



TECHNICAL REPORT

Speech and multimedia Transmission Quality (STQ); Test methods for insert type headsets enabled with structure-borne speech capture

Reference

DTR/STQ-309

Keywordsbone conduction, headset, sound, speech,
structure-borne, testing**ETSI**650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

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

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

Important notice

The present document can be downloaded from the
ETSI [Search & Browse Standards](#) application.

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 on [ETSI deliver](#).

Users should be aware that the present document may be revised or have its status changed,
this information is available in the [Milestones listing](#).

If you find errors in the present document, please send your comments to
the relevant service listed under [Committee Support Staff](#).

If you find a security vulnerability in the present document, please report it through our
[Coordinated Vulnerability Disclosure \(CVD\)](#) program.

Notice of disclaimer & limitation of liability

The information provided in the present deliverable is directed solely to professionals who have the appropriate degree of experience to understand and interpret its content in accordance with generally accepted engineering or other professional standard and applicable regulations.

No recommendation as to products and services or vendors is made or should be implied.

No representation or warranty is made that this deliverable is technically accurate or sufficient or conforms to any law and/or governmental rule and/or regulation and further, no representation or warranty is made of merchantability or fitness for any particular purpose or against infringement of intellectual property rights.

In no event shall ETSI be held liable for loss of profits or any other incidental or consequential damages.

Any software contained in this deliverable is provided "AS IS" with no warranties, express or implied, including but not limited to, the warranties of merchantability, fitness for a particular purpose and non-infringement of intellectual property rights and ETSI shall not be held liable in any event for any damages whatsoever (including, without limitation, damages for loss of profits, business interruption, loss of information, or any other pecuniary loss) arising out of or related to the use of or inability to use the software.

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 2024.
All rights reserved.

Contents

Intellectual Property Rights	4
Foreword.....	4
Modal verbs terminology.....	4
Introduction	4
1 Scope	5
2 References	5
2.1 Normative references	5
2.2 Informative references.....	5
3 Definition of terms, symbols, and abbreviations.....	5
3.1 Terms.....	5
3.2 Symbols.....	6
3.3 Abbreviations	6
4 Headset signal processing and potential improvements with bone conduction sensing	6
5 Measuring human bone conduction signal	7
5.1 Overview	7
5.2 Measurement setup.....	7
5.3 Measurement results.....	8
6 Human bone conduction simulation.....	10
7 Test signals and test arrangements	11
7.1 Overview	11
7.2 Measurements in the presence of background noise	11
7.2.1 Overview	11
7.2.2 Test methods.....	11
7.2.2.1 Speech followed by background noise - Sequence A	11
7.2.2.2 Background noise followed by speech - Sequence B.....	12
7.2.3 Measurement results	12
7.2.3.1 Speech quality analysis - Sequence A.....	12
7.2.3.2 Noise spectrum analysis - Sequence A.....	14
7.2.3.3 Speech quality analysis - Sequence B	17
7.3 Double talk measurements	18
7.3.1 Test method	18
7.3.2 Measurement results	18
8 Conclusion.....	20
Annex A: Inverse transfer functions	21
History	23

Intellectual Property Rights

Essential patents

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

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

Trademarks

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

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

Foreword

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

Modal verbs terminology

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

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

Introduction

As communication devices evolve to meet the demands of consumers and to enhance user experience, the integration of bone conduction technology into in-ear headsets allows significant advances in speech signal processing. Noise suppression and echo control functionalities utilize information of built-in bone conduction sensors inside the ear canal to better identify the talker's voice and separate it from possible degradations, like e.g. ambient noise or a concurrent talker.

However, standardized state-of-the-art test equipment is not able to simulate any transfer path between air-borne and structure-borne sound and thus, the actual performance of headsets cannot be adequately assessed. Recommendation ITU-T P.58 [i.1], which specifies mouth simulator and usage of artificial ears of a HATS, even requires a rather high decoupling between both components. The present document introduces a modified version of standardized acoustic test equipment, which is able to overcome this limitation.

1 Scope

The present document provides test setups and test methods for headsets, which use human bone conduction as an additional input signal to the air-borne transmitted voice of the near-end talker. The work described in the present document includes test setup description, validation procedures and standardized test methods focusing on the time-variant behaviour of headset devices that utilize bone conduction in their signal processing.

2 References

2.1 Normative references

Normative references are not applicable in the present document.

2.2 Informative references

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

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

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

- [i.1] Recommendation ITU-T P.58: "Head and torso simulator for telephonometry".
- [i.2] Recommendation ITU-T P.57: "Artificial Ears".
- [i.3] Recommendation ITU-T P.501: "Test signals for use in telephony and other speech-based applications".
- [i.4] ETSI ES 202 396-1: "Speech and multimedia Transmission Quality (STQ); Speech quality performance in the presence of background noise; Part 1: Background noise simulation technique and background noise database".
- [i.5] ETSI TS 103 281: "Speech and multimedia Transmission Quality (STQ); Speech quality in the presence of background noise: Objective test methods for super-wideband and fullband terminals".
- [i.6] Recommendation ITU-T P.502: "Objective test methods for speech communication systems using complex test signals".

3 Definition of terms, symbols, and abbreviations

3.1 Terms

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

air-borne sound: sound waves travelling through the air

bone conduction: sound conducted through the bones and tissue of the human skull

structure-borne sound: sound waves pulsating and radiating through a solid structure or medium (e.g. a human skull)

3.2 Symbols

Void.

3.3 Abbreviations

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

BGN	Background Noise
DUT	Device Under Test
FFT	Fast Fourier Transformation
G-MOS	Global Mean Opinion Score
HATS	Head And Torso Simulator
NLP	Non-Linear Processing
MEMS	Micro-ElectroMechanical Systems
MOS	Mean Opinion Score
MRP	Mouth Reference Point
N-MOS	Noise (intrusiveness) Mean Opinion Score
S-MOS	Speech (distortion) Mean Opinion Score
SNR	Signal-to-Noise Ratio

4 Headset signal processing and potential improvements with bone conduction sensing

To define appropriate test methods, it is useful to have a look into areas where bone conduction signals may be used to improve the performance of headsets. The main purpose of using bone-conducted sound in signal processing is the improved separation between:

- near-end talker and near-end background noise;
- near-end talker and near-end concurrent talker;
- near-end talker and far-end talker (double talk detection).

A second benefit of using the bone conduction signal could be to directly transmit this signal instead of transmitting the near-end microphone signal, at least in the low frequency domain. The bone conduction signal is almost free of background noise and therefore ideally *noise-cancelled*.

