On the applicability of the SBC codec to support super-wideband speech in Bluetooth handsfree communications

Nathan Souviraà-Labastie, Stéphane Ragot
Orange Labs, Lannion, France
Email: {nathan.souviraalabastie, stephane.ragot}@orange.com

***Abstract*—With the recent standardization of the Enhanced Voice Services (EVS) codec in 3GPP, mobile operators can upgrade their voice services to offer super-wideband (SWB) audio quality (with 32 kHz sampling rate). There is however one important use case which is currently limited by existing standards: handsfree communication with wireless headsets, car kits, or connected audio devices often rely on Bluetooth, and the handsfree-profile (HFP) in Bluetooth is currently limited to narrowband and wideband speech. Following the approach used to extend HFP to support wideband, we study in this paper the applicability of the SBC codec to further extend HFP to SWB. An evaluation of performance is provided taking into account Bluetooth system constraints.**

I. Introduction

The recent standardization of the 3GPP EVS codec [1] now makes it possible for mobile operators to offer FM radio-like audio quality in mobile conversational services, which is a significant improvement over the existing telephone-band or wideband (High-Definition, HD) voice quality. This opens the way to increased naturalness, speech intelligibility or listening comfort, for instance, by taking into account sounds like breath that can now be captured; music quality can also be clearly improved due to the state-of-the-art coding technology included in the EVS codec that can switch between specialized coding modes (speech or music).

Handsfree communication devices such as car kit or wireless headsets often rely on Bluetooth and they are currently limited to wideband (WB) (16 kHz sampling rate) as they do not support yet super-wideband (SWB) (32 kHz sampling rate) for such applications. The objective of this paper is to present the results from a feasibility study on extending the existing Bluetooth SBC codec for the purpose of supporting SWB communications in the Bluetooth Hands-Free Profile (HFP). The rest of the paper is organized as follows: Section II first reviews EVS and Bluetooth technical background before discussing novel ways to use SBC to support SWB. Section IV describes the experimental setup used in this work; results for WB and SWB experiments are presented in Sections V and VI, respectively. Section VII finally concludes the study.

II. Background and motivation

We here present the technical context of the EVS codec and Bluetooth Subband Codec (SBC), before listing some potential configurations of SBC to support super-wideband (SWB) in HFP.

A. Enhanced Voice Services (EVS) codec

The Enhanced Voice Services (EVS) codec standardized by 3GPP in Release 12 [1], [2], [3] has been primarily designed for Voice over LTE (VoLTE) and fulfills the following objectives:

- Enhanced quality and coding efficiency for narrowband (NB) and wideband (WB) speech services;
- Enhanced quality by the introduction of SWB and Full-band (FB) speech;
- Enhanced quality for mixed content and music in conversational applications - EVS is significantly better than previous voice codecs (AMR, AMR-WB);
- Improved robustness to packet loss and delay jitter;
- Backward compatibility to the AMR-WB codec

The EVS codec is specified in 3GPP TS 26.441 to 26.451. The usage of EVS in Voice over IMS (VoIMS) has been introduced in TS 26.114, and the corresponding profile for VoLTE in GSMA IR.92 has been updated in v9.0 to include the support of EVS; this implicitly extended the support of EVS to Voice over Wifi (VoWifi) in IR.51 and Video over LTE (ViLTE) in IR.94. The support of EVS in 3G has been specified in 3GPP Release 13 (see in particular 3GPP TS 26.453 and 26.454).

The EVS codec bitrates are composed of two different sets:

- EVS Primary modes including
  - Fixed bitrates from 7.2 to 128 kbit/s
  - Variable bitrate operation at an average bit rate of 5.9 kbit/s for active speech and about 7-8 kbit/s for music
  - Channel-aware mode at 13.2 kbit/s
- EVS AMR-WB IO modes which are identical to AMR-WB bitrates, from 6.6 to 23.85 kbit/s

In this work we focus on the SWB modes of EVS Primary, which are supported from 9.6 kbit/s to 128 kbit/s. Similar to the study in [4], which used AMR-WB at 23.85 kbit/s, we use the EVS-SWB mode at 24.4 kbit/s.

B. Bluetooth

Bluetooth is a communication standard [5], [6] for short distance data exchange, *e.g.* audio between electronic devices.

