**Abstract:**
Otoacoustic emission (OAE) measurement is a sensitive and effective technique to monitor changes in the inner ear potentially induced by noise exposure. However, outside a controlled testing environment, measurements are challenging since the level of ambient noise might be higher than the low-level OAE response. Therefore, an OAE system was designed, suitable to measure OAEs repeatedly on an individual worker in noisy test conditions. This system features a left and right earpiece, each equipped with a pair of miniature loudspeakers, an external and an internal microphone. In addition to the passive attenuation of the earpiece, adaptive filtering on the distortion product OAE (DPOAE) signals is used to further enhance the ratio between the measured OAE signal and interfering noise. The adaptive filtering technique uses the sound captured by the ipsilateral external and internal microphones as well as from the contralateral internal and external microphones. In this paper, the accuracy of DPOAE signals are studied using different combinations of the four microphones in single and dual adaptive filter topologies, as well as the benefits of adding a fixed transfer function in the adaptive filtering algorithm topology to estimate the acoustic path. Side-by-side comparison shows that a dual-stage adaptive filtering algorithm, using a combination of the contralateral internal microphone with the ipsilateral external microphone, is the most promising approach to denoise the DPOAE signal.
Design considerations for robust noise rejection in Otoacoustic Emissions measured in-field using adaptive filtering

Vincent Nadon\textsuperscript{a,b}, Annelies Bockstael\textsuperscript{b}, Dick Botteldooren\textsuperscript{b}, Jean-Marc Lina\textsuperscript{a}, Jérémie Voix\textsuperscript{a,\ast}

\textsuperscript{a}École de technologie supérieure, 1100, Notre-Dame Ouest, Montreal, Quebec, Canada, H3C 1K3
\textsuperscript{b}Ghent University, WAVES research group, Technologiepark Zwijnaarde 15, B-9052, Ghent, Belgium

Abstract

Otoacoustic emission (OAE) measurement is a sensitive and effective technique to monitor changes in the inner ear potentially induced by noise exposure. However, outside a controlled testing environment, measurements are challenging since the level of ambient noise might be higher than the low-level OAE response. Therefore, an OAE system was designed, suitable to measure OAEs repeatedly on an individual worker in noisy test conditions. This system features a left and right earpiece, each equipped with a pair of miniature loudspeakers, an external and an internal microphone. In addition to the passive attenuation of the earpiece, adaptive filtering on the distortion product OAE (DPOAE) signals is used to further enhance the ratio between the measured OAE signal and interfering noise. The adaptive filtering technique uses the sound captured by the ipsilateral external and internal microphones as well as from the contralateral internal and external microphones. In this paper, the accuracy of DPOAE signals are studied using different combinations of the four microphones in single and dual adaptive filter topologies, as well as the benefits of adding a fixed transfer function in the adaptive filtering algorithm topology to estimate the acoustic path. Side-by-side comparison shows that a dual-stage adaptive filtering algorithm, using a combination of the

\textsuperscript{\ast}Corresponding author's email address: jeremie.voix@etsmtl.ca
Phone number: 1-514-396-8437
Email addresses: vincent.nadon@etsmtl.ca (Vincent Nadon), annelies.bockstael@intec.ugent.be (Annelies Bockstael), dick.botteldooren@intec.ugent.be (Dick Botteldooren), jean-marc.lina@etsmtl.ca (Jean-Marc Lina)
contralateral internal microphone with the ipsilateral external microphone, is the most promising approach to denoise the DPOAE signal.

Keywords: adaptive filtering, noise rejection, noise-induced hearing loss, otoacoustic emissions, hearing health monitoring, digital signal processing

1. Introduction

Risk of noise-induced hearing loss (NIHL) due to occupational noise exposure is often reduced using personal hearing protection devices (HPD). In theory passive hearing protectors are designed to reduce noise levels sufficiently, but in practice, in the workplace, providing personal protectors does not always guarantee prevention of hearing loss. One factor of uncertainty is the exposure level under the hearing protector. Different methods are available to assess passive noise reduction individually [1, 2], but measurements rarely account for variability in noise exposure over an entire working day, which largely depends on proper fit and consistent use of the protector [3, 4]. In addition, susceptibility to NIHL is known to vary among individuals [5]. Therefore, the maximum noise exposure an individual can tolerate before suffering from permanent hearing damage is unknown.

These problems could be tackled by providing a special type of HPD to workers that is able to monitor subtle changes in cochlear status, potentially related to noise exposure, throughout a working day. Such a device would warn the worker or his superior in real-time when a (temporary) change in the worker’s inner ear dynamics is detected. To monitor the daily effect of noise exposure an in-ear hearing protection device featuring otoacoustic emission (OAE) monitoring is a promising solution.

Distortion product OAEs (DPOAEs) are measured by sending two pure tone stimuli to the ear, $f_1$ and $f_2$ with a $f_2/f_1$ ratio of 1.22. Low-level cubic distortion signals (i.e. $f_{dp} = 2f_1 - f_2$) are generated by the active non-linear process of the outer hair cells (OHC), taking place inside the inner ear. The distortion product responses travel back from the inner ear to the outer ear and they can be recorded by an in-ear microphone (IEM) placed inside the ear canal. DPOAE levels are associated with the health of OHC in the cochlea, and these structures are known to be vulnerable to noise expo-
sure [6]. After a certain duration of noise exposure, the OHC inside the inner ear may be fatigued or damaged generally leading to lower amplitude of DPOAEs than when they were healthy, for instance before the exposure [7]. While the measured DPOAE signals are not systematically related to the individual absolute hearing thresholds as assessed with a traditional audiogram, scientific evidences suggests that these individual DPOAE can be used to track the relative change in the inner ear dynamics[6, 8]. Such tracking was demonstrated in Engdahl’s study [7] with the relationship between temporary threshold shift (TTS) variations following noise exposure and DPOAE variations, compared to the pre-exposure baseline. Moreover, Sutton [9] demonstrated the effect of the noise exposure and the recovery progress of DPOAE levels in the following minutes after noise exposure. Therefore, lower DPOAE levels in individuals following a certain dose of noise exposure may indicate a greater susceptibility to NIHL. However, a larger epidemiological study might eventually be necessary with the close DPOAE monitoring method since according to Müller [10], permanent threshold shift (PTS) and TTS can occur due to different ear mechanisms and therefore the relationship might not be as straightforward.

To detect such variations in DPOAE levels, the measurements need to be conducted shortly after the noise exposure. However, measuring DPOAEs outside a controlled environment —i.e. for in-field hearing status close monitoring— is currently strongly hampered by interfering environmental noise [11], as typical DPOAE sound pressure levels (SPL) fall between -20 dB and 20 dB depending on stimuli levels and health of OHC [12]. Therefore, an ambient noise level of 70 dB(A) could make the DPOAE undetectable due to the low signal-to-noise ratio (SNR) [13].