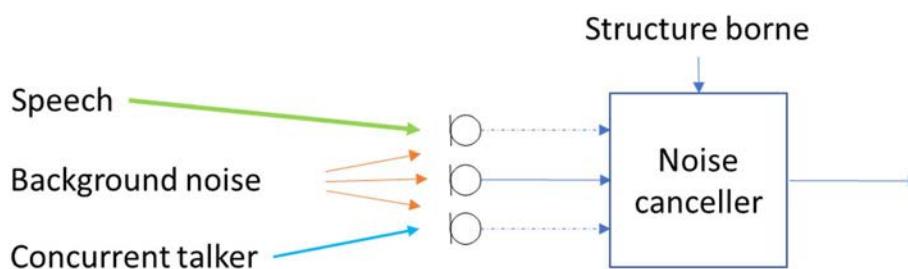


Figure 1: Signals at a near-end noise canceller

Figure 1 illustrates the signals received at a noise canceller. When a bone conduction signal is provided, the separation of near-end speech and impairing signals can be much improved. The noise canceller's background noise estimation can be improved if a clear separation is possible between speech, speech plus noise and noise only. In the presence of speech, the separation between noise and speech is difficult. The noise canceller may diverge, the speech signal may get degraded. When being able to exactly distinguish between speech and noise based on bone conduction signal, the adaptation control can be improved, e.g. adaptation could be frozen or reduced in adaptation speed and the divergence of the noise canceller may be minimized. Consequently, a high-quality speech signal could be provided in conjunction with low background noise. In general, a similar behaviour can be expected with concurrent talkers. If purely the talker wearing the headsets is transmitted, the concurrent talker can be treated like background noise.

Another improvement with bone conduction sensing can be expected in the double talk situation. In almost any communication device, echo cancellation is used to prevent the far-end listener from echo produced at the near-end device. In general, the echo canceller is an adaptive filter modelling the echo path and trying to minimize the echo signal by subtracting the inverse echo signal, as illustrated in Figure 2. The echo canceller may diverge if the impairing near-end speech signal is not reliably detected. The quality of the adaptation control is key, and the better the echo can be separated from the near-end talker signal the better echo cancellation may work in case of double talk. The use of the bone conduction signal may improve this separation since this represents purely the near-end speech signal.

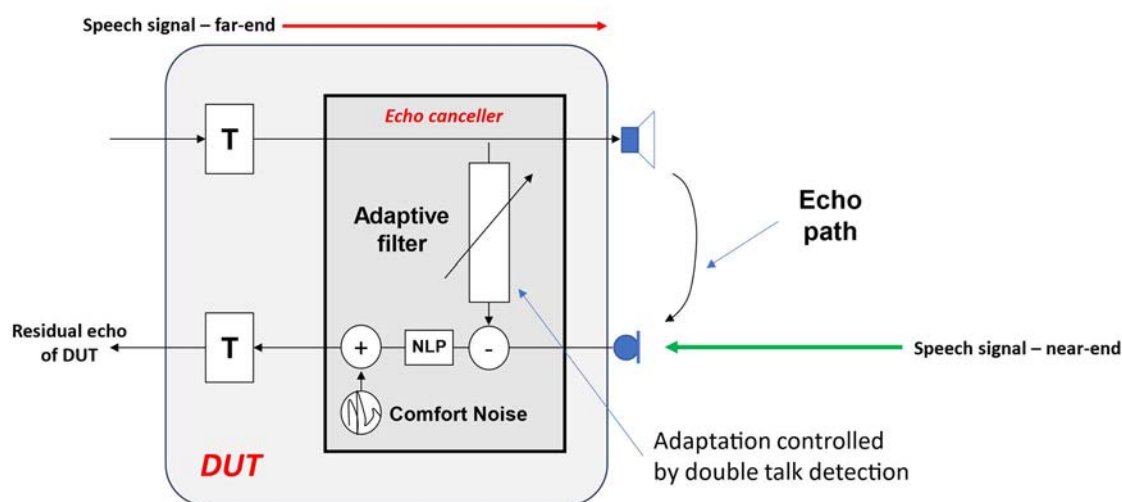


Figure 2: The principal functionality of an echo-canceller and associated signals

5 Measuring human bone conduction signal

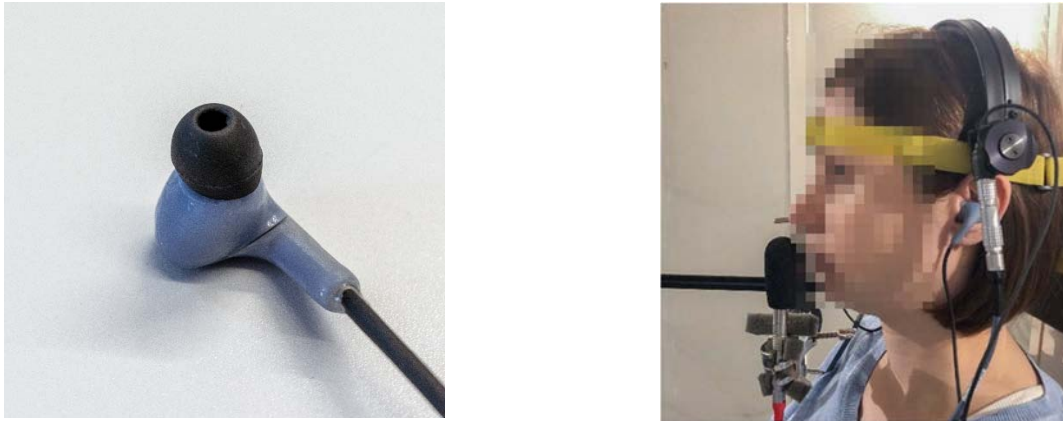
5.1 Overview

The first step in testing headsets with built-in bone conduction sensors is the understanding of human bone conduction signal inside the ear at a location, where bone conduction sensors of headsets are typically placed.

5.2 Measurement setup

To measure the relation between human air-borne signal and bone conduction signal, an experiment was conducted with 11 female and 24 male subjects. The goal was to measure the individual spectra of the air-borne sound and the voice transmitted through bone conduction. The measurements provided an overview of the spread of the individual differences and were the basis for deriving average transfer functions.

The air-borne speech signal was captured using a free field microphone at the MRP of each test person. For assessing human bone conduction instrumentally, a piezoelectric MEMS voice pickup sensor was placed in a mock-up of an in-ear headset which was worn by the test persons during the measurements. The speech sequence spoken by the test subjects consisted of eight German sentences according to clause B.3.7 of Recommendation ITU-T P.501 [i.3], which were repeated three times, having the mock-up mounted on both the left and the right ear. During the second repetition, the test subject applied some pressure on headset to ensure a proper positioning of the headset. The last repetition was used to derive the average transfer functions, but the results showed no significant differences from the first repetition. Figure 3 shows the mock-up of the in-ear headset and the test setup for one individual.



**Figure 3: Left: mock-up of the in-ear headset
Right: test setup for measuring human air-borne and bone-conducted speech signals**

5.3 Measurement results

Figure 4 shows the magnitude of the average Fourier Transform of the individual bone-conducted speech signal compared to the acoustic signal at MRP. The analysis was performed using $1/12^{\text{th}}$ octave band resolution, which is based on a periodogram (Hann window, FFT size of 16K, 66 % overlap). Figure 5 illustrates the average of these signals for female and male talkers. Figure 6 shows the average transfer function for male and female speakers. The difference is mainly in the frequency range below 300 Hz, where female voices generally do not provide much signal energy due to the higher fundamental frequency. Consequently, the average male spectrum was chosen as the target spectrum for structure-borne signal simulation. Note that due to the insufficient SNR of the sensors that were available at time of the experiment, the bone conduction signal could only be measured reliably up to 1 kHz. The inverse transfer functions which were used to generate the bone conduction signal from an air-borne speech signal can be found in Annex A.