It is using the unlicensed 2.4 GHz band (ISM) also shared with Wifi. Each of the 79 channels has a bandwidth of 1 MHz ; time slots are 625 microseconds long. Two data rate modes are available: a mandatory mode, called the Basic Rate (BR) and an optional one, called the Enhanced Data Rate (EDR). The Basic Rate mode uses binary GFSK modulation and achieves data rate of 1 Mbps. The Enhanced Data Rate uses two types of modulation: π/4-DQPSK and 8DPSK that can reach data rates of 2 Mbps and 3 Mbps, respectively.

One main feature of a Bluetooth device is the different profiles it supports; audio applications require one of the following profiles:

- Hands-Free Profile (HFP) or Headset Profile (HSP) for conversational scenario
- Advanced Audio Distribution Profile (A2DP) for streaming scenario
- Hearing Aids (HA) profile

In this paper, we focus on HFP as our interest is the SWB conversational services using the EVS codec.

1) SBC: SBC [7], [8] is based on subband Adaptive Pulse Code Modulation (APCM) coding. The input signal is decomposed in 4 or 8 subbands using a critical sampled cosine modulated polyphase filterbank. The filter length is 40 or 80 with an analysis/synthesis delay of 37 or 73 samples by exploiting the polyphase structure. The filterbank is actually using the same structure as the MPEG-1 PQMF filterbank [9] with a shorter prototype filter. Each subband is normalized by scale factors and by block companding APCM quantization. The bit allocation per subband is determined based on the coded scale factors by either the SNR or the Loudness method which aims at minimizing distortion. Table 6.14 to Table 6.17 of the A2DP test specification [10] describe all possible SBC encoder settings used to verify the conformance (sampling frequency, channel mode, block length, number of subbands and method).

A bitpool value is recommended to ensure good quality but remains adjustable to obtain different bitrates. In mono channel mode (as in the rest of this paper), the bitpool is recommended to be set at 5 times the number of subbands at 16 kHz and 4 times at other sampling rates, hence ensuring a sufficient quality as mentioned in the A2DP test specification section 6.5.1.2.3 [10]. The formulae that compute the bitrate depending on these features and parameters can be found in section 12.9 of the A2DP specification [11]. In mono, the bitrate can be expressed in kilobits per second (kbit/s) as

$$bitrate = \frac{8 \times frame\_length \times fs}{subbands \times block\_length}$$

where $frame\_length$ is number of bytes per frame and equals

$$frame\_length = 4 + (4 \times subbands)/8 + \lceil nb\_blocks \times bitpool/8 \rceil.$$

2) mSBC: The modified SBC (mSBC) is a modification of the A2DP SBC coder with minimal changes (new block length and frame header). It has 8 subbands, a bitpool of 26, a block length of 15, hence yielding a codec bitrate of 60.8 kbit/s and a frame duration of 7.5 ms ($8 subbands \times 15 samples/16\ kHz$).

Packet Type	User Payload (bytes)	Symetric Max Rate (kbit/s)	T_{eSCO} for	
			96 kbit/s	128 kbit/s
EV3	1-30	96	4	
2-EV3	1-60	192	8	6
3-EV3	1-90	288	12	8

TABLE I: eSCO single-slot packets (EV3, 2-EV3 and 3-EV3).

It is the mandatory codec in HFP to support WB and was designed to fit in extended Synchronous Connection-Oriented (eSCO) packets.

3) eSCO logical transport: As SCO does not support retransmission and is not adapted to new codecs, the logical transport of mSBC been defined only with eSCO. We follow the same approach in this work. We hereafter focus on single-slot packet types, which is illustrated in Figure 1, and in particular on EV3 (BR), 2-EV3 (EDR) and 3-EV3 (EDR) packet types for which Table I provides a summary of the features. EV3 is the only single-slot eSCO mandatory packet type in the Bluetooth core. A T_{eSCO} period extends from an eSCO instant to another (see Figure 1) and is composed of an even number of slots with a minimum of 4 slots (a slot is 0.625 ms long). The maximum rate is obtained for the lowest allowed T_{eSCO} and the maximum payload size (see Table I). Small values of T_{eSCO} imply more frequent communication that involve less retransmission reliability and higher interferences with other radio communication like WiFi which may be a problem for the coexistence of Bluetooth and WiFi networks.

