Project: The Verge – ANC Headphone showdown - Summary Report

Background

Evaluating ANC can be done subjectively, but testing ANC objectively is rarely done in a proper acoustic way. Often, the methods and setups are way too simplistic and contrived, such as using one or two loudspeakers to generate some deterministic and steady state noise (White or Pink Noise), and then evaluating the overall attenuation as a single number – or spectrally by displaying an fft plot.

Instead, HEAD acoustics proposes 4 simple tweaks to the test setup in order to get results that correlate much better with real world experiences:

1) Use realistic background noises – like those captured in airplane cabins, busy street intersections, train and metro stations, etc.
2) Using Hearing Models to analyze the residual noise at the ear drum – the human brain is not an FFT analyzer!
3) Analyzing temporal aspects – similarly to above, the human brain is sensitive to events in the time domain, and the ANC needs to be able to handle transients well.
4) Using advanced models to inspect effects on the audio/speech playback quality – if the ANC is effective enough, your podcast can be enjoyed more thoroughly.

In connection with The Verge, this project aims to measure the Active Noise Cancellation (ANC) performance of various headsets in different environments to provide an in-depth analysis of their true effectiveness.

Testing environment

4 noise environments were identified that people will often encounter where they will use their consumer ANC headphone. These environments were all recorded for about 30+ seconds to capture a completely representative sample of the noise that includes transients, starts and stops, and continuous noise segments.

Using the unobstructed ears of the Head and Torso Simulator, a specific 10 second sample of the noises were measured, and the sound pressure levels reported in the brackets for each background noise listed below:

1) Pink Noise [81.38 dBA]
2) Loud Music [93.16 dBA]
3) Railway platform [77.56dBA]
4) Crossroads [68.60dBA]
5) Airplane Cabin [82.16 dBA]

The technology used for recording the background noise environments and reproducing them in the lab is extensively documented in ETSI TS103 224:
https://www.etsi.org/deliver/etsi_ts/103200_103299/103224/01.06.01_60/ts_103224v010601p.pdf

In short, an 8-microphone array is used to capture the noise environments, and then an 8-speaker setup is used in conjunction with 8 microphone array to equalize the room and the sound field for playback. The result is a sound field where magnitude and phase between original source material and reproduced material is as close as possible, and the devices under test in the lab will think they are in the real world.

Devices under test

The following 5 devices were provided by The Verge and tested:

1) BOSE QC-1
2) BOSE NC700
3) BOSE QC-45
4) SONY WH-1000XM5
5) APPLE AirPod Max

Product Performance Parameters

Key information from the data collected can be used to compare the product performance between devices.

Some example benchmarks:

- **Overall attenuation in dB**: One number representing the sound level at the ears without a product, and with the product. Expressed in dB.
- **Compare spectrum of attenuation**: A plot to show passive - and active – attenuation in the frequency domain to make a judgement on mechanical and DSP acoustic capabilities.
- **Level vs. time**: A look at device performance over time to track attack times and convergence. Especially interesting in transient noise environments.
- **LAeq**: A very common way for government/industry to rate noise levels and calculate noise exposures. Expressed in dB SPL.
- **Loudness level reduction**: Apply psychoacoustics to the noise levels at the ears to examine human perception and get a much better gauge on the subjective experience. A typical hearing model applied is that of Zwicker from ISO532-1: https://webstore.ansi.org/standards/iso/iso5322017 Results are expressed in phon.
- **Speech Intelligibility Index (SII)**: Use the noise levels measured at the ear drum to estimate how much effort it would be to have conversations in that environment. More info on SII can be found in ANSI S3.5: https://webstore.ansi.org/standards/asa/ansiasas31997r2017 Results are expressed in percent.
- **Speech AND Noise Suppression Quality metrics**: Measure speech in noise at the ear drums and evaluate how both the quality of the reproduced speech, along with the noise suppression quality. The ANC must be effective, without ruining the audio playback. We're relying on ETSI TS 103 281: https://www.etsi.org/deliver/etsi_ts/103200_103299/103281/01.03.01_60/ts_103281v010301p.pdf Results are expressed on a MOS scale from 1 to 5, 1 being the worst, 5 being the best.