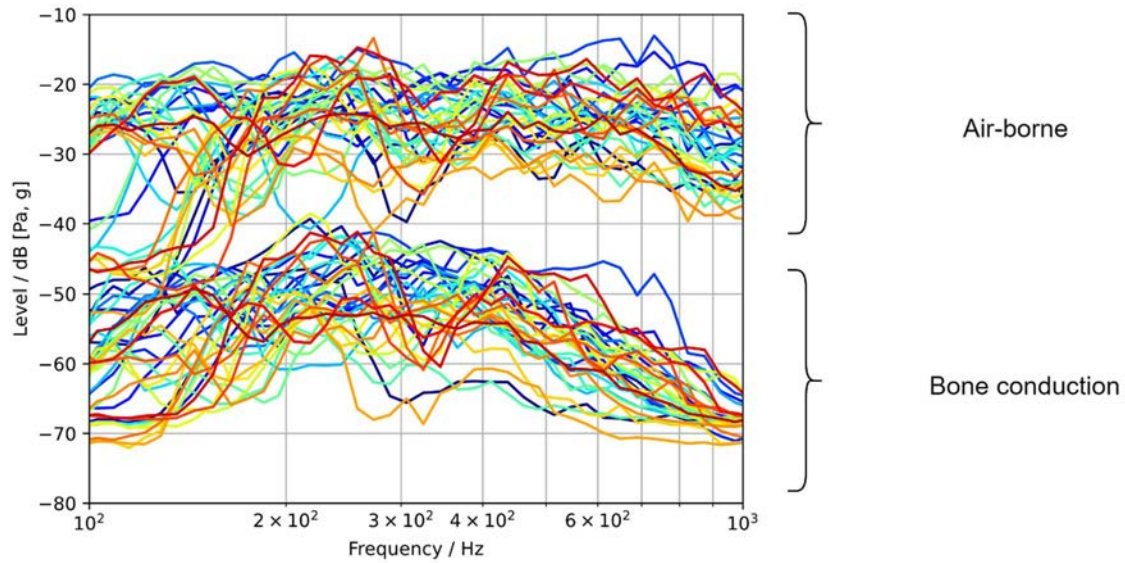


Figure 4: Magnitude of average Fourier Transform of air-borne and bone-conducted signal for each test subject

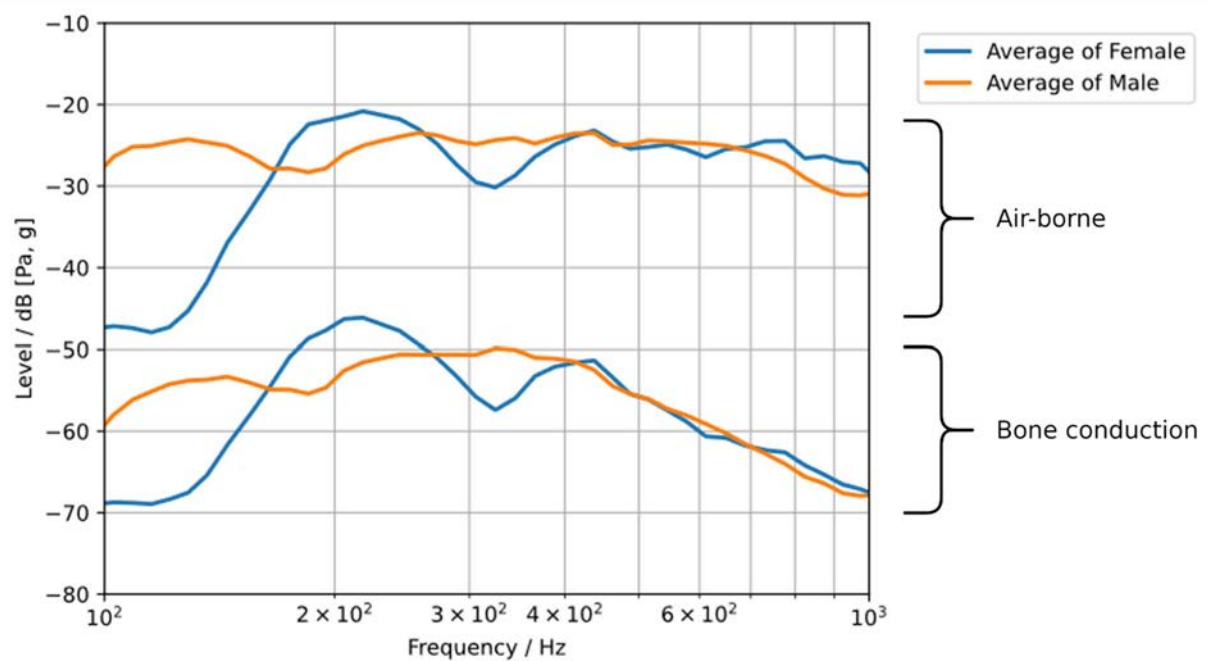


Figure 5: Average magnitude of air-borne and bone-conducted signals for male and female speakers

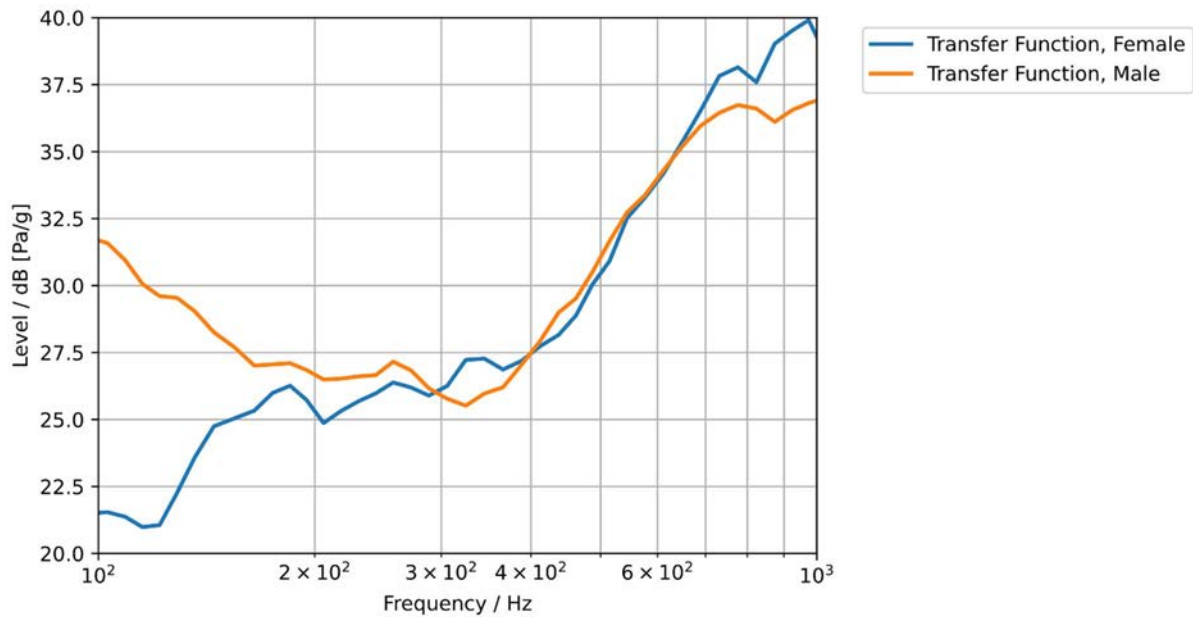


Figure 6: Average transfer functions for male and female speakers

6 Human bone conduction simulation

The simulation of human bone conduction is realized using a HATS complying with Recommendation ITU-T P.58 [i.1], which is equipped with modified type 4.4 artificial ears according to Recommendation ITU-T P.57 [i.2]. The modification consists of an actuator, which allows a realistic simulation of the average human bone-conducted sound for the headset under test. The actuator is integrated close to the position in the ear canal, where also the sensor was located (see clause 5.2). To validate the simulation on HATS, the same sensor as described in clause 5 was used. Figure 7 shows the average spectra of the original and simulated bone-conducted speech signal. It can be seen that the simulation matches the target response within a tolerance of ± 3 dB.