Standard OAE probe tips provided with commercial systems are pre-moulded, that is, not optimized for each individual’s ear canal. Therefore they usually provide a limited amount of passive attenuation, which is not suitable for noisy industrial test environments [14]. To further improve the SNR, software solutions like standard sample rejection techniques based on noise threshold levels and time averaging methods [15] can be applied in case of limited disturbance [12]. However, these techniques were shown to be insufficient in more realistic occupational noise settings such as in industrial environments [13], especially for the range below 1500 Hz [12]. It is important
to minimize the error in the DPOAE level induced by the noise in the averaging algorithm, so that it is smaller than the real DPOAE level fluctuations in response to the noise exposure. Otherwise, the real fluctuations caused by the inner ear changes would be undetectable.

In response to these problems encountered with time averaging methods, several adaptive filtering techniques use a contralateral in-ear microphone (IEM-C) as a physiological noise reference and an ipsilateral outer ear microphone (OEM-I) as an external background noise reference to remove the noise captured with the microphone in the tested ear [12, 16–20]. The adaptive filtering noise rejection algorithm increases the SNR for the entire frequency spectrum of interest while reducing the test time needed compared to standard time averaging methods. Compared to the other multimicrophone adaptive filtering studies [12, 16], the topology presented in the current paper has the main advantage of relying on two identical OAE probes each including the same internal and external microphones giving the possibility of combining the 4 microphone signals available. Moreover, the external microphones are integrated within the OAE probes so that a greater spatial coherence between external (OEM) and in-ear (IEM) noise signals can be achieved. In previous studies, the number of combinations was limited with the maximum of 3 microphones used [12, 16], therefore the adaptive filtering topologies did not have as much flexibility as the proposed method.

The aim of the current paper is to compare different microphone combinations and algorithm configurations of adaptive filters to enhance the DPOAE measurements in noise conditions. Comparisons are first done in White noise conditions and then the optimal configuration is validated for more realistic noise scenarios by testing the adaptive filtering noise rejection algorithm in Industrial conditions. The optimal configuration was also tested in Quiet conditions to make sure DPOAE signals were not altered.

2. Description of the OAE system with noise rejection

In order to determine the optimal microphone combination for robust noise rejection of external and internal noise, the signals from OEM-I, OEM-C and IEM-C captured in human subjects using the OAE system detailed below were post-processed in
various combinations. The adaptive filter topology used was either with a single-stage (Stage 1 only) or a dual-stage (Stage 1 and Stage 2), more details are found in the following section. After processing with the various noise rejection algorithms, DPOAE levels and residual background noise levels were obtained with the signal extraction algorithm.

2.1. OAE recording equipment

In order to measure in-field DPOAE response in presence of high sound pressure levels industrial noise, a complete earpiece-embedded OAE probe was designed. For this designed earpiece, two high-quality miniature balanced armature loudspeakers were used to send the two pure-tone stimuli, one loudspeaker per stimulus, to eliminate possible interference between stimuli for minimal sound distortion reducing the risk of measuring a false DPOAE. These loudspeakers, Model WBFK from Knowles (Itasca, IL, USA), have a wide-band frequency range of 0.1-8 kHz with resonance frequencies around 4 kHz and 7 kHz according to the manufacturer’s specifications based on measurements in an IEC711 type 1 coupler. One miniature microphone is placed towards the ear canal in order to measure the DPOAE response and residual noise within the ear canal. A second miniature microphone is placed on the outer face of the earpiece to measure the external background noise (see Fig. 1). In order to seal the ear canal for proper DPOAE measurements and also to protect the human subject when measuring in higher levels of background noise, roll down foam ear tips were mounted on the two probes.

Stimuli to evoke the DPOAE response in the left and right earpieces were generated from the Auditory Research Platform (ARP) [21]. The pairs of two stimulus tones $f_1$ and $f_2$ are presented over the entire frequency range looped in descending frequency order, from $f_2 = 6169$ Hz to $f_2 = 1000$ Hz to match the frequency range of the clinical device previously used as a reference. Eight points per octave band were collected for a total of 22 pairs of primary tones. The stimuli are set with a $f_2/f_1$ ratio of 1.22 with sound pressure levels of $L_1 = 65$ dB and $L_2 = 55$ dB [22].

The ARP was set to send the stimuli with a rise and fall time of 200 ms [14]. According to the recommendations in previous studies [12], the continuous stimulation
period should be longer than a few hundred milliseconds in order to provide enough time for the adaptive filtering noise rejection algorithm to converge, thus giving maximum performance for the algorithm. The duration of the plateau of steady stimulation of the designed system was set at 1.4 seconds. With this duration, sufficient samples were collected for the adaptive filtering within a more than reasonable time for total test duration of about 1.5 minutes. Measurements with the developed system were stopped after the complete frequency range had been acquired twice in Quiet conditions and four times in noisy conditions, but signals were post-processed only for two repetitions in this study.

The data acquisition system setup includes a PXI-4462 card (National Instruments, TX, USA) used for the acquisition of the four microphone signals. The PXI-4462 provides accurate time synchronization between microphone signals with sampling frequency set at 48 kHz. The data acquisition card is connected to a desktop computer equipped with MATLAB® (Mathworks, MA, USA) to capture and post-process the signals using various algorithm configurations. This way, different approaches could be systematically compared in their ability for DPOAE signal identification and noise rejection.
2.2. OAE noise rejection algorithm

Figure 2 illustrates the adaptive filtering noise rejection algorithm used. A Finite Impulse Response (FIR) band-pass filter with centre frequency at $2f_1 - f_2$ and a bandwidth of 200 Hz was used to filter out the desired signal from the tested ear’s IEM signal. The same filter is also used on the other microphone signals (see OAE enhancer in the block diagram of the proposed noise rejection algorithm in Fig. 2) for homogeneity in the subsequent signal processing.

2.2.1. Fixed transfer function

Based on typical implementation of noise control algorithms [23 [24], a precise transfer function with fixed coefficients $\hat{H}(z)$, modelling the probe’s attenuation transfer function $H(z)$, was combined with an adaptive filter in the signal processing scheme. The $\hat{H}(z)$ transfer function is used to predict the contribution of the ambient background noise disturbance on the internal microphone, based solely on the input of the external microphone. This way, the high number of fixed coefficients would minimize the error and the adaptive coefficients could compensate quickly for slight variations in the probe attenuation transfer function over time [23].
To identify the probe attenuation transfer function $H(z)$ for each subject, an adaptive filter is set up in an offline system identification configuration, as illustrated in Fig. 1 where OEM and IEM microphones are exposed to a high level wide-band noise source, i.e. white noise. Eq. 1 represents the filtering process of the reference noise signal $x(n)$ with the filter coefficients $\hat{w}$ to obtain the filtered signal $y(n)$ which will ideally be similar to the desired signal $d(n)$ to minimize the error, calculated with Eq. 2, and therefore identify the system after adaptation of the filter coefficients with Eq. 3 based on the convergence step size, the error signal $e(n)$ and the reference signal $x(n)$. Here a step size $\mu = 0.5$ was used to calculate the fixed coefficients of the FIR filter. The number of coefficients for the FIR filter was empirically set to $N = 600$ as a compromise between filter length and improvement in transfer function accuracy. This high filter order ensured a precise transfer function over a wide frequency range.