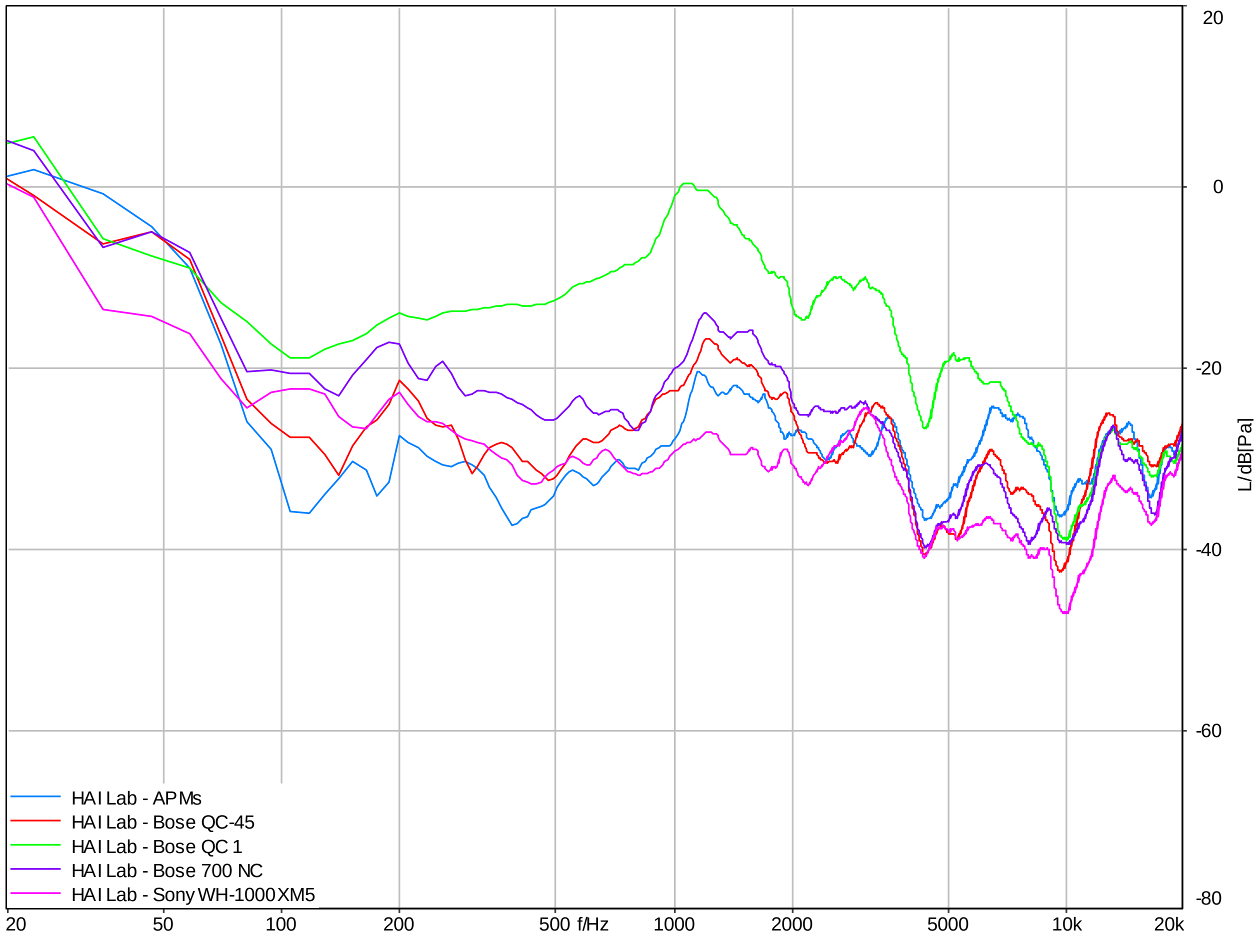
Results

A table of results for each of the performance parameters is presented below. The best performing result in each condition and metric is highlighted in green. Close performing results are colored light green.

	Noise condition	BOSE QC-1		BOSE NC700		BOSE QC45		Sony WH-1000XM5		Apple AirPod Max	
ANC mode		OFF	ON	OFF	ON	OFF	ON	OFF	ON	OFF	ON
Overall Attenuation [dB]	Cond #1	9.4	7.84	16.31	20.35	16.29	24.92	19.38	30.42	19.69	26.11
	Cond #2	5.54	4.21	12.81	20.24	11.37	22.17	12.58	28.46	15.64	24.06
	Cond #3	5.75	6.68	12.23	19.41	11.76	22.79	14.49	29.51	16.28	25.95
	Cond #4	6.12	5.24	11.17	17.91	11.12	21.12	13.25	28.16	12.61	23.54
	Cond #5	0.72	9.9	2.84	21.32	2.26	25.73	3.32	26.27	7.11	29.96
Laeq [dB SPL]	Cond #1	71.91	73.15	66.43	61.77	64.93	56.75	61.86	49.69	61.26	54.29
	Cond #2	93.51	91.67	89.49	80.2	89.28	75.35	88.2	68	84.83	69.88
	Cond #3	72.11	70.43	66.67	58.82	65.31	53.64	62.2	46.6	60.71	50.73
	Cond #4	61.33	61.81	58.25	50.16	56.53	54.91	54.72	37.86	55.25	44.5
	Cond #5	8216	73.49	79.93	63.45	76.76	43.37	78.2	54.46	75.4	51.65
Loudness Level Reduction [phon]	Cond #1	8.8	11.3	14.6	21.5	15.3	24.5	19.2	28.6	16	24.7
	Cond #2	4.4	5.7	11	18.8	9.6	20.8	11.9	27	13.8	23.5
	Cond #3	6	9.7	11.5	20	12	23.5	15.7	29	13.8	25.3
	Cond #4	6.6	10	11.9	20.6	12.1	23.8	15.8	30	13	25.5
	Cond #5	2.4	10.5	3.9	19.5	3.9	23.4	5.1	26.5	7.2	27.6
SII [%]	Cond #1	6.85	14.09	44.02	56.42	45.33	66.61	62.9	71.25	45.7	62.48
	Cond #2	7.83	8.3	24.29	36.91	27.13	38.37	39.36	51.88	25.4	42.31
	Cond #3	13.98	19.82	50.24	63.7	52.92	67.63	70.87	79.89	54.53	72.83
	Cond #4	46.83	47.95	77.96	84.66	79.13	87.42	91.52	95.55	82.83	92.23
	Cond #5	39.77	35.69	57.79	73.1	58.31	74.79	66.78	86.41	63.43	85.37
NOISE Suppression Quality [MOS 1..5]	Cond #1		1.5		2.3		2.3		2.6		2.4
	Cond #2		1.9		1.7		1.7		1.9		1.8
	Cond #3		1.7		2.7		2.7		3		2.8
	Cond #4		2.3		3.3		3.3		3.5		3.3
	Cond #5		1.8		3		2.9		3.3		3.3
SPEECH Quality [MOS 1..5]	Cond #1		2.8		3.7		3.6		3.6		4.2
	Cond #2		2.5		3.2		3.1		3.3		3.6
	Cond #3		3		3.7		3.6		3.6		4.2
	Cond #4		3.4		3.8		3.7		3.6		4.3
	Cond #5		3.1		3.8		3.6		3.5		4.3