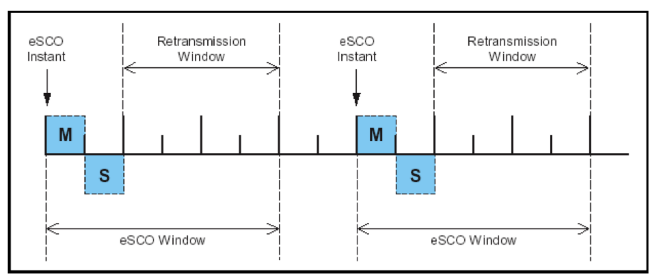

Fig. 1: Example of two eSCO windows for single-slot packets with $T_{eSCO} = 8$ (Figure 8.9 from Bluetooth specification [6]).

III. Potential codecs for SWB support in Bluetooth HFP

This section aims at discussing the codec that would be most appropriate to support SWB in HFP. We put forward some novel configurations of SBC at 24 and 32 kHz sampling rate and discuss their adaptation with eSCO packets, and the alignment in time of codec frames and logical transport packets. Table II gathers several options. The main features to align with eSCO are the codec rates and the frame durations:

- The codec rate should be lower than the eSCO maximum rate.
- The codec frame duration should be equal to the T_{eSCO} or at least being a multiple of T_{eSCO}.

	SBC codec					eSCO logical transport				
Codec configuration	**Sampling Rate**	**Block length**	**Bitpool (bits)**	**Frame duration (ms)**	**Rate (kbit/s)**	**Packet type**	**T_{eSCO} (slots)**	**T_{eSCO} (ms)**	**Rate (kbit/s)**	**Packets per codec frame**
mSBC	*16*	*15*	*26*	*7.5*	*60.8*	*EV3*	*6*	*3.75*	*64*	2
						2-EV3	*12*	*7.5*	*64*	1
A	32	30	21	7.5	92.8	EV3	4	2.5	96	3
						3-EV3	12	7.5	96	1
B	32	20	20	5	92.8	EV3	4	2.5	96	2
						2-EV3	8	5	96	1
C	24	15	27	5	94.4	EV3	4	2.5	96	2
						2-EV3	8	5	96	1
G.722.1C[12]	32	-	-	20	48	EV3	8	5	48	4
(or G.719[13])	(48)					2-EV3	16	10	48	2

TABLE II: Some SBC configurations (8 subbands) fitting in eSCO single-slot packets.

- Furthermore, it could be beneficial to also have alignment with the EVS 20 ms frames to reduce transcoding delays in super-wideband (SWB).

The mSBC configuration is given in Table II as a reference configuration fitting in eSCO. The mSBC frame duration is 7.5 ms, which corresponds to 12 eSCO slots of 0.625 ms. To fit in eSCO, some padding is added to increase the payload size to 30 bytes (resp. 60 bytes), *i.e.* 64 kbit/s with a T_{eSCO} of 6 (resp. 12) in EV3 (resp. 2-EV3) packet types. As a result, each mSBC frame is transported by two EV3 packet (7.5ms) in basic rate (see last column of Table II). It can be noticed that the number of packets used to encode one SBC frame does not change the eSCO resulting bitrate. G.722.1 Annex C [12] (G.722.1C) or G.719 [13] are also included as examples for prospective comparison (see the related discussion in Section VI-E).

Following the same logic, we propose two SBC configurations (noted B and C) which are aligned with T_{eSCO} and EVS frame duration at 96 kbit/s plus one configuration (noted A) that use the same frame duration as mSBC but with double block length (as sampling rate is doubled). It can be noticed that informal listening leads us to consider 96 kbit/s as a critical lower bound to ensure sufficient audio quality for SWB with SBC.

IV. Experimental setup

A. SBC codec

In this work, we used the SBC implementation from the BlueZ project [14]. Some minor source codec modifications were applied to improve the accuracy of filterbank calculations and pass all conformance tests specified in [10]. In addition, the codec was modified to support extra block lengths (20, 30). Note that the SBC implementation from Bluedroid [15] was also considered in this initial phase of this study, and verified to pass conformance tests with proper settings. However, we selected the BlueZ version because it already supported mSBC.