To verify the proper identification of the transfer function $H(z)$, the transfer function $\hat{H}(z)$ estimated with the offline adaptive identification, was compared with the target probe transfer function $\hat{H}'(z)$ estimated from the difference between the ipsilateral OEM and IEM power spectra here used as a benchmark. Although the fixed transfer function could solely be estimated from spectrum differences between the OEM and the IEM in a real-time processing scheme, such an approach would be computationally more complex and therefore unsuitable for a low-cost DSP implementation. As observed in Fig. 3, the transfer function $\hat{H}(z)$ and its phase response estimated with the adaptive identification process is in agreement with the target transfer function $\hat{H}'(z)$. Only the left probe transfer function is shown in Fig. 3 as only this probe was used for DPOAE measurements on test subjects.

To include the probe attenuation transfer function as a fixed transfer function in the signal processing, different topologies can be used as seen in standard implementations of acoustic noise control algorithms [23, 24]. In this article, two placements of the transfer function estimate $\hat{H}(z)$ with the adaptive filter were tested: the fixed transfer function placed in parallel with the adaptive filter (see Fig. 2) and in series before the adaptive filter (see Fig. 4).
Figure 3: a) magnitude and b) phase wrapped on 1080° (6π) of the estimated transfer function \( \hat{H}(z) \) with the adaptive system identification configuration and the target probe attenuation transfer function \( \hat{H}'(z) \) for two subjects, here shown as an example. The root-mean-square error (RMSE) for the Subject 1 and Subject 2 transfer function magnitudes are respectively 5.64 dB and 1.91 dB over the frequency range of interest of 0-8 kHz.

Figure 4: The fixed transfer function placed in series with the adaptive filter
2.2.2. Adaptive Filtering Noise Rejection processing

The Normalized Least-Mean Squares (NLMS) adaptive filter in Stage 1 (Fig. 2), using the adaptive filter’s Eq. 1 with $i = 1$, models the transfer function between the contralateral IEM $x_1(n)$ and the tested ear IEM $d_1(n)$, where $n$ represents the samples, which represents acoustical differences between the two ear canals:

$$y_1(n) = \hat{w}_1^T(n)x_1(n)$$  \hspace{1cm} (1)

$$e_1(n) = d_1(n) - y_1(n)$$  \hspace{1cm} (2)

$$\hat{w}_{i=1,2}(n + 1) = \hat{w}_i(n) + \mu \frac{e_i(n)x_i(n)}{x_i^T(n)x_i(n)}$$  \hspace{1cm} (3)

In Eq. 3 with $i = 1$, the error signal $e_1(n)$ is used to correct the adaptive filter’s coefficients represented by the vector $\hat{w}_1$ in order to model accurately the residual noise disturbance in the tested ear’s IEM. The noise disturbance considered here for Stage 1 is mainly caused by physiological noise arising from vital functions like breathing and heartbeats. In Eq. 3 where $x^T(n)$ is the transposed form of the noise reference signal $x(n)$, $i = 1$ represents Stage 1 (Fig. 2).

In Stage 2 ($i = 2$) in Eq. 3 of the dual adaptive filter topology, the fixed transfer function $\hat{H}(z)$, previously estimated offline, models the physical transfer function $H(z)$ of the probe in an occluded ear canal. This transfer function is combined with the Stage 2 adaptive filter, so that the adaptive filter compensates for slight variations in the fixed $\hat{H}(z)$ transfer function over time due to variations in the earplug attenuation.

In Eq. 3 with $i = 2$, the error signal $e_2(n)$ is used to correct the adaptive filter’s coefficients ($\hat{w}_2$) in order to model the ambient noise disturbance captured by the tested ear’s IEM. The error signal $e_2(n)$ calculated from the difference between the desired $d_2(n)$ and the fixed filter $y'_2(n)$ as well as adaptive filter $y_2(n)$ output signals in Eq. 4 is also used as the output of the noise rejection algorithm since it mainly consists of the DPOAE signal without the noise signals. Here $y_2(n)$ is the ambient noise $x_2(n)$ filtered with the adaptive filter in Eq. 5 and $y'_2(n)$ is this same ambient noise filtered with the
fixed transfer function in Eq. [6]. The proposed adaptive filtering approach implies that the OAE response is not picked up by the OEM or contralateral IEM, which is normally the case since the DPOAE responses are unilaterally generated and the probes provide high passive attenuation, therefore they seal the low level DPOAE responses within the tested ear.

\[ e_2(n) = d_2(n) - y_2(n) - y'_2(n) \]  \hspace{1cm} (4)

\[ y_2(n) = \hat{w}_2^T(n)x_2(n) \]  \hspace{1cm} (5)

\[ y'_2(n) = \hat{w}'_2^T(n)x'_2(n) \]  \hspace{1cm} (6)

**Adaptive filter parameters selection.**

As expected from adaptive filters theory [25], a small step size (\( \mu \)) to adjust the coefficients in the adaptive filters (Eq. [3]) makes the adaptive filter converge slowly, but the error \( e(n) \) resulting from the difference between the desired signal \( d(n) \) and the transfer function output \( y(n) \) is smaller. Therefore, less error is induced in the DPOAE signal magnitude extraction.

The step sizes of the adaptive filters used to process the data were first calculated following the methods shown in [25]. Afterwards, for both stages of the proposed adaptive filtering noise rejection algorithm (Fig. 2), the step size \( \mu \) was optimized per subject in an adaptive procedure by starting from the smallest step size \( 1 \times 10^{-5} \) and iteratively increasing until no additional noise reduction effect was observed. The observed effective range for the step size \( \mu \) was always found to be below 0.7.

The step size for Stage 1 ended up being the most sensitive parameter. For cases when the probe did not provide sufficient attenuation and therefore the adaptive filtering algorithm was more sensitive, the step size of Stage 2 was increased up to \( \mu = 0.5 \), otherwise it was left at \( \mu = 0.01 \) by default.

The final step size selected for each test was based on the criteria that a noise rejection effect was observed without overcompensating, reducing DPOAE response magnitude [12], therefore achieving the optimal SNR. The overcompensation was evaluated
based on the baseline DPOAE level measured in quiet conditions. The adaptation process was started 3.6 seconds before the first recorded OAE level. This convergence time was sufficient even for smaller step sizes.

