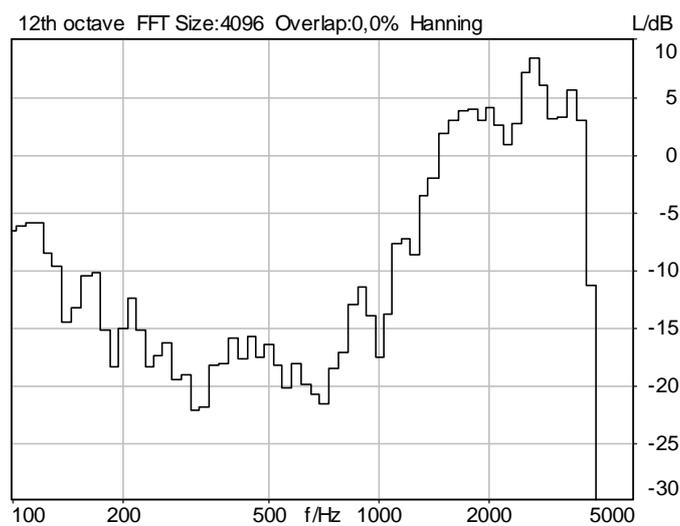


Figure A.1: ERL(f) for the narrowband echopath simulation

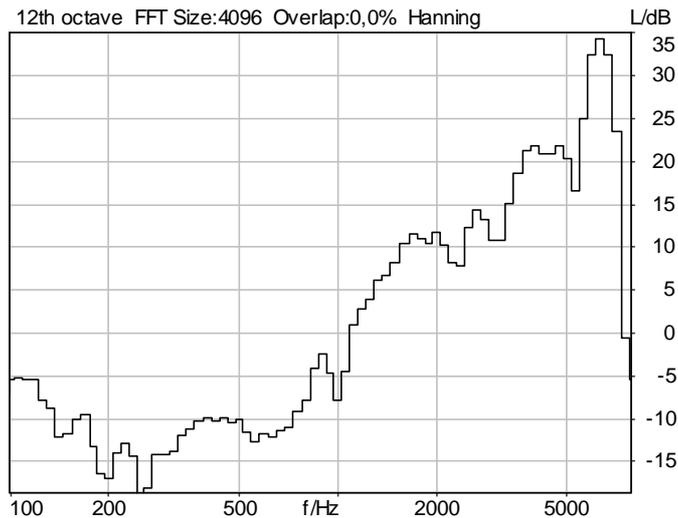


Figure A.2: ERL(f) for the wideband echopath simulation

In wideband the echo attenuation is calculated by calculating the level difference between electrical input and electrical output of the terminal. Since this calculation is based on speech and is dependent on the spectral energy distribution a CSS signal with speech like power density according to Recommendation ITU-T P.501 [16] is used. The spectral representation of this test signal in comparison to the spectra of English speech sentences as found in Recommendation ITU-T P.501 [16] is shown in figure A.3.

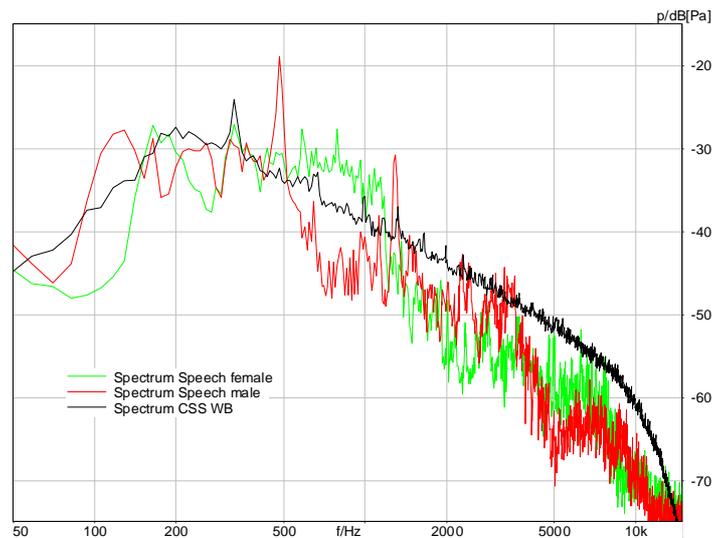


Figure A.3: CSS test signal (black) for determining the echo attenuation for the wideband echopath simulation

Annex B (normative): Test signal for immunity to DTMF false detection in send direction

The type of signal used is a multi speaker signal chosen for containing various frequency and critical for DTMF decoders. The signal is a recording of a polish theatre play used in a French standard for test of Telephone Answering Machines. The duration of the test signal is 38 minutes. For convenience the signal is divided into 94 parts of 25 s duration each, to be played in series.