B. Setup for objective quality evaluation

A set of 30 speech signals of 8 seconds sampled at 48 kHz was used as input of tested codecs. This set is balanced between 3 male and 3 female speakers uttering 5 different double-sentences of 8 seconds in French language. The codec outputs are processed to set up different evaluations or speech specific frequency responses.

Perceptual Objective Listening Quality Assessment (POLQA) [16], [17] was used to estimate the Mean Opinion Score of the Listening Quality Objective (MOS-LQO). We used the solution from OPTICOM (version 2.4) with high accuracy mode and level adjustment in SWB mode. It can be noticed that POLQA scores are referred to as MOS-LQO$_s$ scores which range up to 4.5 for WB signals (up to 7 kHz bandwidth) and up to 4.75 for SWB signals. Moreover in this SWB mode, the reference signals always have a 32 kHz sampling rate even when the tested signals have other sampling rates. We also used the WB version [18] of Perceptual Evaluation of Speech Quality (PESQ) [19], [20] to compare our results (see Section V) with a previous study [4].

Informal listening of the codec outputs have been conducted in parallel to the objective experiments on SWB to confirm results (see Section VI).

C. Setup for subjective quality evaluation

For subjective tests, we used 16 test signals sampled at 32 kHz, divided into 4 items in 4 categories: clean speech, classical music (vocal, instrumental), modern music (vocal, instrumental), and mixed contents. Each of the stimuli was 7 to 10 seconds long.

V. Wideband experiments

In this section, we calibrate our experimental procedure on known wideband (WB) behaviour. We first examine the WB objective performance (POLQA) of the SBC codec under different recommended settings and in particular the mSBC setting. mSBC is then compared to G.722 while transcoding with AMR-WB. The objective is to establish a yardstick established in [4].

A. From SBC to mSBC

We hereafter call recommended configurations those described in section 6.5.1.2.3 of the A2DP test specification [10] using the bitpool recommended in section 12.9 of the A2DP specification [11]. These operating points define the different right tails of the curves in Figure 2. The rest of the curves are then obtained by reducing the bitpool (and hence the bitrate) starting from these points.

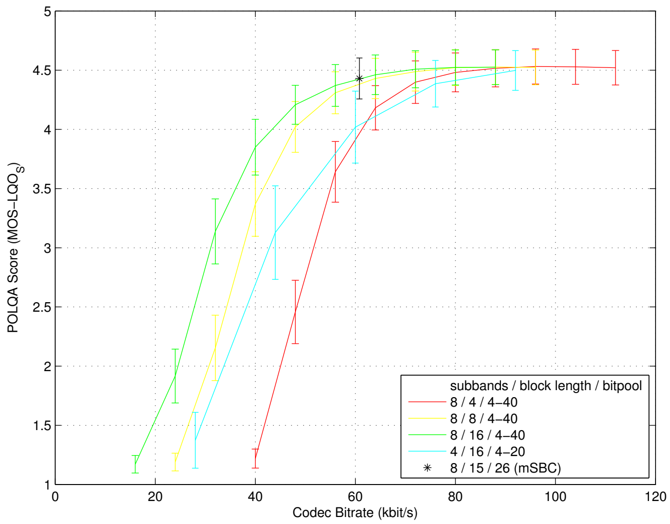

Fig. 2: POLQA scores (MOS-LQO$_s$) for SBC operating at 16 kHz (Wideband) with a bitpool ranging from 4 to recommended values with steps of 4 and the loudness bit allocation method.

Figure 2 is composed of mean POLQA scores and their standard deviations on the 30 speech samples described in Section IV-B. In addition to the above described curves, the mSBC result is represented in black. One can verify that scores decrease as bitrates is reduced and the recommended configurations always yield the maximum score, whereas mSBC is slightly under this maximum value. It can also be noticed that mSBC has similar performance to the configuration with a block size of 16 at 64 kbit/s and also with 8 subbands but a frame duration of 8 ms. It was also verified that the SNR bit allocation method leads to similar results and the configuration with a block size of 12 leads to a curve in between those of block sizes of 8 and 16.