Looking at the overall attenuation spectrally can also highlight the performance differences of devices.

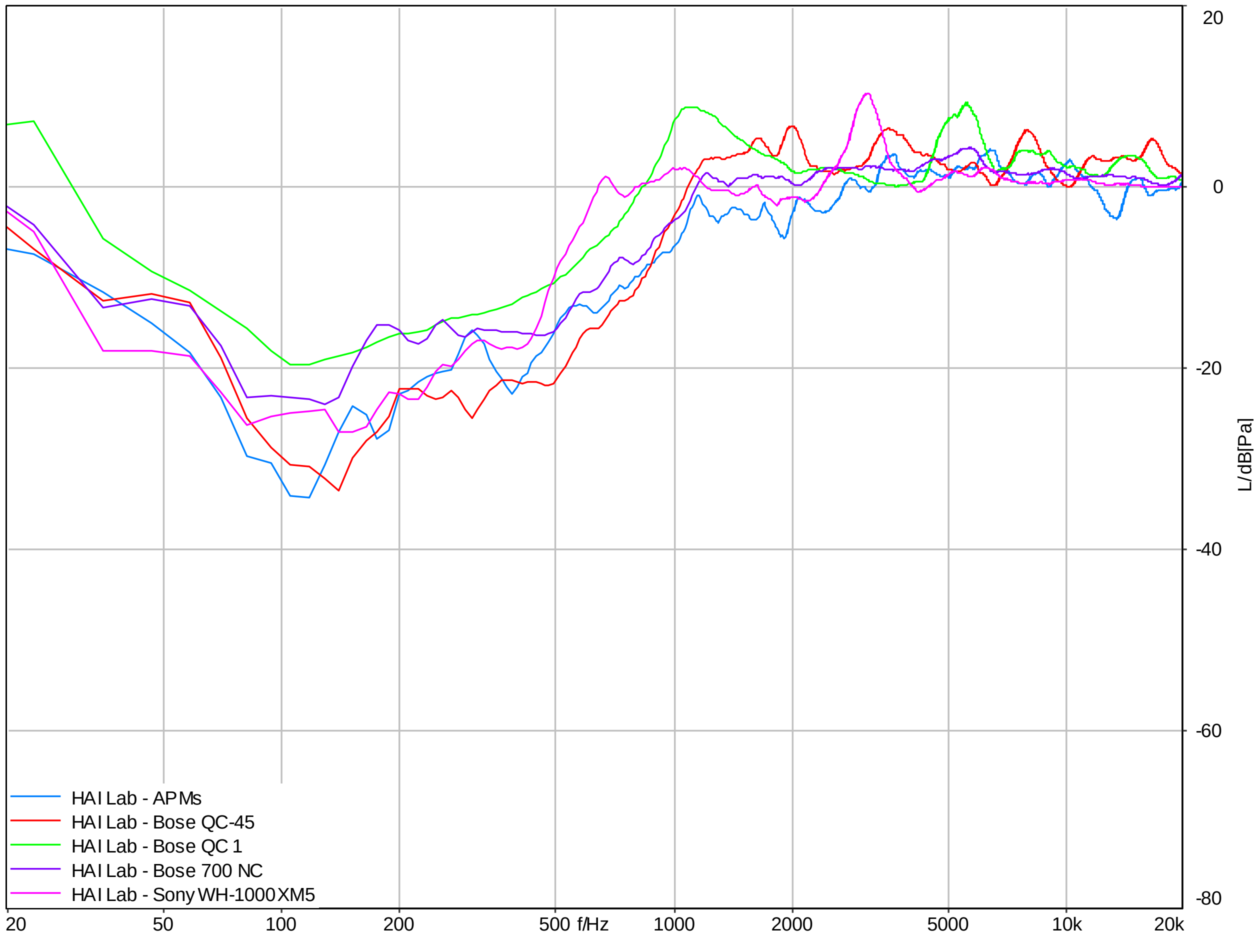
The total attenuation for each of the 5 devices in BGN Cond #5 (Airplane Cabin) is presented in the plot below. In these plots any data point below 0 dB is the device attenuating the noise from reaching the ear drum:



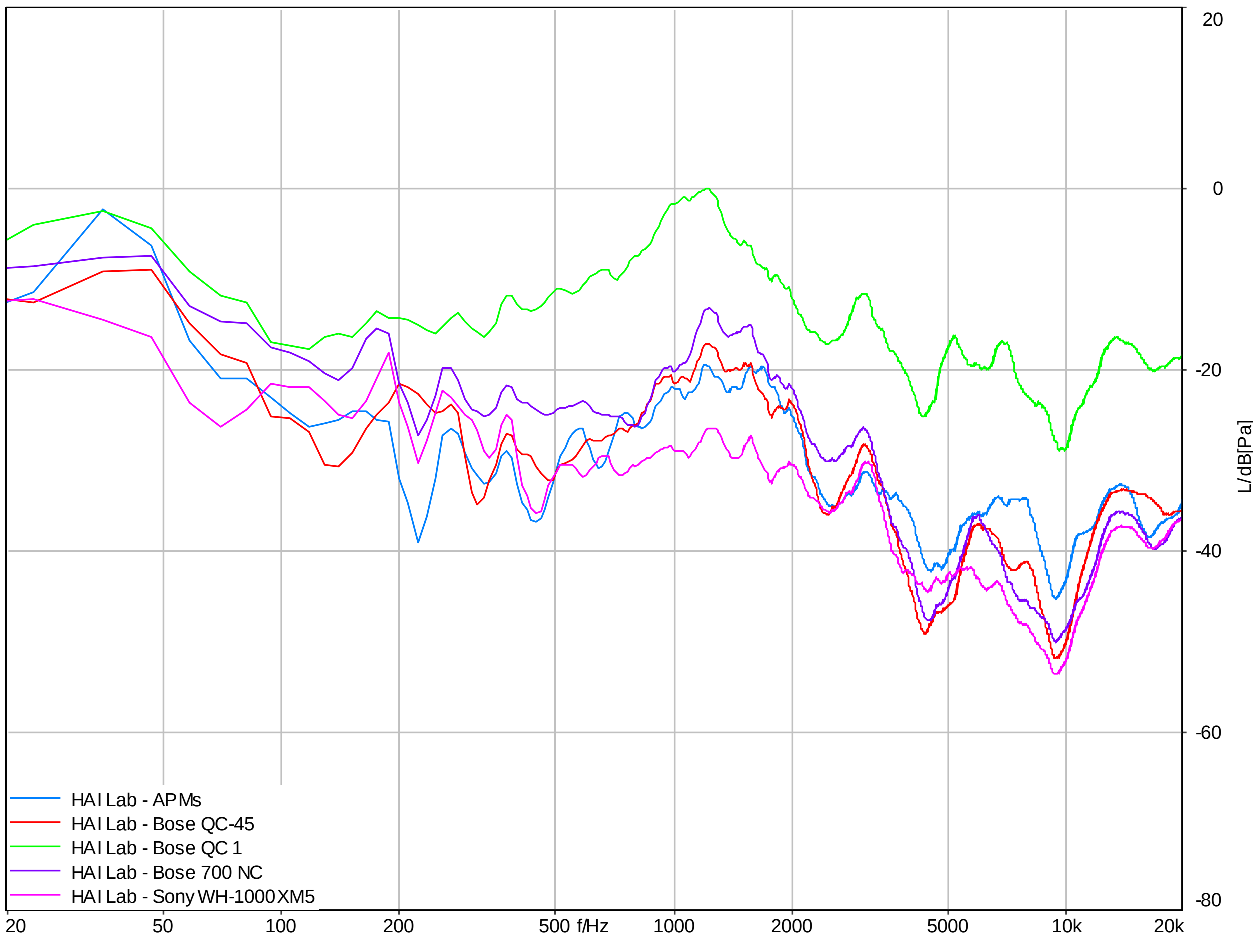
This noise environment is generally low frequency dominated, and the APMs and QC-45 offer very good ANC which helps with low frequency attenuation (in many places greater than 30dB attenuation).

The QC-1s are completely overmatched by the modern headphones.