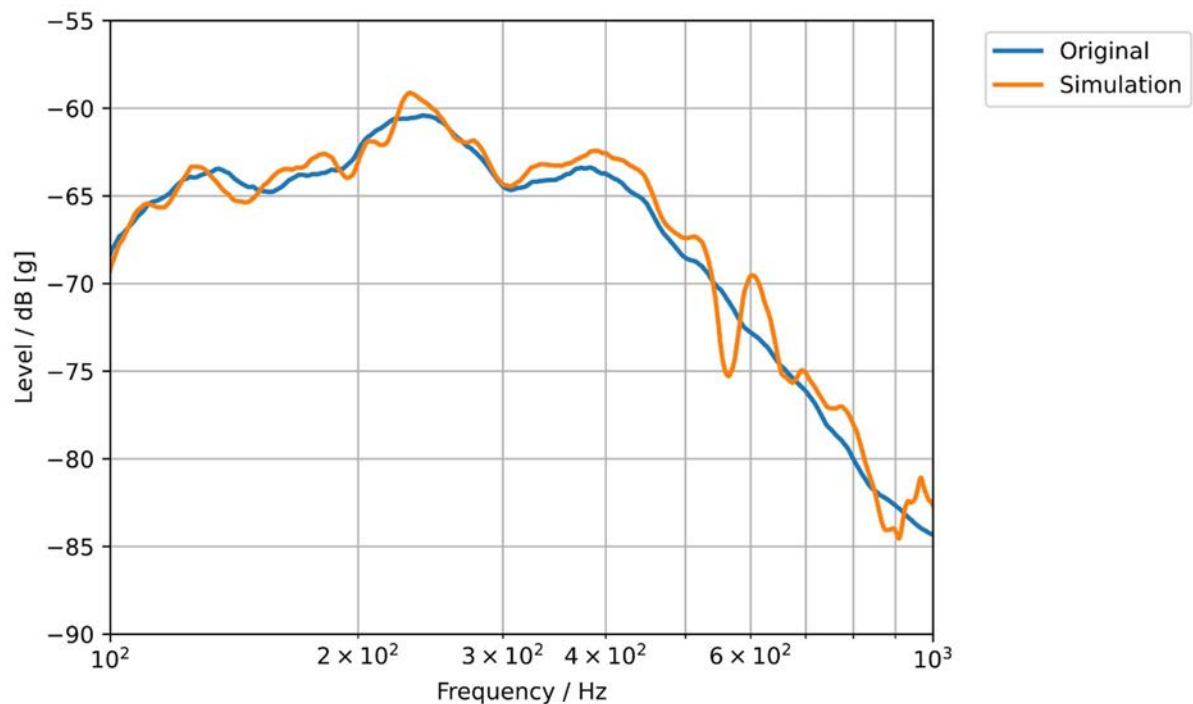


Figure 7: Comparison of simulated and original human average bone-conducted signal

7 Test signals and test arrangements

7.1 Overview

To evaluate the performance of the noise cancellers, specific test sequences were created to show performance differences with and without simulation of bone conduction. Those test sequences include measurements with interference signals and double-talk measurement. An interference signal could be a background noise or a speech signal from a concurrent talker.

For the test sequences defined in this clause, multiple in-ear headsets from various manufacturers were measured. If applicable, different generations of the same headset were tested.

Table 1: Test devices

Device number	Manufacturer	Manufacturing year
1	A	2019
2	B	2020
3	A	2022
4	B	2022
5	C	2022
6	B	2023
7	D	2023
8	B	2023

7.2 Measurements in the presence of background noise

7.2.1 Overview

Two test sequences were created so that speech and interference signal do not start simultaneously. The speech signal corresponds to the English sentences described in ETSI TS 103 281 [i.5] and was calibrated to -4,7 dBPa at MRP. The signal was played back via the artificial mouth of the HATS, which was equipped with the bone conduction simulation. Several binaural background noise scenarios were used to simulate noisy environments according to ETSI ES 202 396-1 [i.4].

7.2.2 Test methods

7.2.2.1 Speech followed by background noise - Sequence A

The first test sequence (denoted as *Sequence A* hereafter) focuses on the potential divergence of the noise canceller which may happen if suddenly a background noise situation occurs. In this situation, the device may not be able to separate speech and noise, resulting in degraded speech and/or insufficiently cancelled noise.

Two analyses were carried out to investigate the performance of the noise canceller in detail, as indicated in Figure 8:

- S-MOS (speech distortion), N-MOS (noise intrusiveness) and G-MOS (overall quality) are calculated according to ETSI TS 103 281 [i.5] on the signal parts that overlap with the noise.
- The noise spectrum is analysed immediately after the playback of the speech sequence.

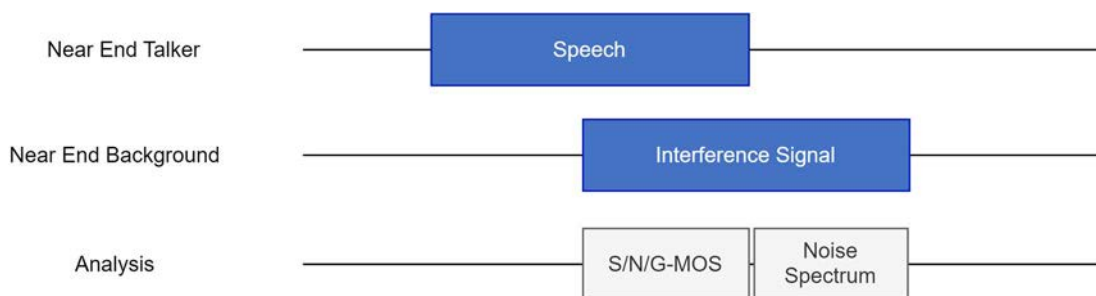


Figure 8: Test sequence A and corresponding analysis

7.2.2.2 Background noise followed by speech - Sequence B

The focus of the second test is on the potential divergence of the noise canceller, when speech suddenly occurs during the background noise playback (denoted as *Sequence B* hereafter), as illustrated in Figure 9. Here the focus is set to the degradation of the speech signal, as well as on a possible change in noise reduction performance. S-MOS, N-MOS, and G-MOS were calculated for simultaneously active speech and background noise parts.