The signal level has to be adjusted to -20 dBm0 RMS (which corresponds approximately to a speech level of -10 dBm0 when measured with a 20 s averaging).

The signal is applied to the input of MGW and production of DTMF codes is recorded.

The signal is part of the present document and can be downloaded from:

- <http://docbox.etsi.org/STQ/Open/ES%20202%20718%20Test%20signal/>.

Annex C (informative): Example IP Delay variation for jitterbuffer quality measurements

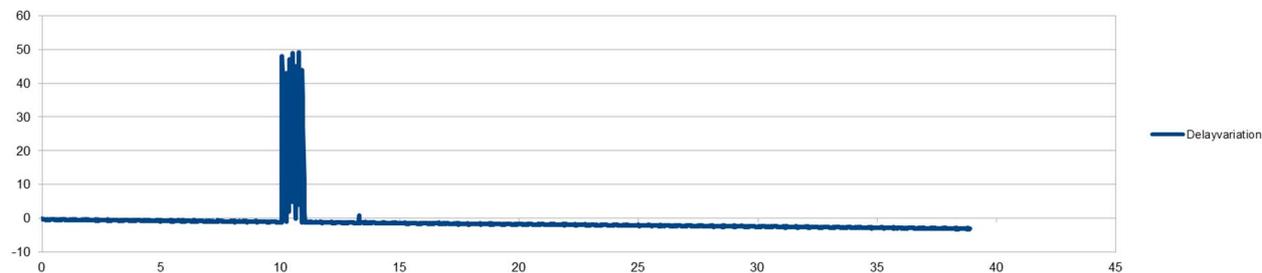


Figure C.1: Example IP delay variation (IPDV) introduced for the measurement shown in clause 6.3.3 (IPDV in ms vs. time in seconds)

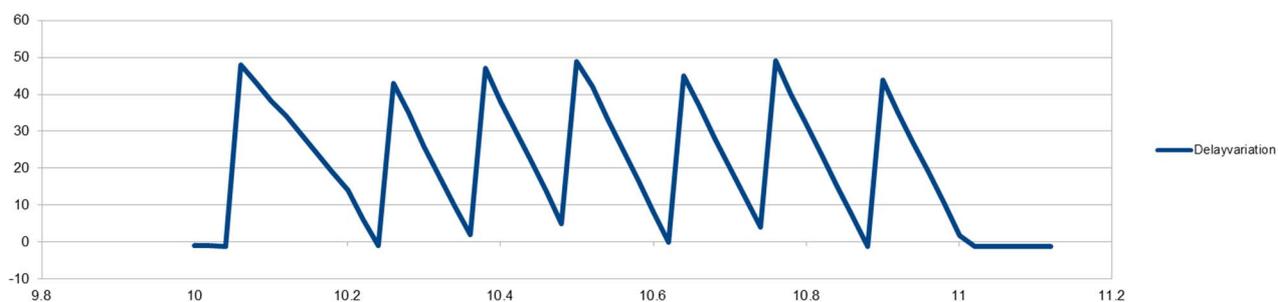


Figure C.2: Zoom in into the IP delay variation part (IPDV in ms vs. time in seconds)

Annex D (informative): Bibliography

- Recommendation ITU-T P.64: "Determination of sensitivity/frequency characteristics of local telephone systems".
- Recommendation ITU-T P.79: "Calculation of loudness ratings for telephone sets".
- Recommendation ITU-T P.581: "Use of head and torso simulator (HATS) for hands-free terminal testing".
- IEC 61260: "Electroacoustics - Octave-band and fractional-octave-band filters".
- ETSI ES 202 739: "Speech and multimedia Transmission Quality (STQ); Transmission requirements for wideband VoIP terminals (handset and headset) from a QoS perspective as perceived by the user".
- ETSI TR 102 648-1: "Speech Processing, Transmission and Quality Aspects (STQ); Test Methodologies for ETSI Test Events and Results; Part 1: VoIP Speech Quality Testing".
- ETSI EG 201 377-2: "Speech Processing, Transmission and Quality Aspects (STQ); Specification and measurement of speech transmission quality; Part 2: Mouth-to-ear speech transmission quality including terminals".
- ETSI ES 201 235 (Parts 1 to 4): "Access and Terminals (AT); Specification of Dual-Tone Multi-Frequency (DTMF) Transmitters and Receivers".

History

Document history		
V1.1.1	October 2011	Publication
V1.2.1	October 2014	Publication
V1.3.1	September 2017	Publication
V1.4.1	March 2020	Membership Approval Procedure MV 20200508: 2020-03-09 to 2020-05-08
V1.4.1	May 2020	Publication