The number of coefficients \( N \) for both adaptive filter stages were determined based on experiments on one representative test subject’s noisy DPOAE measurements. Using \( N = 40 \) coefficients for Stage 1 and Stage 2 seemed appropriate for most subjects.

2.3. Otoacoustic response level extraction algorithm

Amplitude demodulation [26, 27] was used to extract the DPOAE levels from the recorded IEM signal of the tested ear. The SNR of the extracted DPOAE is evaluated from the ratio between the DPOAE level and the noise magnitude \( \hat{b} \). This noise magnitude is based on the average of the energy calculated in Eq. 7 plus two standard deviations as calculated in Eq. 8 over a total of 20 DFT bins, calculation of \( \hat{b} \) is presented in summary in Eq. 9. The noise estimator, also referred here as Noise2sd, is similar to what is often used in clinical systems as discussed in [14]. Here \( \hat{b} \) is estimated from the energy present in the 10 adjacent DFT bins above and below the expected DPOAE response frequency bin after discarding \( n_1 = n_2 = 25 \) bins in order to avoid the effect of spectral leakage of the DPOAE signal on these surrounding bins [12]. In Eq. 7, \( n_1 \) and \( n_2 \) are respectively the last discarded DFT bin below and above the DPOAE frequency bin. One frequency bin for the DFT represents a frequency step of \( \Delta f = f_s/M \), where \( M \) is the number of samples included in the DFT and \( f_s \) is the sampling frequency. The \( \hat{b} \) magnitude is converted to decibels afterwards.

\[
\bar{a} = \frac{1}{10} \sum_{i=n_1-11}^{n_1-1} (E_2(i))^2 + \frac{1}{10} \sum_{i=n_2+1}^{n_2+11} (E_2(i))^2 
\]  

\( \bar{a} \)
\[ \sigma = \sqrt{\frac{1}{n-1} \left( \sum_{i=n_1-1}^{n_1-1} ((E_2(i))^2 - \bar{a})^2 + \sum_{i=n_2+1}^{n_2+1} ((E_2(i))^2 - \bar{a})^2 \right)} \]

with \((E_2(i))^2\), energy of the DFT of the \(e_2(n)\) signal;

\[
\frac{M}{2} - 1 < i < M;
\]

\[ \hat{b} = \bar{a} + 2\sigma \]  

3. Experimental procedures

3.1. Experimental setup

For the DPOAE measurements, 8 otologically normal test-subjects (2 females and 6 males) with normal hearing were recruited. Their hearing threshold levels were measured by tonal audiometry at the octave band frequencies from 250 Hz to 8000 Hz performed with a clinical Interacoustics AC40 audiometer (made in Denmark) with TDH-39 transducers (Telephonics, made in USA) and tympanometry using a Madsen Zodiac 901 tympanometer (made in Denmark). Following the experimental test protocol approved by the “Comité d’éthique de la recherche de l’ÉTS”, the internal review board of the university [28]. Subjects had a stable DPOAE response (SNR of 3 dB or higher) over the whole frequency range of interest when measured with a clinical reference system [14].

For each subject, only the left ear was tested for DPOAEs, hence IEM-I and OEM-I were the signals from the left earpiece whereas IEM-C and OEM-C refers to the right earpiece. It is assumed that the physiological noise captured inside the ear canal is similar for both left and right ear canals. As IEM-C does not capture the DPOAE responses evoked at the tested ear, it can serve as a noise reference for these physiological noises [12]. While presenting the external noise, both the test ear and the contralateral ear were occluded by the OAE probes’ roll down foam tip so that measurements with external background noise would not induce temporary threshold shifts.
The earpieces were calibrated in a Head and Torso Simulator (HATS) for microphone level accuracy and adequate stimuli level at the eardrum. To set the electroacoustic conversion gain of the whole signal chain in MATLAB, including all components and data acquisition cards, the HATS was calibrated with a 1000 Hz pure tone of 94 dB(SPL). To match the sound pressure level at the eardrum with the IEM, a white noise signal was generated by one of the OAE probe miniature loudspeakers to measure the transfer function between this microphone and the HATS microphone to establish the gain adjustment necessary.

Measurements were carried out in a double-wall audiometric booth compliant with ANSI S3.1 [29] and ISO 8253-1 [30] where background noise was presented through four loudspeakers. At first, measurements were conducted without noise to establish a quiet baseline condition. Secondly, to cover the whole frequency spectrum of interest to evaluate the noise rejection performances of the different microphone combinations, a white noise sequence was played at 70 dB(A). Finally, to verify that the optimal microphone combination had a similar performance in more realistic measurement settings, an industrial noise sequence of aperiodic machinery noises was played at 70 dB(A). The sequence was selected from the NOISEX database [31] and had a crest factor of 18.8 dB, dynamic range of 78.3 dB, kurtosis of 3.9 and C-A weighted signal power ratio of 2.7 dB, calculated with a MATLAB Sound Analysis script [32]. Both white and industrial noise sequences were played through the desktop PC soundcard, and the sound pressure levels were set at the subject’s position (without any subject present) using a B&K 4189 free-field microphone (Bruel & Kjær, Nærum, Denmark) and Trident mX v6.8.0 software (Nelson Acoustics, Elgin, TX, USA) for processing.

The noise rejection algorithm was therefore evaluated in three different noise conditions; (a) the test conditions without external noise played through the loudspeakers, referred to as Quiet, (b) the condition where white noise is played (i.e. White), and (c) the condition where industrial noise is played (Industrial). Throughout the measurements, both probes were kept in the subject’s ears and not removed between the different test conditions.
3.2. Algorithm and microphone configurations

In order to determine the optimal microphone combination for robust noise reduction of external and internal noise, signals from OEM-I, OEM-C and IEM-C were post-processed in various combinations with either a single-stage (Stage 1 only) or a dual-stage (Stage 1 and Stage 2) adaptive filter topology.

Using four microphones simultaneously would require an additional adaptive filtering stage, therefore more computational resources in the future DSP application. It is also illogical to use the OEM-C to remove the ambient noise in IEM-C prior to the use of IEM-C to reduce physiological noise in IEM-I, since the transfer function between IEM-I and IEM-C (Stage 1) is already rejecting this incoherent ambient noise signal. For these reasons, the four simultaneous microphones configuration was not investigated.