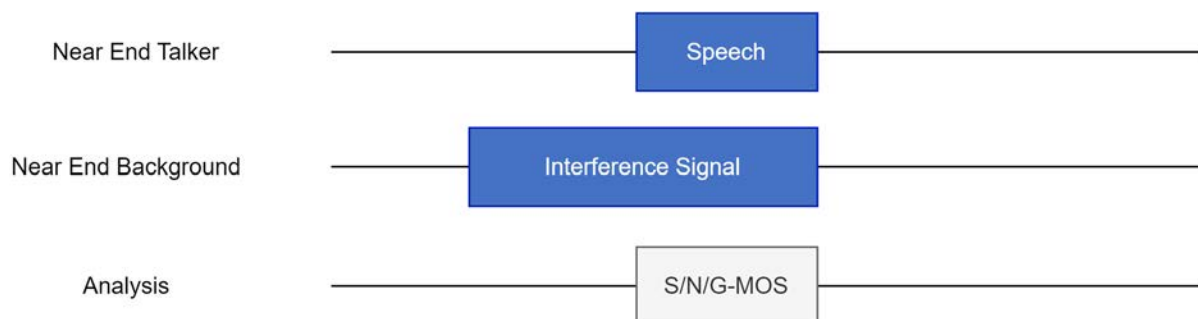


Figure 9: Test sequence B and corresponding analysis

7.2.3 Measurement results

7.2.3.1 Speech quality analysis - Sequence A

This clause provides the measurement results from clause 7.2.2.1 based on the MOS values calculated using the method described in ETSI TS 103 281 [i.5], as shown in Table 2 to Table 5. Improvements in terms of MOS due to the bone conduction simulation are highlighted in orange. The test results for almost all devices show a clear improvement in S-MOS (speech quality) and indicate that a bone conduction sensor can improve noise cancellation in the device by e.g. detecting active/non-active speech more reliably, which results in an improved speech quality in the presence of background noise.

Table 2: MOS results for device 1 and 2, test sequence A

Device number	1				2			
	Pub		Midsize car 130 km/h		Pub		Midsize car 130 km/h	
Structure-Borne	Off	On	Off	On	Off	On	Off	On
S-MOS	2,7	2,8	3,0	3,1	2,5	2,8	2,0	3,6
N-MOS	2,7	2,7	3,7	3,6	2,8	2,7	3,9	4,1
G-MOS	2,1	2,1	2,7	2,6	2,0	2,1	1,8	3,2
Average SNR [dB]	5,38	5,32	20,70	26,27	5,81	7,43	23,71	33,86

Table 3: MOS results for device 3 and 4, test sequence A

Device number	3				4			
BGN scenario	Pub		Midsize car 130 km/h		Pub		Midsize car 130 km/h	
Structure-Borne	Off	On	Off	On	Off	On	Off	On
S-MOS	2,3	2,6	3,6	3,4	2,6	3,2	3,3	3,6
N-MOS	2,9	3,0	3,5	3,8	3,5	3,7	3,9	3,9
G-MOS	1,8	2,0	3,1	3,0	2,2	2,7	2,9	3,2
Average SNR [dB]	2,81	7,34	12,98	26,47	6,72	18,22	22,16	23,03

Table 4: MOS results for device 5 and 6, test sequence A

Device number	5				6			
BGN scenario	Pub		Midsize car 130 km/h		Pub		Midsize car 130 km/h	
Structure-Borne	Off	On	Off	On	Off	On	Off	On
S-MOS	1,9	2,3	2,2	2,7	2,0	2,9	2,7	3,4
N-MOS	4,0	4,1	4,7	4,7	4,7	4,6	4,8	4,7
G-MOS	1,8	2,0	2,3	2,8	2,3	2,7	2,9	3,3
Average SNR [dB]	13,13	19,90	32,09	31,26	15,37	25,05	29,81	37,58

Table 5: MOS results for device 7 and 8, test sequence A

Device number	7				8			
BGN scenario	Pub		Midsize car 130 km/h		Pub		Midsize car 130 km/h	
Structure-Borne	Off	On	Off	On	Off	On	Off	On
S-MOS	2,1	2,9	3,4	3,5	2,2	3,1	2,5	3,4
N-MOS	4,4	4,4	4,8	4,7	3,7	3,8	4,7	4,7
G-MOS	2,1	2,6	3,2	3,4	2,1	2,7	2,7	3,3
Average SNR [dB]	21,38	27,41	34,82	31,79	8,35	17,55	35,22	38,77

7.2.3.2 Noise spectrum analysis - Sequence A

This clause illustrates the noise spectra calculated according to clause 7.2.2.1 for different test devices. The spectra were calculated using a Hann window, FFT size of 16K, and 66 % overlap. A 1/12th octave smoothing was applied to the resulted spectra. In Figure 10 and Figure 11, it can be noticed that devices 1, 2, and 5 seem to use the information provided from the bone conduction simulation in order to attenuate the background noise transmitted in the uplink after speech is stopped. The attenuation in device 1 takes place in the frequency range up to 500 Hz, whereas devices 2 and 5 attenuate the noise across the whole frequency range. Other devices do not utilize the bone-conducted speech signal and probably apply noise reduction based on the air-borne signal. This is the reason why there is no such difference in the noise spectrum when simulating the human bone conduction speech.