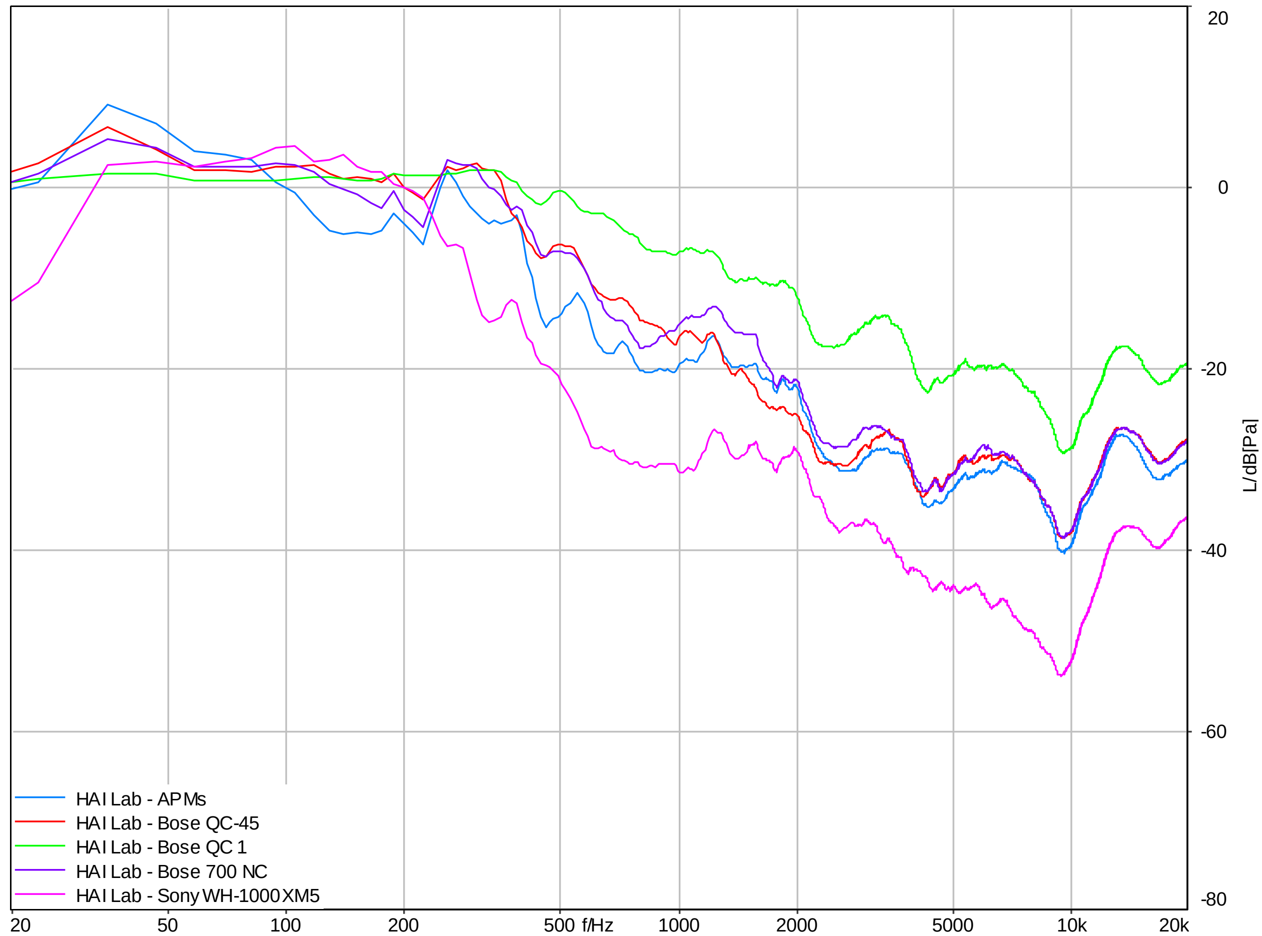
If we wanted to look at the active attenuation only, it would look like this:



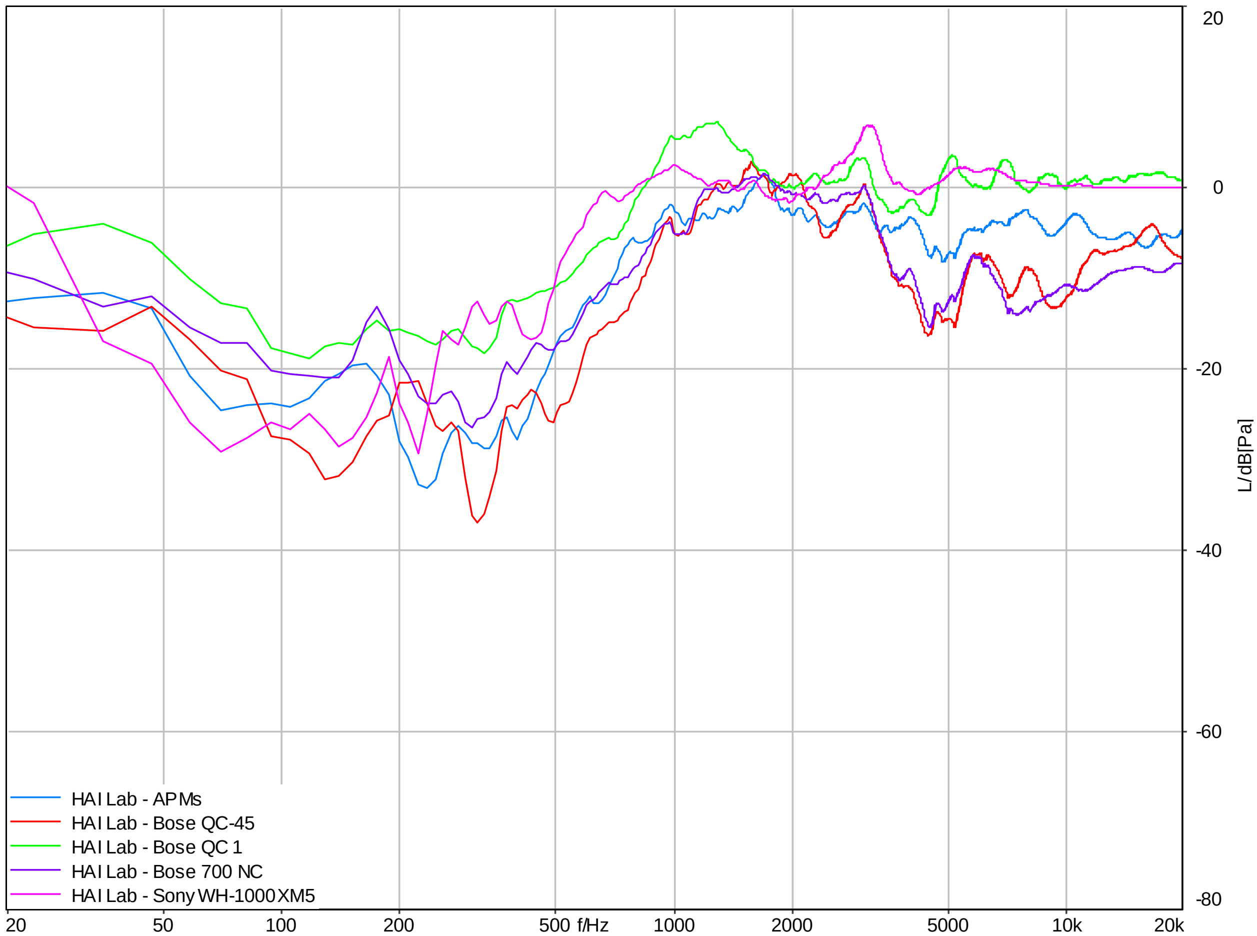
Next the max attenuation for BGN Cond #2 (Loud Music) is presented for each device:



The music environment – while containing lots of low frequency energy – also contains many more transient effects than the airplane cabin noise. This will stress the headphones differently, and developers may have to consider other techniques to improve overall ANC performance.

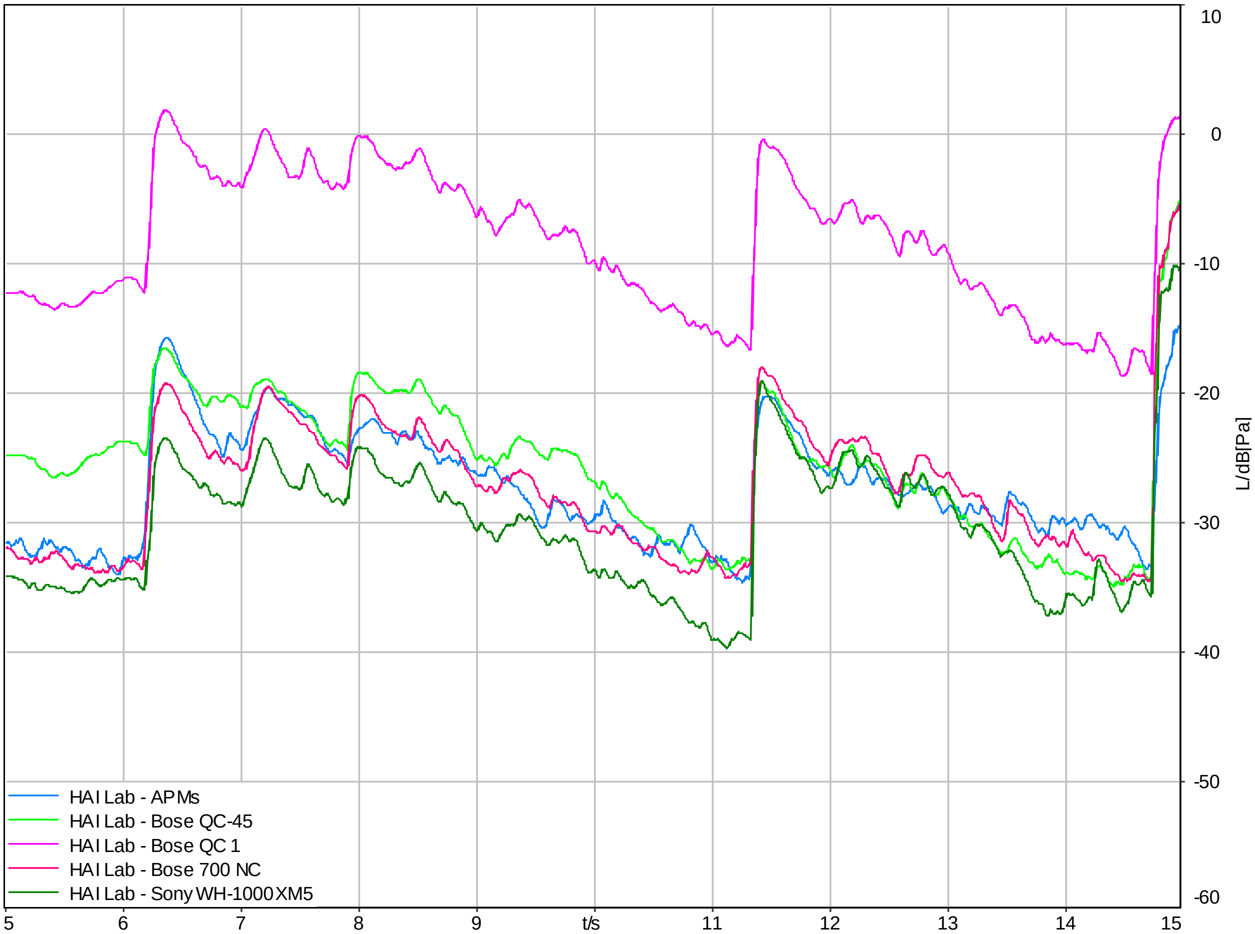


Passive attenuation for BGN Cond #2 (Loud music). The WH-1000XM5 have tremendous passive attenuation.