At first, tests were conducted with configurations #1 to #3 (see Table 1) which only use one of the three available microphones for a single-stage noise reduction scheme. Configurations #4 to #7 use a combination of two out of three available microphones for a dual-stage noise reduction scheme. For configurations #1 to #5, the algorithms were tested without the fixed transfer function. The OAE probe estimated transfer function ($\hat{H}(z)$) can only be applied on an OEM signal (as in configurations #4 and #5) since the transfer function models the difference between the OEM-I and IEM-I and is meant to mimic the residual external noise inside the ear canal using an OEM. Configurations #6 and #7 consisted of the dual-stage topology #5 (IEM-C and OEM-I) tested with the fixed transfer function in parallel (see Fig. 2) and in series (see Fig. 4) respectively, in order to see the effect of this additional filter on the measured DPOAE response magnitude. Although the fixed transfer function might provide benefits, removing this transfer function from the processing scheme in configurations #6 and #7 would simplify the DPOAE system since no offline identification of the $H(z)$ transfer function would be necessary, thus minimizing the complexity of in-field measurements and processing time.

After determining the configuration that would provide the most promising results in the White conditions, this optimal configuration was applied in the Quiet test conditions and in Industrial conditions since it is expected that the adaptive filters’ per-
Table 1: Tested algorithm and microphone configurations. Here “C” denotes the contralateral earpiece microphone and “I” denotes the ipsilateral microphone. The fixed transfer function is placed in Parallel and in Series with the Stage 2 adaptive filter.

<table>
<thead>
<tr>
<th>Configuration #</th>
<th>Stage 1</th>
<th>Stage 2</th>
<th>Fixed transfer function</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>1</td>
<td>OEM-C</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>2</td>
<td>OEM-I</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>3</td>
<td>OEM-I</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>4</td>
<td>OEM-C</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>5</td>
<td>OEM-C</td>
<td>OEM-I</td>
<td>N/A</td>
</tr>
<tr>
<td>6</td>
<td>OEM-C</td>
<td>OEM-I</td>
<td>Parallel</td>
</tr>
<tr>
<td>7</td>
<td>OEM-C</td>
<td>OEM-I</td>
<td>Series</td>
</tr>
</tbody>
</table>

formance would be similar to the corresponding microphone combinations in White or Industrial. Since the probes were not refitted between the noise conditions, the transfer functions identified with the adaptive filters are expected to stay similar. It is also expected that ambient noise characteristics would only affect initial adaptive filter training and tracking capabilities which can be demonstrated with one, optimal, combination since such tracking capabilities are mostly determined by the filter order and step size \[25]. Optimizing the filter order and step size for every possible situation would be a time-consuming task \[25] and would not demonstrate the effect of the microphone combinations, therefore it would not achieve the main goals of this study. The Quiet condition is used to verify that the noise rejection algorithm does not alter the DPOAE signal when ambient noise is absent.

4. Experimental Results

4.1. Performance of algorithm and microphone configurations

To assess the performance of different algorithms and microphone configurations, the recorded DPOAE responses and noise levels were compared to those recorded in the Quiet conditions with configuration #0, i.e. the baseline results (Fig. 5(a)). The differences between this baseline and the recordings in noisy test conditions were plotted in Fig. 6 for each configuration presented in Table 1. The optimal configuration was established based on the minimal difference with the baseline, therefore when the background noise had almost no effect on the signal. Ideally, minimal or no influence of the background noise and minimized errors in signal processing on DPOAE levels are
Figure 5: a) DPOAE responses in Quiet test conditions with adaptive filtering noise rejection OFF (baseline) and b) DPOAE level differences from the baseline when adaptive filtering is ON. Median values across the 8 subjects tested are shown for DPOAE signals (blue squares) and noise (Noise2sd, red circles) with error bars indicating the respective first and third quartile.

desired. An upper limit of 3 dB was considered reasonable, since changes of 2-7 dB in DPOAE levels can be observed after noise exposure [7]. Moreover, based on previous work by Keppler [33] the minimal detectable difference for actual clinical systems was found to be around 3 dB. For quiet conditions DPOAE measurements in the current study, the DPOAE levels were higher than the noise levels as shown in Fig. 5(a) and met this minimal detectable difference criterion.

When the external noise level was high and adaptive filtering noise rejection was not used (configuration #0 in Table 1), the noise levels were clearly elevated compared to the baseline, most differences in noise levels lie between 15 dB and 25 dB as shown in Fig. 6(a). Also in this figure, DPOAE response levels were clearly shifted by 5 dB when compared to the baseline. Such results were expected since the noise rejection algorithm was not applied on the signals while white noise was played in the external loudspeakers. The elevated noise levels make it impossible to separate DPOAE signals from the background noise, leading to an artificial increase in DPOAE amplitude. Therefore, the DPOAE levels estimated in White conditions would normally be rejected because, in practice, artificial increase in DPOAE signals due to the measurement con-
Figure 6: Difference from baseline (Quiet, configuration #0) for each DPOAE response extracted with each configuration in Table 1 for White conditions
Figure 7: Difference from the baseline (Quiet, configuration #0) for a) configurations #0 and b) #5 (Table 1) in Industrial conditions. In c), results for White (Fig. 6(f)) are superimposed on results for Industrial conditions for comparison.

Conditions could hide true variations in DPOAE responses actually resulting from noise exposure. Similar findings were observed for measurements in Industrial conditions (Fig. 7(a)).

Fig. 6(b) to 6(h) show that the adaptive filtering noise rejection algorithm was clearly beneficial when working in elevated levels of background noise. The registered noise levels for configurations #1-7 were still significantly higher compared to the baseline (Paired Wilcoxon test with Bonferroni correction, all obtained p-values are lower than the significance level \( \alpha \) set at 0.05), but they fall within an overall more acceptable range (mostly around 0 to 5 dB absolute levels) and the DPOAE magnitude remains stable (around 0 to 3 dB difference when referring to the baseline). The Wilcoxon test is a non-parametric test and is used here because the assumption that the observations come from a normal distribution does not hold, and the number of observations is too small to apply the central limit theorem, hence, underlying assumptions for parametric testing are not fulfilled. The Bonferroni correction was applied by multiplying the p-values by the number of comparisons. This is done to account for the fact that making multiple comparisons increases the likelihood of incorrectly rejecting the null hypothesis.

Only when configuration IEM-C was used in a single-stage approach (Table 1 #1), the DPOAE responses differ significantly from measurements in Quiet conditions (Paired
Wilcoxon test with Bonferroni correction, \( p < 0.0001 \). Therefore, since configuration #1 was the only one using IEM-C in a single-stage, these results indicate that using the external microphone as noise reference signals in an adaptive noise rejection algorithm has more effect on the overall DPOAE level stability than using the internal microphone signals.