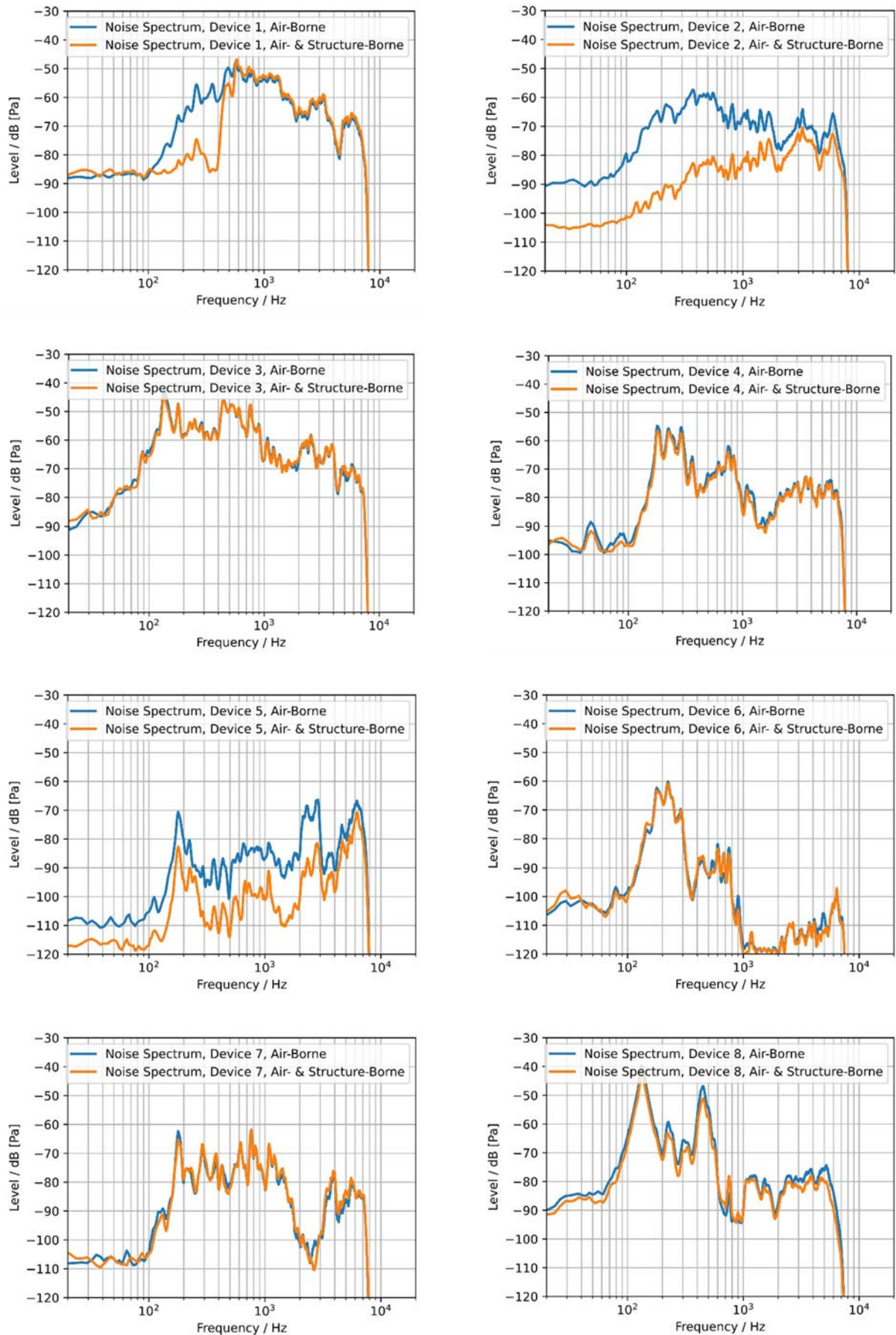


Figure 10: Noise spectrum in uplink with and without bone conduction simulation
BGN: Pub

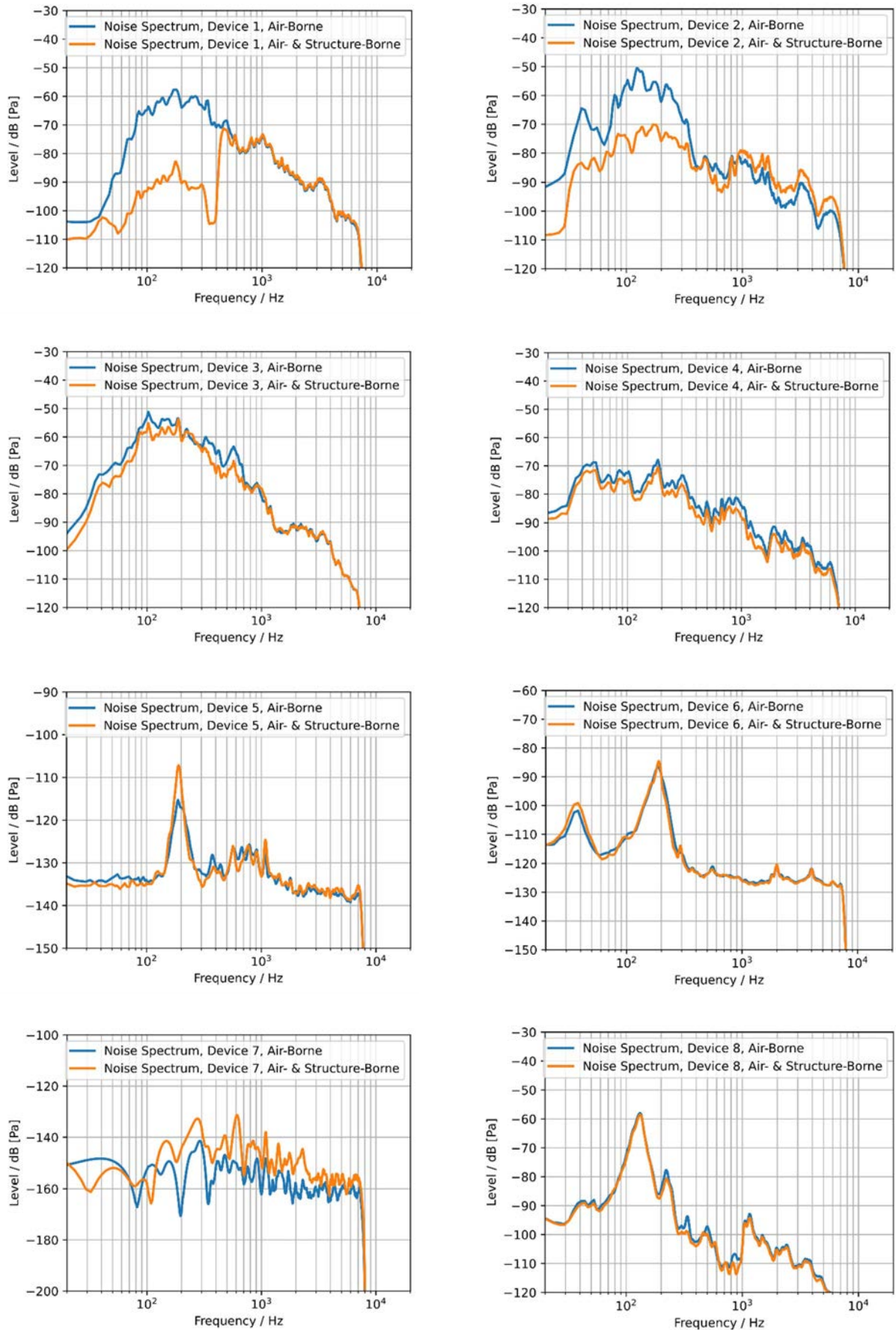


Figure 11: Noise spectrum in uplink with and without bone conduction simulation
 BGN: Midsize Car 130 km/h

7.2.3.3 Speech quality analysis - Sequence B

In this clause, results from clause 7.2.2.2 are provided in Table 6 to Table 9. The performance of the device under test in noisy environment is evaluated with and without bone conduction simulation. Similar to the results in clause 7.2.3.1, improvements can be noticed in MOS values with bone conduction simulation, which indicates that these devices benefit from the simulated bone conduction signal to improve the speech quality in noisy environments.

Table 6: MOS results for device 1 and 2, test sequence B

Device number	1				2			
	Pub		Midsize car 130 km/h		Pub		Midsize car 130 km/h	
Structure-Borne	Off	On	Off	On	Off	On	Off	On
S-MOS	2,5	2,7	2,5	3,1	2,2	2,3	2,0	3,5
N-MOS	2,7	2,7	3,8	3,6	2,7	3,2	4,1	4,1
G-MOS	1,8	2,0	2,4	2,6	1,8	2,1	1,9	3,1
Average SNR [dB]	3,76	4,70	19,54	21,23	3,62	7,05	22,52	27,83

Table 7: MOS results for device 3 and 4, test sequence B

Device number	3				4			
	Pub		Midsize car 130 km/h		Pub		Midsize car 130 km/h	
Structure-Borne	Off	On	Off	On	Off	On	Off	On
S-MOS	2,8	2,5	3,6	3,1	2,4	3,4	3,7	3,8
N-MOS	2,8	3,0	3,5	3,8	3,8	3,7	3,5	2,8
G-MOS	2,0	1,8	3,2	2,7	2,0	3,0	3,3	3,0
Average SNR [dB]	3,41	5,17	14,7	26,92	18,36	24,23	15,53	9,98