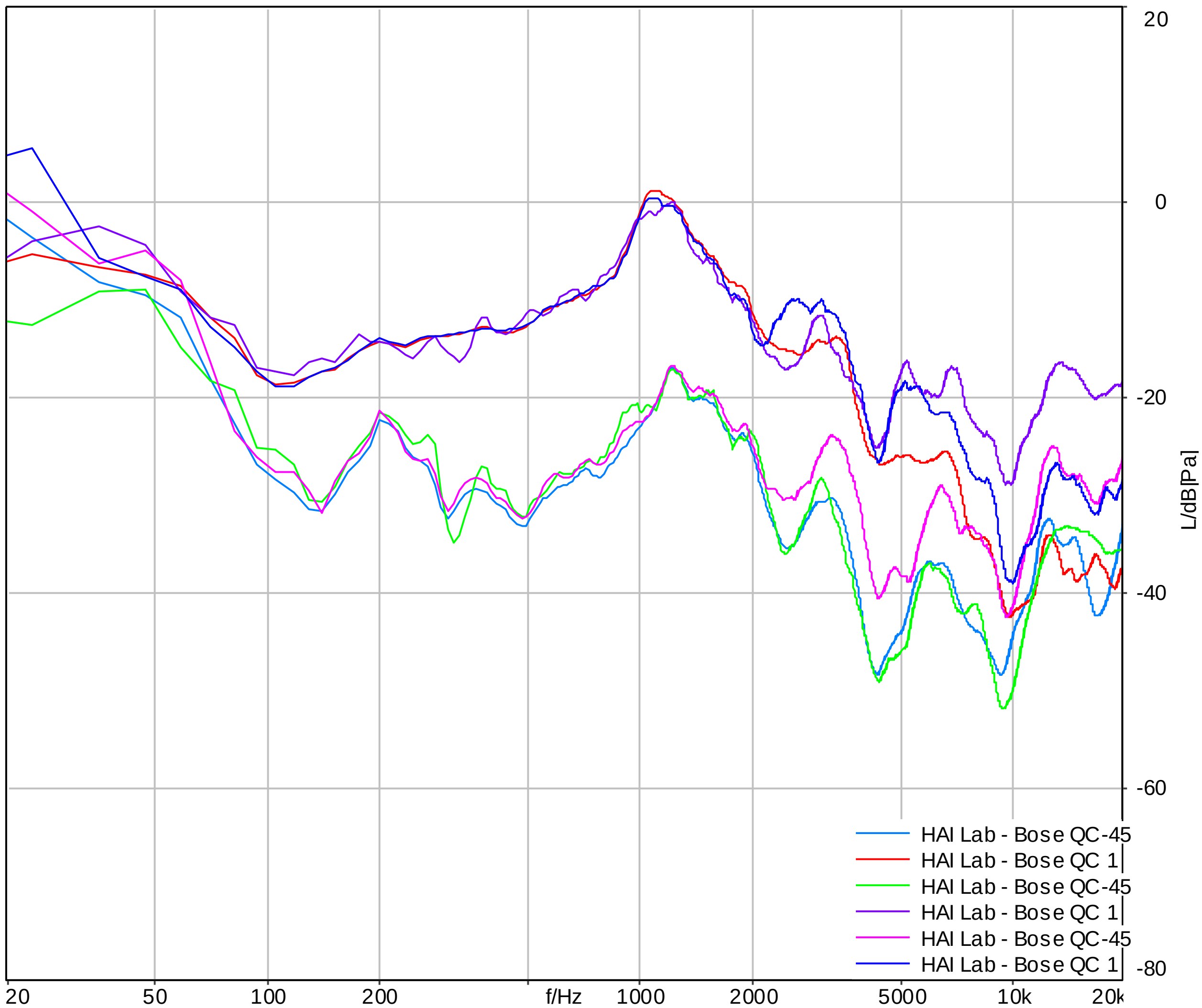
Active attenuation for BGN Cond #2 (Loud music). Notice the 3+kHz attenuation for the QC-45 and NC700 (and even APMs to an extent). That helps even the playing field in terms of total attenuation.

If we're interested in examining the headphones' performance across time, we can use a Level vs. Time plot. Here, for BGN Cond #2 (loud Music) we show a time section that features five distinct impulses (3 piano notes closely grouped at 6.1s, 7.1s and 8s, additional piano note at 11.3s, and base drop at 14.7s).



For next generation adaptive ANC headphones, this will be much more interesting to examine. At this point, none of the headphones really exhibit signs of improved attack or convergence time, indicated by how fast (and how far) the curve drops from its peak.