B. SBC/AMR-WB transcoding at 16 kHz

In this section, we reproduce the study in [4] on AMR-WB transcoding evaluated with PESQ (see Table IIIa and Table IIIb). We then provide the results of POLQA on the same task (see Table IIIc). In [4], mSBC was compared to G.722 [21] (at 64 kbit/s and a block size of 1024 samples) and AMR-WB [22] (at 23.85 kbit/s and DTX off). mSBC and G.722 in double transcoding with AMR-WB are compared with the conclusion that mSBC “maintains a higher average speech quality level compared to G.722”. While our PESQ results are comparable with those in [4], the results from POLQA are less explicit as the original POLQA scores of G.722 and mSBC are close, hence leading to less differences on the transcoding.

VI. Super-wideband experiments

In this section, we examine the super-wideband (SWB) objective and subjective performance of the SBC codec under one novel and different supported settings. The corresponding

	Mean	Std	Max	Min
mSBC	4.46	0.05	4.56	4,35
G722	3.87	0.16	4.21	3.30
G722-AMRWB-G722	3.03	0.27	3.60	2.44
mSBC-AMRWB-mSBC	3.68	0.39	4.16	2.63

(a) PESQ scores obtained in [4].

	Mean	Std	Max	Min
AMRWB	3.61	0.42	4.05	2.33
mSBC	4.09	0.10	4.26	3.91
G722	3.70	0.12	3.87	3.39
G722-AMRWB-G722	3.01	0.36	3.48	1.91
mSBC-AMRWB-mSBC	3.56	0.39	3.93	2.42

(b) PESQ scores (MOS-LQO$_w$).

	Mean	Std	Max	Min
AMRWB	4.16	0.26	4.52	3.56
mSBC	4.43	0.18	4.73	4.04
G722	4.47	0.16	4.74	4.14
G722-AMRWB-G722	4.01	0.26	4.32	3.31
mSBC-AMRWB-mSBC	3.98	0.29	4.41	3.26

(c) POLQA scores (MOS-LQO$_s$).

TABLE III: Objective results of mSBC, G.722 and AMR-WB in an AMR-WB transcoding scenario

frequency responses are analyzed, indicating an insufficient response in high frequencies. SBC is then tested with EVS transcoding in SWB. Finally, a wider discussion is opened on other aspects.

A. SBC at 32 kHz

Figure 3 is constructed in the same way as Figure 2, except that SBC operates with SWB signals and two new block lengths (20 and 30) are tested. The evolution of the POLQA scores with bitrate of the two unsupported block lengths (20, 30) show a clear advantage compared to other block lengths for bitrates lower than 60 kbit/s, but almost no differences can be observed for bitrates higher than 96 kbit/s. However, informal listening tests demonstrate that a bit rate higher than 96 kbit/s is required to preserve frequency content above 14 kHz and at 96 kbit/s SBC reaches a critical point for audio quality in SWB (at any block length). POLQA accuracy should therefore be used with caution to attest SWB audio quality with SBC as it estimates almost perfect quality at 96 kbit/s for every block length.

B. Frequency response

In order to quantitatively confirm these informal observations, the frequency response of SBC is shown in Figure 4 for different bitpools and in Figure 5 for different configurations at 96 kbit/s (see Table IV for configuration details). Both figures represent empirical transfer functions and are obtained after concatenating the 30 speech samples and comparing the periodograms of codec input and output signals.

One can observe in Figure 4 a significant attenuation in the frequency response for bitpools from 2 to 24 (112 kbit/s), hence confirming informal listenings. We can also observe that at 96 kbit/s changing the block length to 20 (configuration B) or 30 (configuration A) does not lead to significant improvements compared with the supported block length of 16.

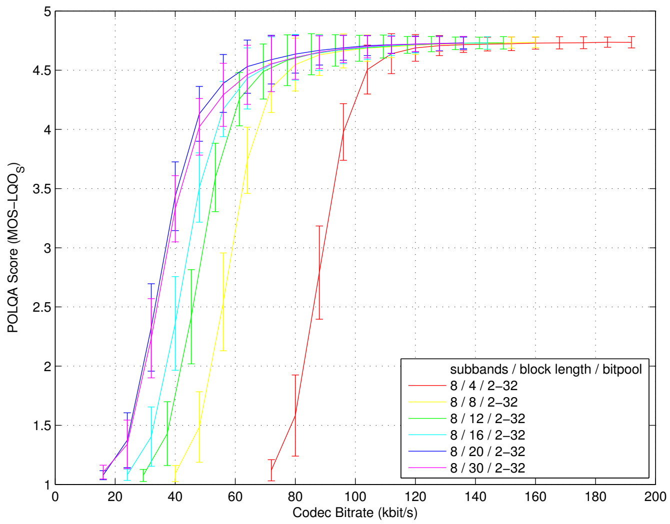

Fig. 3: POLQA scores (MOS-LQO$_s$) for SBC operating at 32 kHz (SWB) with a bitpool ranging from 4 to the recommended values with steps of 2.