In terms of noise rejection in *White* conditions, configurations with OEM-I in a single-stage (configuration #3, Fig. 6(d)), or dual-stage algorithms (configuration #5, see Fig. 6(f)) significantly outperforms the other configurations for noise level reduction (Paired Wilcoxon test with Bonferroni correction between configuration #5 in *White* and #0 in *Quiet*, \( p < 0.05 \)). In addition, while the dual-stage configuration #5 did not decrease the median noise level when compared to the single-stage configuration #3, its performance was more consistent across subjects which can be seen from the smaller interquartile range in higher frequencies, in Fig. 6(f) compared to Fig. 6(d). Indeed, this additional IEM picks up mostly the physiological noise from the non-stimulated ear and these noise levels vary between subjects [34].

When comparing configurations #5, 6 and 7, the best noise rejection performance was obtained without the fixed transfer function (#5 in Table 1) as presented in Fig. 6(f) or when it was put in parallel (#6); although Fig. 6(g) shows that for the frequencies above 4000 Hz, the parallel fixed transfer function had weaker performance than with no fixed transfer function (#5). Absolute differences in DPOAE levels with the baseline averaged over the tested frequencies are shown in Table 2 along with the SNR improvement provided by the various configurations. As observed in this Table, putting the transfer function in parallel for high ambient noise levels increased the SNR as also shown in Fig. 6(g) (on average 0.5 dB when compared with configuration #5, Fig. 6(f)), but variations remained systematic between configurations #6 and #5 in *White* (\( p < 0.05 \)) and the absolute difference with the DPOAE signals measured in the baseline was increased. When the transfer function was put in series (#7), registered noise level’s medians were systematically increased (Paired Wilcoxon test with Bonferroni correction between configurations #7 and #5 in *White*, \( p < 0.01 \)), with on average a 1 dB difference over the whole frequency range as seen in Fig. 6(h) compared to configuration #5 in Fig. 6(f). As shown in Fig. 3 the fixed transfer function
can differ between subjects and generally had a resonance frequency around 3-4 kHz. When no fixed transfer function was implemented (#5) in White, no systematic difference with the DPOAE amplitude in Quiet with configuration #5 was found ($p > 0.05$), although at distinct higher frequency points the difference was noticeable when comparing DPOAE differences in Fig. 6(f) and 5(b). Such difference, which was more noticeable in presence of higher noise levels, could be explained by a poorer performance of the adaptive filtering algorithm in higher frequencies [24] possibly due to the time variability of the acoustic transfer function towards higher frequencies.

Table 2: Absolute differences in DPOAE levels with the baseline averaged over the tested frequencies for the various configurations and the SNR improvement provided by the microphone configurations compared to the results without adaptive filtering in White conditions

<table>
<thead>
<tr>
<th>Configuration #</th>
<th>Difference with Baseline (dB)</th>
<th>SNR improvement (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.4</td>
<td>5.6</td>
</tr>
<tr>
<td>2</td>
<td>2.7</td>
<td>6.6</td>
</tr>
<tr>
<td>3</td>
<td>2.9</td>
<td>7.2</td>
</tr>
<tr>
<td>4</td>
<td>2.3</td>
<td>6.5</td>
</tr>
<tr>
<td>5</td>
<td>2.2</td>
<td>7.2</td>
</tr>
<tr>
<td>6</td>
<td>2.8</td>
<td>7.6</td>
</tr>
<tr>
<td>7</td>
<td>2.4</td>
<td>7.1</td>
</tr>
</tbody>
</table>

Configuration #5 was established as the optimal combination due to the demonstrated beneficial effect of IEM-C for reducing physiological noise variability across subjects and OEM-I for the stability of DPOAE signals, as well as the lack of substantial benefits using the fixed transfer function $\hat{H}(z)$. To validate this optimal microphone configuration in realistic test conditions, configuration #5 was tested in Industrial conditions. Comparing results in Fig. 7(c) showed that on average the DPOAE response difference between White and Industrial was smaller than 1 dB, the difference with the baseline for White was 2.2 dB as shown in Table 2 and 2.8 dB for Industrial. Although noise levels measured in Industrial conditions with configuration #5 remain significantly higher ($p < 0.01$) than in White conditions with configuration #5, comparing Fig. 7(b) with Fig. 6(f) shows that the adaptive filtering noise rejection algorithm was capable of reducing the noise to acceptable levels (between -5 and +5 dB absolute noise levels) for more accurate DPOAE measurements. Activating the adaptive filter-
ing algorithm using configuration #5 improved the SNR average by 7.2 dB in the White (Table 2) and 6.8 dB in the Industrial conditions compared to switching the adaptive filtering off.

4.2. Effect of adaptive filtering in Quiet test conditions

As illustrated in Fig. 5(a) and 5(b) an average difference of 0.5 dB was found between DPOAE response signals registered with and without the adaptive filtering algorithm (comparing configuration #5 with #0 in Table 1) in the Quiet test conditions. A pairwise Wilcoxon test between configurations #5 and #0 in the Quiet reveals that the difference was non-significant \( p = 0.09 \). Fig. 5(b) shows that the major differences were found at the lowest frequency where the adaptive filtering algorithm slightly reduces the amplitude, and around 4000 Hz where the opposite was found, again due to the poorer performance of adaptive filtering in higher frequencies mentioned in Section 4.1. Differences in DPOAE amplitude were on average clearly below clinical test-retest variability \( [33] \), which indicates that the noise rejection algorithm performance could stay active even in the Quiet test conditions. Slight differences found in the Quiet test conditions with and without adaptive filtering could be explained by variations in the noise captured inside the ear canal such as breathing noise, which is more visible in the low frequencies due to the occlusion effect, that influences the DPOAE signal detection.

5. Discussion

To improve the noise rejection capabilities of the OAE measurement equipment in noisy environments, results of the various configurations of the adaptive filtering topology were compared and analyzed. The noise rejection configurations used in this study include: (1) different microphone configurations combining various noise reference signals, reducing the amount of noise signal disturbing the measurement system; (2) a fixed transfer function calculated offline which can be included in the adaptive filtering algorithm to potentially provide a more precise estimate of the OAE probe transfer function.
For the different microphone configurations, the optimal configuration to be used for noise rejection in White conditions appears to be a combination of the OEM-I together with IEM-C (#5). As expected, using the ipsilateral external microphone at the test ear gave the most coherent estimate of the external disturbance that will be picked up by IEM-I as opposed to the OEM-C which was influenced by the head’s acoustic shadow. In addition, the contralateral in-ear microphone was shown to reduce the inter-subject variability, suggesting it helped reduce physiological noises. In the envisioned system, the user will be able to select either the left or right probe to subsequently measure DPOAEs in each ear, the contralateral probe will always be used for noise rejection.