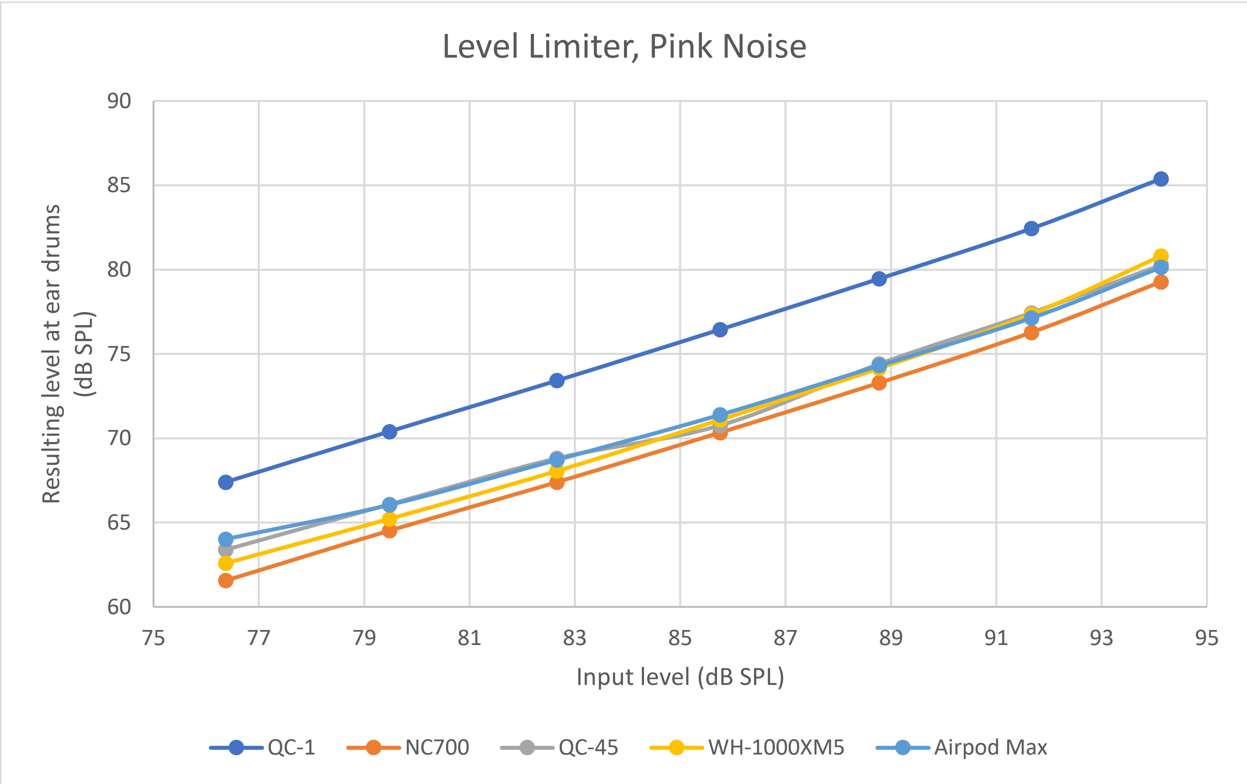
Another point of view that might be interesting is to see the evolution of devices within a brand. Here are the BOSE QC-1 (oldest) and QC-45 (newest) products' total ANC performance compared across 3 different noise scenarios:



It's pretty clear to see the improvement from one to the other. And while the ANC generally stays consistent, the higher frequencies tend to differ a bit depending on the noise environment.

Additionally, we looked at how linearly the devices behave as we increase the ambient noise. Apple makes claims about doing this in transparency mode, but it's clear they don't when ANC is set to max. As a basic approach, only Pink Noise was used as a noise source. Measurements were made at the ear drum with and without the device inserted.

If the devices are linear, we would expect a straight line increasing from bottom left towards top right.



In the plot above, all the devices are applying max ANC all the time (straight line).

With next gen Headphones, that advertise adaptive ANC, this is something we would expect to change. Especially in non-stationary noise environments.

Conclusions

A tremendous amount of data was collected in the test session. A summary of some initial conclusions is listed here:

1) If you're looking for the best headphones to block out the outside world, you can't go wrong with the Sony WH-1000XM5. It offers the best passive attenuation, and great ANC, resulting in sold performance in all noise environments.
2) A close second is the Apple AirPod Max. They take the lead in BGN Cond #5 (Airplane Cabin noise) with their superior ANC in the 100-500Hz range playing a big part in this low frequency dominated noise source.
3) The BOSE QC-45 comes in right behind the above two headphones.
4) Our psychoacoustic metrics to a large part support the above, but perhaps give a slight further advantage to the XM5s.
    a. Note that for BGN Cond #2 (Loud Music), all the headphones struggled, and our MOS metric rated the QC-1s in tied first place!! This is more an indictment of how difficult it is to suppress noise of this amplitude and character (everyone did poorly) than how good the QC-1s are.
5) When evaluating the speech quality of audio playback in the presence of these 5 noise environments, the clear favorites are the AirPod Max headphones. Even with ANC set to max, they do a great job of allowing the speech playback to come through undisturbed, unaltered, and sounding natural.
6) None of the headphones have truly adopted adaptive ANC features and the sophistication in terms of situational awareness.