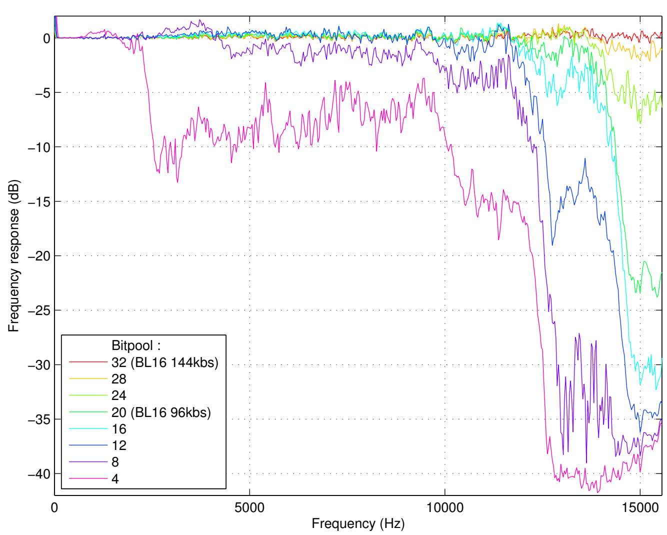

Fig. 4: Frequency response for SBC operating at 32 kHz, 8 subbands and a block length of 16 for different bitrates (induced by different bitpools).

C. SBC/EVS transcoding at 32 kHz

SBC is evaluated with POLQA on single transcoding with EVS (EVS-SBC and SBC-EVS) and double transcoding (SBC-EVS-SBC) (see Table IV). EVS is operating at 24.4 kbit/s with DTX on and three different configurations of SBC operating at 96 kbit/s are tested: one supported configuration for a block length of 16 (BL16 96kbs) and the two unsupported configurations (A and B) proposed in Section III (see Table II for details). One can note the recommended configuration for a block length of 16 (BL16 144kbs) and operating at 144 kbit/s that is given for comparison.

While the recommended configuration does not degrade the

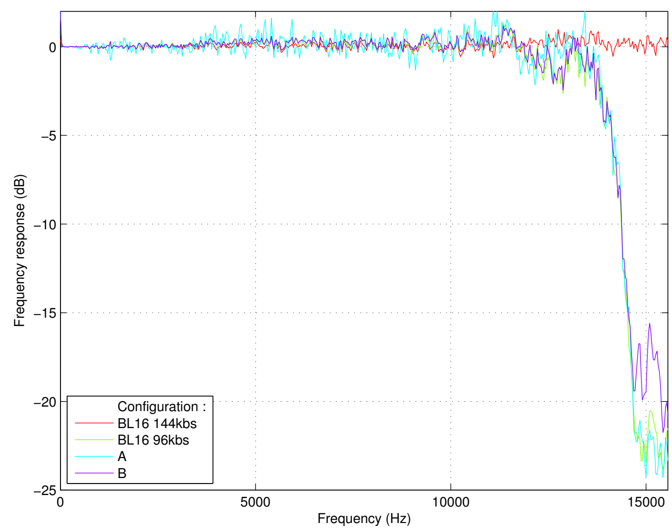

Fig. 5: Frequency response for SBC operating at 32 kHz for different configuration at 96 kbit/s (see Table II).

Configuration	**BL16 144kbs**	**BL16 96kbs**	**A**	**B**
Block length	16	16	30	20
Bitpool	32	20	21	20
Bitrate (kbit/s)	144	96	92.8	92.8
EVS	4.66	4.66	4.66	4.66
SBC	4.73	4.68	4.67	4.67
SBC-EVS	4.67	4.63	4.64	4.62
EVS-SBC	4.66	4.59	4.57	4.58
SBC-EVS-SBC	4.65	4.49	4.47	4.47

TABLE IV: POLQA scores (MOS-LQO$_s$) of different SBC configurations in EVS transcoding scenarios.