In this study, the implementation of the different fixed transfer function topologies in White conditions did not provide the expected benefits (see Fig. 6(g) & 6(h) and Table 2). This might be explained by a drift of the OAE probe earplug transfer function from the transfer function estimated at the beginning of the experiments, arising from a small change in the initial probe placement over the duration of the test sequence due to various movements of the human subject. Such drift in the fixed transfer function directly affects the series fixed transfer function topology (configuration #7, see Fig. 6(h)) since the error arising from the drift, directly propagates to the adaptive filters’ noise reference signal \( x_2(n) \) in Eq. 5. Moreover, an important difference in the resonance magnitude of the fixed transfer function between subjects, as shown in Fig. 3, can lead to various performances of the noise rejection algorithm depending on how much drift in the earplug transfer function was induced by changes in probe placement over time. Although in the parallel topology (configuration #6) the adaptive filter could bypass and correct the erroneous signal resulting from the bad fixed transfer function \( y'_2(n) \) in Eq. 6, since the reference input of this adaptive filter \( x_2(n) \) is the OEM signal and not \( \hat{H}(z) \)’s output, results did not show substantial benefit (see Fig. 6(g)). In the future, the transfer function identified offline, \( \hat{H}(z) \), may be used to initialize the coefficients of the second stage adaptive filter instead of using the transfer function in parallel with the adaptive filter. In this way, the adaptive filter will start immediately after the system is turned on, with a precise estimate of the individual’s OAE probe transfer function and should quickly account for slight variations in the probe fit and prevent the fixed
transfer function from inducing additional error in the signal, as seen with the tested parallel and series topologies.

The optimal configuration in the White conditions was also tested in the Quiet conditions in order to make sure that the adaptive filtering would not alter the DPOAE signal when noise levels were low (< 60 dB). In the application of DPOAE monitoring in-field, such artificial change due to processing could mask true early DPOAE changes (2-7 dB [7]) and therefore hide subtle, but important information about the functionality of the inner ear. In this study, it was shown that the optimal adaptive filtering noise rejection configuration (#5) can be used in very low background noise (Fig. 5(b)) with minimal influence on the initial DPOAE signal magnitude without adaptive filtering (#0). With reference to a baseline measured in optimal Quiet conditions at the beginning of a working day, true DPOAE variations due to noise exposure that exceeds the system’s normal test variability, would indicate the user to take appropriate action to prevent development of NIHL. Therefore, although the adaptive filtering noise rejection algorithm was not required in Quiet test conditions, it is advisable to maintain the algorithm active to minimize the system’s normal test variability.

Compared to previous work [12] and despite the limited passive attenuation actually achieved by the foam plug of the OAE probe [14] which affected the adaptive filters’ step size selection and overall noise rejection performance due to the magnitude of the residual noise in the IEM. The proposed system was able to achieve significant disturbance rejection in the presence of wide-band background noise presented from a realistic level up to 70 dB(A). This promising result was achieved in great part from the position of the ipsilateral outer microphone (OEM-I) at the immediate vicinity of the earplug. Such positioning ensured a greater spatial coherence than would typically be possible with the setup presented in [12] where the outer microphone was very distant from the actual OAE measurement probe. While the system was tested for levels between 70 and 75 dB(A), DPOAE measurements in higher noise levels could also be possible, but such levels were simply not tested for the moment. Therefore, further studies with the current or an improved version of the proposed DPOAE system, including DPOAE probes with better passive attenuation, might demonstrate that DPOAE measurements are also possible in noise sound pressure levels higher than
75 dB(A). In any case, even though current North American legislation for noise exposure in the workplace recommends wearing hearing protectors when noise levels exceed 85 dB(A) or 90 dB(A), it is foreseen that the envisioned system will be worn by the worker all day long to reduce the cumulative noise exposure with the probes’ passive attenuation. Moreover, the system could still perform DPOAE measurements after an excessive noise exposure dose is detected and while ambient noise levels remain around 70-75 dB(A) or lower, when the worker is away from noisy machinery, for example.

6. Conclusions

To enhance DPOAE measurements in noisy conditions, comparisons of different combinations of three microphones and algorithm configurations of adaptive filters were conducted using white noise sequences. After finding the optimal configuration, it has been validated for more realistic noise scenarios by testing the adaptive filtering noise rejection algorithm with industrial noise sequences. Moreover, the optimal configuration was tested in Quiet conditions to make sure it would not alter DPOAE signals when no external noise was present.

Results of tests conducted in this study have shown that in order to achieve optimal noise rejection in noisy background conditions, the ipsilateral OEM and contralateral IEM should be used as noise references respectively for the external background noise and the residual physiological noise. The results obtained in this study have also shown that the fixed transfer function estimating the OAE probe’s attenuation transfer function, typically used in conventional noise rejection systems, did not provide the expected benefits in the current tested processing scheme and should be used for a different purpose, perhaps for the initialization phase of the second adaptive filter or simply excluded from the design. To narrow down the range of step sizes for the adaptive filters, the OAE probe fit will be checked more thoroughly in the future. The proposed approach was shown effective in ambient noise levels of 70 dB(A), where measurements without adaptive filtering noise rejection would normally fail. The optimal adaptive filtering noise rejection algorithm can be kept operational in Quiet as well
as in *Industrial* test conditions in order to allow a constant and reliable detection of any cochlear changes during occupational noise exposure.

**Acknowledgements.**
Annelies Bockstael is a postdoctoral fellow of the Research Foundation-Flanders (FWO); the support of this organization is gratefully acknowledged. The ETS-affiliated authors are thankful to the Natural Sciences and Engineering Research Council (NSERC) Individual Discovery Grant Program for part of the funding as well as for technical support from the *Sonomax-ETS Industrial Research Chair in In-Ear Technology* for prototyping the experimental OAE probes. Vincent Nadon is especially grateful to the Institut Robert-Sauvé en santé et sécurité au travail (IRSST) for its support for the project.

**References**


[28] École de technologie supérieure, “Comité d’éthique de la recherche (CÉR) website”, Online, URL: http://www.etsmtl.ca/Recherche/Chercheurs/Ethique-recherche-humains/Comite-d-ethique-de-la-recherche-(CER) (date last viewed 01/01/15).


Reviewer's comments:

Reviewer #1:

1 - In my previous review, I asked that throughout the paper all equations should only be referenced after they have been introduced in the text. This has not been completed, despite claims that it has. This is a minor style issue.

The authors apologize for this misunderstanding. They previously thought the reviewer was referring to the order of presentation in the text vs equation positioning.

Eq. 1-3 are now first put into context in Section 2.2.1. Following equations (3-6 and 7-9) are also put into context and then referenced in the text.

2 - Minor typesetting comment - when citing multiple references, please adopt the simpler formatting [12,16-20] - for example on p4 para.2 '...with the microphone in the tested ear[12,16,17,18,19,20]...'. This should be used throughout the paper.

This has been corrected.