Table 8: MOS results for device 5 and 6, test sequence B

Device number	5				6			
	Pub		Midsize car 130 km/h		Pub		Midsize car 130 km/h	
Structure-Borne	Off	On	Off	On	Off	On	Off	On
S-MOS	1,8	3,1	2,1	3,1	2,0	2,8	2,5	3,3
N-MOS	4,4	3,7	4,7	4,7	4,6	4,7	4,7	4,7
G-MOS	1,9	2,6	2,0	3,0	2,3	2,8	2,7	3,2
Average SNR [dB]	19,59	21,55	34,36	32,16	15,47	29,20	27,43	36,86

Table 9: MOS results for device 7 and 8, test sequence B

Device number	7				8			
	Pub		Midsize car 130 km/h		Pub		Midsize car 130 km/h	
Structure-Borne	Off	On	Off	On	Off	On	Off	On
S-MOS	1,9	2,7	3,1	3,4	2,2	3,2	2,5	3,5
N-MOS	4,4	4,2	4,7	4,7	4,0	3,9	4,7	4,7
G-MOS	1,9	2,4	2,9	3,4	2,3	2,8	2,7	3,4
Average SNR [dB]	17,59	24,09	34,68	33,23	12,64	25,44	34,85	37,36

7.3 Double talk measurements

7.3.1 Test method

For double talk testing, the double talk speech test signal from Recommendation ITU-T P.501 [i.3] is used, as shown in Figure 12. The attenuation range in double talk as described in Recommendation ITU-T P.502 [i.6] is calculated. The detailed level vs. time analysis of the double talk signal allows a detailed investigation of potential improvement of the double talk performance with bone conduction functionality. The signal of the competing speaker (upper part of Figure 12) is used for the mouth playback.

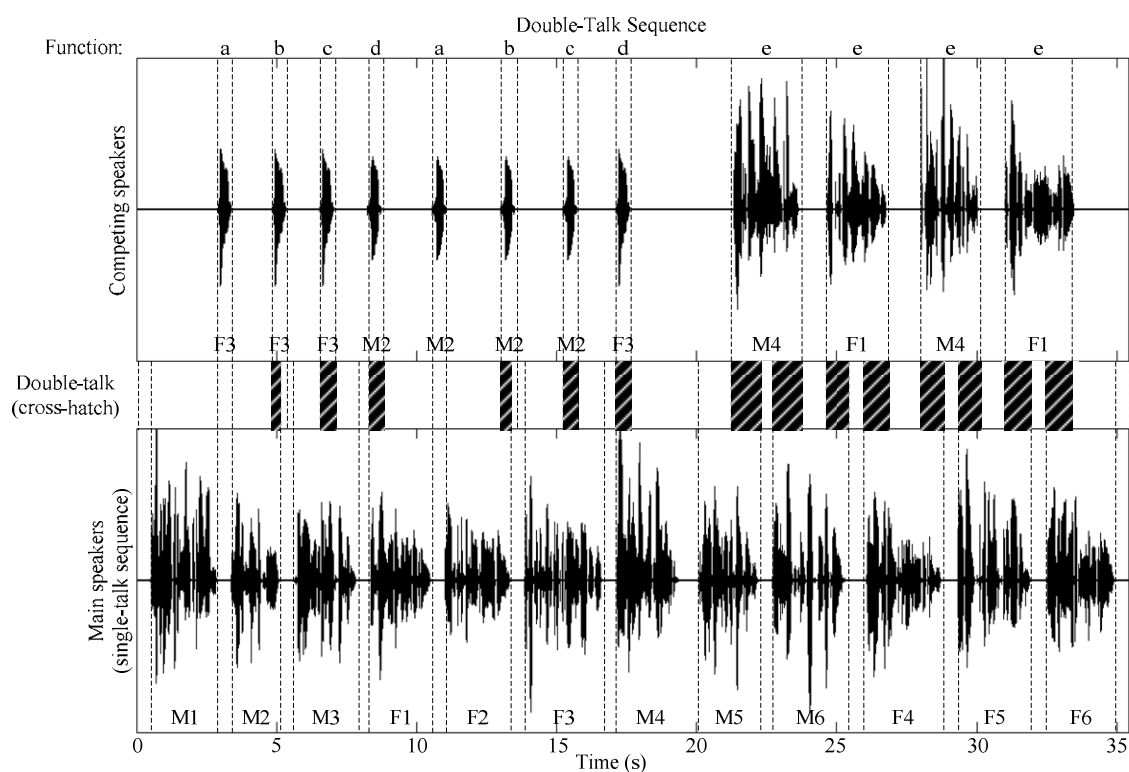


Figure 12: Double-talk speech sequence from Recommendation ITU-T P.501 [i.3]

7.3.2 Measurement results

The double talk attenuation for each device is illustrated in Figure 13. Although the attenuation in sending in double talk is already quite low for most devices, it can be seen, that with the use of structure borne, these headsets are more precise in the detection of the send signal and introduce even less attenuation. The attenuation graph with bone conduction simulation active (orange) shows even less attenuation and an almost transparent behaviour of the headset in this double talk situation.