EVS scores, the three other configurations imply some degradation. All transcoding scenarios are affected, in particular the SBC-EVS-SBC case. Moreover, this last scenario does not accumulate linearly (in terms of POLQA scores) the distortion generated by the two single transcodings. While these SBC configurations have already been observed to be quality limited, this seems to reveal that SBC is even more poorly transcoding with EVS at 96 kbit/s. These observations have been confirmed by subjective listening evaluations reported in Section VI-D.

D. Subjective evaluation

In this section, we present the results of a subjective MUltiple Stimuli with Hidden Reference and Anchor (MUSHRA) [23] test conducted by four expert listeners. The listeners were asked to rate each test sample from 0 to 100; the different versions of the stimuli are:

- original sampled at 32 kHz, provided as a reference and repeated as an hidden reference (noted 'ref')
- a 7 kHz low-pass filtered version (noted '7kHz anchor')
- original coded with EVS at 24.4 kbit/s (noted 'EVS')
- original coded with configuration A (noted 'A')
- original coded with configuration B except that bitpool is 24 (noted 'B,bitpool=24')

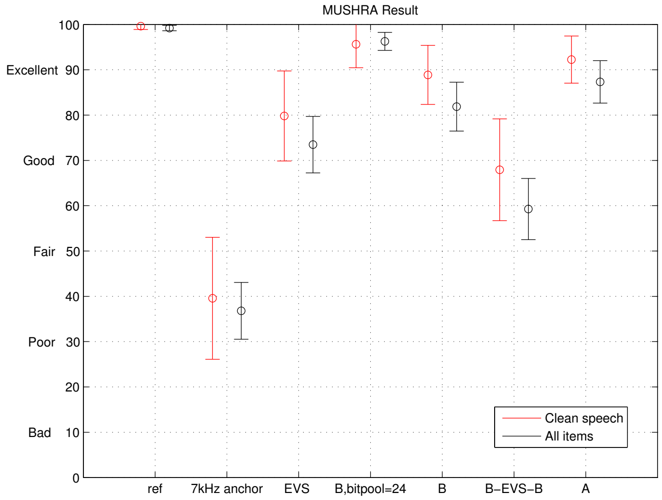

Fig. 6: Subjective test results.

- original coded with configuration B (noted 'B')
- original double transcoded with configuration B and EVS (24.4 kbit/s) - configuration B (noted 'B-EVS-B')

The mean results and 95% confidence intervals are gathered in Figure 6 separately for clean speech (4 samples) and all items (16 samples); the clean speech case is isolated for comparison with POLQA scores. Note that testing conditions did not include any bad condition (e.g. with packet losses or severe distortions); subjects were left free to use the grading scale, given labels ranging from Bad (0–20) to Excellent (80–100). Subjects used the grading scale in slightly different ways with regards to bandwidth limitation, some using almost the full scale and others not, hence scores did not always cover the same range and confidence intervals are quite large. Clean speech results show a similar trend as in POLQA scores in Table IV, except that POLQA tends to over-estimate the relative quality of tested conditions. Subjective test results confirm that the selected SBC conditions have better quality than EVS at 24.4 kbit/s, but the SBC-EVS-SBC condition is worse than EVS.

E. Discussion on complexity, delay, and sampling rate

While SWB experiments mainly focused on audio quality and frequency responses, other factors are important in codec design. For instance, the different types of induced delay (codec delay, latency due to eSCO logical transport, transcoding delay) should be taken into account. For instance, the use of two EV3 packets to encode one mSBC frame induces a eSCO logical transport delay of 7.5 ms as well as an extra latency due to retransmission cycles. Codecs like G.719 or G.722.1C induce important delays compared to SBC, their complexity and memory footprints are also of concern even if they are known to be small.