3 - Equations and symbols used - see for example p12. I think it would benefit the clarity of the paper if the authors adopted a more formal mathematical typsetting. As it stands there appears to be a mix of this and pseudo-code. For example, the general convention for typesetting the FFT of a variable x say, would be the uppercase X(f). An example from the paper Noise2sd is an estimate of the noise floor - general convention again would be to indicate an estimated value by an overline, hat or tilda over a variable such as n(t). 'Noise2sd' reads like a matlab variable name, and presents the work more like pseudo-code.

According to the reviewer’s recommendation, Eq. 7-9 have been modified. The estimate of the noise floor is now indicated by b^ since the variable n is already used to refer to the frequency bins of the FFT.

4 - It is not clear if a statistical signal detection algorithm was used prior to reporting DPOAE levels. It makes little sense reporting DPOAE level, if noise dominates. A good example to use could be an F-test or magnitude squared coherence test.

The authors would like to thank the reviewer for this constructive comment. The authors have used a more step-by-step experimental approach instead of a purely statistical approach.

Based on previous work by [Keppler2010] the minimal detectable difference for a clinical system was found to be around 3dB. For quiet conditions DPOAE measurements in the current study, the DPOAE level average was higher than the average noise level as shown in Fig. 5a and met this minimal detectable difference criterion.
According to the system self-distortion test that the authors performed prior to the analysis of the results, by placing the probe in a head and torso simulator and turning the DPOAE primary tone signals on, no distortion product was detected at the DPOAE frequency. Therefore DPOAE levels that were detected in humans did not result from a self-distortion of the system that could have been caused by possible nonlinearities. Based on this system self-distortion test and the FFT levels at the specific DPOAE frequencies calculated from the primary pure-tones, the authors validated that the DPOAE level was higher than the noise floor in humans for normal test conditions. The DPOAE levels measured with the FFT were similar to those detected with the Amplitude Modulation algorithm and the influence of different noise floor levels on the estimated DPOAE levels were assessed in [Nadon2014].

In higher ambient noise level conditions, the DPOAE levels were compared with the DPOAE measurement in quiet conditions, i.e. the baseline. Absolute differences in DPOAE levels with the baseline averaged over the tested frequencies for the various configurations in Table 2 show that DPOAE levels vary by less than 3dB for most combinations in noise conditions and this could be due to several factors such as probe fit and other measurement conditions. The normal probability plots in Fig.11 in [Nadon2015] also confirm that the DPOAE amplitude deviations from the DPOAE baseline are centered around 0 dB difference when adaptive filtering is used. Moreover, according to Fig.9 in [Nadon2015] the DPOAE signals are more stable and above the noise floor even for 75dB(A) industrial noise conditions when using adaptive filtering. For each algorithm combination presented in the current paper, the DPOAE level deviations from the baseline are shown in Fig. 6.

These corrections were applied to the text in section 4.1 that presents Figure 5a. The text in section 4.1 now reads as:

“Ideally, minimal or no influence of the background noise and minimized errors in signal processing on DPOAE levels are desired. An upper limit of 3 dB was considered reasonable, since changes of 2-7 dB in DPOAE levels can be observed after noise exposure [Engdahl2004]. Moreover, based on previous work by [Keppler2010] the minimal detectable difference for actual clinical systems was found to be around 3 dB. For quiet conditions DPOAE measurements in the current study, the DPOAE levels were higher than the noise levels as shown in Fig.5a and met this minimal detectable difference criterion.”
Reviewer #2:

-I remain concerned about very wide range of the adaptive step sizes used. I think it would be fair to add a sentence about this in the conclusions, in terms of suggested future work or the like.

To answer the reviewer’s comment the following sentence was added in the conclusions:

“To narrow down the range of step sizes for the adaptive filters, the OAE probe fit will be checked more thoroughly in the future.”

Detailed comments:

Introduction:
-P2, last line: "... cochlea, and these structures ..."

This has been corrected.

-P3, l.7 from bottom: Re-phrase: "Standard OAE probe tips provided with commercial systems are pre-moulded, that is, not be optimized for each individual's ear canal. Therefore they usually provide a limited amount of passive attenuation, which is not suitable for noisy industrial test environments [14]."

This has been corrected.

-P4, l.14-15: "... giving the possibility of combining the 4 microphone signals available."

This has been corrected.

-P4, l.16: "... greater spatial coherence of the noise signal ..."? Coherence represents a relation between two signals: the noise signal and what?

To answer the reviewer’s comment the sentence was modified:

“Moreover, the external microphones are integrated within the OAE probes so that a greater spatial coherence between external (OEM) and in-ear (IEM) noise signals can be achieved.”

Description:
-P4, l.2 from bottom: "... internal noise, the signals ..."

This has been corrected.

-P5, l.1: Delete "microphone"

This has been corrected.

-P5, l.10: "... per stimulus, to ..."

This has been corrected.
Add spaces around =

This has been corrected.

"... filter out the desired signal ..."

This has been corrected.

To improve readability while still keeping the same y-axis scale for both Subjects presented in the same figure. The phase was wrapped with a modulo of 1080°, that corresponds to 6*π, instead of using -180 to 180°.

This has been corrected.

I suggest to avoid the use of FFT/fft here. The FFT is a method to compute the DFT, where the latter is the concept referred to. Also, in the equations a notation like F(e) should be used to denote the DFT of the signal e.

The term FFT was replaced with DFT in the text according to the reviewer's recommendation. The equation notation was modified according to Reviewer 1's recommendation # 3.

Experimental results:

To address the reviewer's comment, the sentence was rephrased and now reads as:

"Ideally, minimal or no influence of the background noise and minimized errors in signal processing on DPOAE levels are desired."

The following text was added at the end of paragraph 3 in Section 4.1:

"The Wilcoxon test is a non-parametric test and is used here because the assumption that the observations come from a normal distribution does not hold, and the number of observations is too small to apply the central limit theorem, hence, underlying assumptions for parametric testing are not fulfilled. The Bonferroni correction was applied by multiplying the p-values by the number of comparisons. This is done to account for the fact that making multiple comparisons increases the likelihood of incorrectly rejecting the null hypothesis."

Adaptive filtering does not perform poorer at higher frequencies by nature. The observed poorer performance is more likely due to more complicated and more time-variable acoustical transfer functions towards higher frequencies. Please rephrase accordingly.
According to the reviewer’s recommendation the sentence was rephrased and now reads as:

“Such difference, which was more noticeable in presence of higher noise levels, could be explained by a poorer performance of the adaptive filtering algorithm in higher frequencies [Kuo2006] possibly due to the time variability of the acoustic transfer function towards higher frequencies.”

P20, Table 2: One decimal will do, thanks.

This has been corrected.

P21, l.11: "... marginally significant ...". Earlier the authors specifically state that their selected confidence level is 0.05. Thus an observed p-value of 0.09 signals a non-significant difference.

This has been corrected.