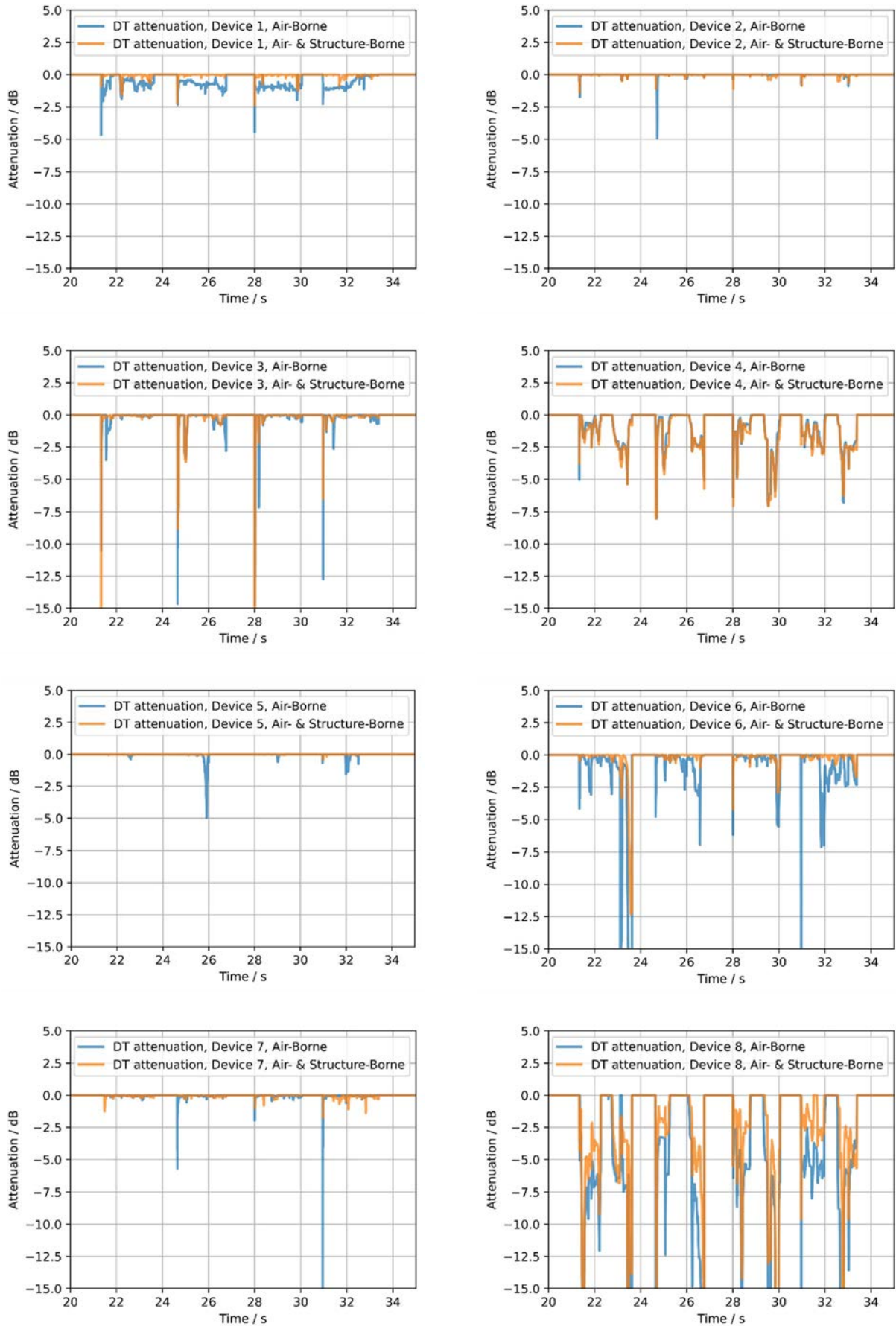


Figure 13: Comparison of double talk attenuation in send direction with and without bone conduction simulation

8 Conclusion

The present document presents a new method for testing in-ear headset devices which utilize bone conduction sensing and identifies new test sequences to evaluate the performance of noise canceller. The conducted experiments resulted in the following main findings:

- The actual performance of DUTs featuring bone conduction sensors may be significantly degraded when testing without bone conduction simulation.
- Most DUTs rely strongly on the information provided to its bone conduction sensor to detect speech and allow a better control of the noise canceller.
- When measured with bone conduction simulation, speech quality and noise reduction were improved in the uplink for almost all test devices. Devices with bone conduction sensors can be tested and optimized more reliably in a lab where it is possible to fine-tune the noise cancelling algorithms thus optimizing their utilization.
- Most of tested DUTs already have a low double talk attenuation in send direction. However, some headsets were more precise and introduced less attenuation in the uplink signal when simulating the bone conduction signal.
- The results indicate that it is beneficial to simulate the human bone conduction signal when measuring headsets, regardless of whether they are actually equipped with an appropriate bone conduction sensor. For devices not providing this functionality, performance is not further affected.

Annex A: Inverse transfer functions

Figure A.1 shows the inverse of the average transfer functions illustrated in Figure 6. These transfer functions can be used to convert a human air-borne speech signal into its corresponding bone conduction signal. Due to the insufficient SNR of the sensors that were available at the time of the experiment, the bone conduction signal could only be measured reliably in the frequency range between 100 Hz and 1 000 Hz. Thus, when using these transfer functions, it is assumed to have a behaviour similar to a bandpass outside this frequency range.

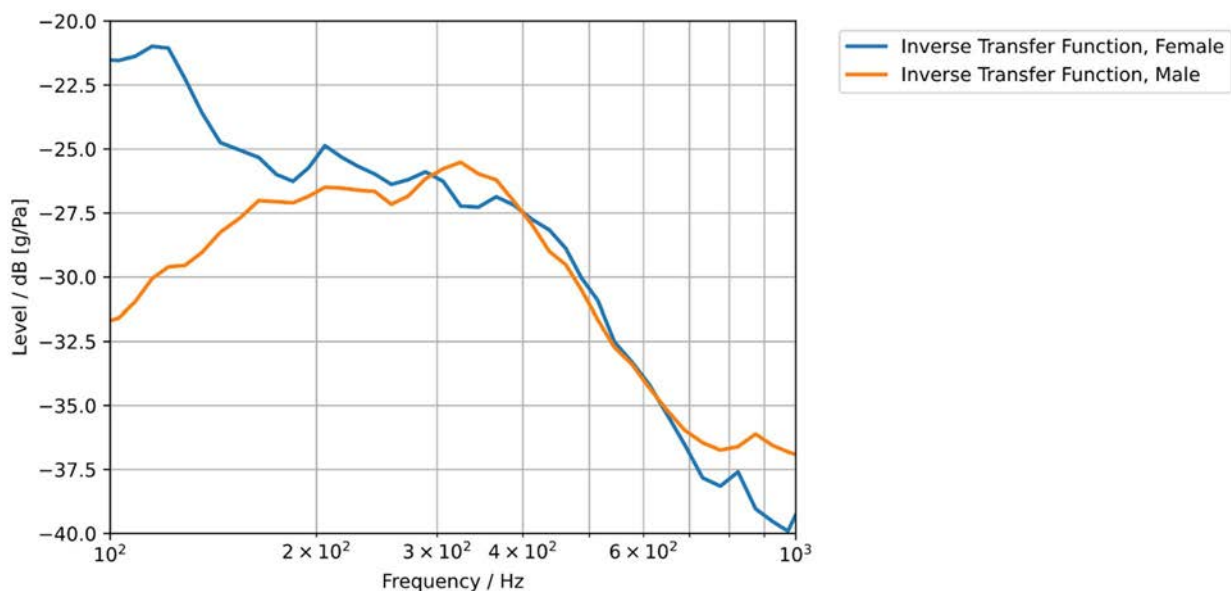


Figure A.1: Inverse average transfer functions for female and male speakers

The magnitudes of both transfer functions in $1/12^{\text{th}}$ octave bands between 100 Hz and 1 000 Hz can be found in Table A.1.

Table A.1: Values of inverse average transfer functions for female and male speakers

Frequency / Hz	Level / dB [g/Pa]	
	Female	Male
102,87	-21,54	-31,59
108,87	-21,37	-30,96
115,12	-20,99	-30,06
121,61	-21,06	-29,61
128,59	-22,25	-29,55
136,09	-23,58	-29,05
144,75	-24,75	-28,26
154,75	-25,04	-27,69
164,78	-25,33	-27,01
174,78	-25,99	-27,06
184,80	-26,26	-27,10
194,80	-25,73	-26,85
205,75	-24,87	-26,49
217,74	-25,32	-26,53
230,23	-25,68	-26,61
243,23	-25,98	-26,66
257,18	-26,38	-27,16
272,18	-26,20	-26,84
288,37	-25,89	-26,16
305,86	-26,26	-25,78
324,56	-27,23	-25,52
344,55	-27,27	-25,97
365,75	-26,87	-26,21
388,24	-27,19	-27,05
411,95	-27,74	-27,94
436,94	-28,16	-28,99
462,01	-28,88	-29,52
487,00	-30,04	-30,50
514,37	-30,91	-31,65
544,36	-32,52	-32,74
576,73	-33,30	-33,38
611,72	-34,19	-34,31
649,11	-35,35	-35,18
689,09	-36,54	-35,97
731,50	-37,83	-36,45
776,48	-38,15	-36,74
823,91	-37,59	-36,61
873,89	-39,03	-36,11
924,02	-39,52	-36,55
974,00	-39,91	-36,81

History

Document history		
V1.1.1	August 2024	Publication