In the case of 24 kHz sampling rate (e.g. configuration C in Table II), an SBC block length of 15 or 30 would induce a 5 or 10 ms frame, which is a submultiple of the EVS frame length (20ms); it would also yield to a lower complexity and good audio quality around 96 kbit/s. This solution would be technically elegant, however, this lower sampling rate implies an audio bandwidth limitation of 12 kHz that does not match any of the audio bandwidths natively supported by EVS; it would not comply with expected SWB acoustic frequency response, and one may expect potential transcoding artefacts in the SBC-EVS-SBC transcoding case.

VII. Conclusion

The different experiments presented in this paper show that the SBC codec is not fully adapted to extend the HFP Bluetooth profile to SWB, in particular for eSCO on basic rate (BR). The new SBC configurations (A and B) at 96 kbit/s have insufficient frequency response in high frequencies and SBC-EVS-SBC transcoding induces artefacts that are also detected by objective evaluation. Compared with mSBC, the bit rate of 96 kbit/s would push eSCO transport (in basic rate) to its limit, which may bring issues of coexistence with other systems (e.g. Wifi). A work-around would be to use SBC in enhanced data rate (EDR), and in this case higher bit rates than 96 kbit/s would ensure sufficient performance.

References

[1] M. Dietz and al., "Overview of the EVS codec architecture," in Proc. ICASSP, pp. 5698–5702, April 2015.
[2] 3GPP TS 26.441, "EVS Codec General Overview, version 12.0.0," 2014.
[3] S. Bruhn and al., "Standardization of the new 3GPP EVS codec," in Proc. ICASSP, pp. 5703–5707, April 2015.
[4] W. Kargus, J. Spiewla, G. Spittle, X. Sun, and W. Zuluaga, "Objective Evaluation of Wideband Speech Codecs for Voice Communication Over Bluetooth," in Audio Engineering Society Convention 129, Nov 2010.
[5] Bluetooth SIG, "The Bluetooth Core specification (vol.1), v4.2.," 2014.
[6] Bluetooth SIG, "The Bluetooth Core specification (vol.2), v4.2.," 2014.
[7] F. de Bont, M. Groenewegen, and W. Oomen, "A high-quality audio coding system at 128 kb/s," in Audio Engineering Society Convention 98, Feb. 1995.
[8] C. Hoene and M. Hyder, "Optimally using the Bluetooth subband codec," in Proc. IEEE 35th Conference on Local Computer Networks (LCN), pp. 356–359, Oct 2010.
[9] M. Bosi and R. E. Goldberg, Introduction to Digital Audio Coding and Standards. Kluwer Academic, 2002.
[10] Bluetooth SIG, "Advanced Audio Distribution Profile (A2DP) v1.0-1.3.1, Test Suite Structure (TSS) and Test Purposes (TP)," 2015.
[11] Bluetooth SIG, "Advanced Audio Distribution Profile v1.3.1," 2015.
[12] ITU-T Rec. G.722.1, "Low-complexity coding at 24 and 32 kbit/s for hands-free operation in systems with low frame loss," May 2005.
[13] ITU-T Rec. G.719, "Low-complexity, full-band audio coding for high-quality, conversational applications," 2008.
[14] Standalone SBC library (v1.3). http://www.bluez.org/sbc-13.
[15] Bluedroid source code. https://android.googlesource.com/platform/external/bluetooth/bluedroid.
[16] ITU-T Rec. P.863, "Perceptual Objective Listening Quality Assessment," 2014.
[17] ITU-T Rec. P.863.1, "Application guide for Recommendation ITU-T P.863," 2014.
[18] ITU-T Rec. P.862.2, "Wideband Extension to Recommendation P.862 for the Assessment of Wideband Telephone Networks and Speech Codecs," 2005.
[19] ITU-T Rec. P.862, "Perceptual Evaluation of Speech Quality (PESQ), An Objective Method for End-to-End Speech Quality Assessment of Narrowband Telephone Networks and Speech Codecs," 2001.
[20] ITU-T Rec. P.862.1, "Mapping Function for Transforming P.862 Raw Result Scores to MOS-LQO," 2003.
[21] ITU-T Rec. G.722, "7kHz audio-coding within 64 kbit/s," 1988.
[22] ITU-T Rec. G.722.2 Annex C, "14 kHz mode at 24, 32, 48 kbit/s," 2003.
[23] ITU-R BS. 1534-3, "Method for the subjective assessment of intermediate quality levels of coding systems," 2015.