

Rapid estimation of high-parameter auditory-filter shapes

Yi Shen,^{a)} Rajeswari Sivakumar, and Virginia M. Richards

Department of Cognitive Sciences, University of California, Irvine, 3151 Social Science Plaza, Irvine, California 92687-5100

(Received 15 April 2014; revised 10 July 2014; accepted 12 August 2014)

A Bayesian adaptive procedure, the quick-auditory-filter (qAF) procedure, was used to estimate auditory-filter shapes that were asymmetric about their peaks. In three experiments, listeners who were naive to psychoacoustic experiments detected a fixed-level, pure-tone target presented with a spectrally notched noise masker. The qAF procedure adaptively manipulated the masker spectrum level and the position of the masker notch, which was optimized for the efficient estimation of the five parameters of an auditory-filter model. Experiment I demonstrated that the qAF procedure provided a convergent estimate of the auditory-filter shape at 2 kHz within 150 to 200 trials (approximately 15 min to complete) and, for a majority of listeners, excellent test-retest reliability. In experiment II, asymmetric auditory filters were estimated for target frequencies of 1 and 4 kHz and target levels of 30 and 50 dB sound pressure level. The estimated filter shapes were generally consistent with published norms, especially at the low target level. It is known that the auditory-filter estimates are narrower for forward masking than simultaneous masking due to peripheral suppression, a result replicated in experiment III using fewer than 200 qAF trials.

© 2014 Acoustical Society of America. [<http://dx.doi.org/10.1121/1.4894785>]

PACS number(s): 43.66.Yw, 43.66.Dc, 43.66.Ba [FJG]

Pages: 1857–1868

I. INTRODUCTION

For human listeners, peripheral frequency tuning is often probed psychophysically using masking experiments. A common experimental paradigm incorporates a model of peripheral filtering, and it is assumed that listeners detect a pure-tone target based on the output of a single hypothetical “auditory filter.” Only the portion of masker energy that falls within the pass band of the auditory filter contributes to masking, and the detection threshold is determined by the target-to-masker intensity ratio at the output of the auditory filter. Patterson (1976) introduced a behavioral technique to estimate the shape of the auditory filter using masking experiments. In these experiments, pure-tone targets are detected in the presence of a masker that consists of two noise bands, one on either side of the target frequency. According to the auditory-filter model, moving the noise bands closer to the target reduces the detectability of the target. Therefore, the auditory-filter shape can be estimated by measuring detection thresholds for various positions of the noise bands and then fitting the model of the auditory filter to the resulting experimental data.

Besides the estimation of the auditory filter shape, another means of estimating peripheral tuning behaviorally is to determine the level of a narrowband masker required to make a fixed-level pure-tone target inaudible as a function of the masker’s center frequency. The resulting function is referred to as the psychophysical tuning curve (PTC, e.g., Vogten, 1978).

Although frequency selectivity is a fundamental aspect of auditory perception, neither the auditory filter nor the PTC are routinely measured. One factor that hinders the

wide application of these measures of spectral resolution is the time required to estimate these functions. The estimation of both the PTC and the auditory filter involves repeatedly measuring detection thresholds across various masking conditions, yielding testing times in excess of 1 h. For listeners who are naive to psychoacoustic experimentation, both pre-test training and a larger number of repetitions for threshold estimation might be needed.

Recently, several novel techniques have been developed to improve the efficiency in assessing frequency selectivity. These techniques can be divided into two general categories. In the first category, Søk et al. (2005) successfully modified the traditional procedure for collecting the PTC and reduced the testing time to less than 10 min. This approach adaptively maintains the level of a narrowband masker to just mask a pure-tone target using Békésy tracking as the center frequency of the masker slowly sweeps across frequencies. The resulting trace of the masker level at threshold as a function of the masker frequency is taken as estimates of the PTC. Given its efficiency, this Békésy-tracking-based estimation of the PTC has been applied in a number of recent clinical studies (Malicka et al., 2009; Charaziak et al., 2012).

For the second category, a Bayesian adaptive procedure has been proposed (Shen and Richards, 2013) to improve the efficiency of assessing frequency selectivity. This approach estimates auditory-filter shapes while maintaining the basic experimental procedures relative to traditional procedures. Compared to the traditional experimental procedures, the Bayesian adaptive procedure aims to estimate the model parameters directly, rather than estimating detection thresholds and then using thresholds to estimate the model parameters. The resulting adaptive procedure, the quick-auditory-filter (qAF) procedure was initially evaluated using a model of the auditory filter that was symmetric in frequency, and was found to reduce the testing time for the estimation of the

^{a)}Author to whom correspondence should be addressed. Electronic mail: shen.yi@uci.edu

auditory-filter shape from more than an hour to approximately 10 min (Shen and Richards, 2013).

Thus far, the qAF procedure has only been tested using symmetric auditory-filter shapes, even though the symmetric model of the auditory filter is valid primarily at low sound intensities. At moderate and high intensities, auditory filters with asymmetric shapes better capture behavioral masking results and better represent cochlear processing. For asymmetric auditory-filter shapes, auditory-filter models with five free parameters are frequently adopted (Rosen *et al.*, 1998; Oxenham and Shera, 2003). An auditory-filter model with increased complexity (i.e., with a relative large number of model parameters) could mean that longer testing time might be required for individual parameters to converge and there might be a heightened potential for computational instability. The first focus of this study is to modify the qAF adaptive procedure to enable the efficient estimation of auditory-filter models with five free parameters.

A second focus of the current study is to evaluate the usefulness of the qAF procedure for naive listeners. Shen and Richards (2013) showed that the qAF procedure is sensitive to response errors made during the first a few trials of testing. Naive listeners having no experience with the experimental task are more likely to make early errors. Consequently, when naive listeners are tested, the qAF procedure occasionally exhibits unstable behaviors and relatively slow convergence of the parameter estimates. The current study introduces a number of experimental techniques to reduce the adverse effect of early response errors on parameter estimates and improve the stability of the qAF procedure.

An important feature of cochlear processing is its nonlinearity, which provides the auditory system with high sensitivity and fine spectral resolution. Estimates of auditory filters reflect such nonlinearities. For example, for normal-hearing listeners, the auditory filter is narrower when estimated using simultaneous masking than using forward masking owing to the presence of suppression for simultaneous maskers. The third focus of this study is to determine whether the qAF procedure is sufficiently sensitive to detect the known difference in the auditory-filter shapes for simultaneous versus forward masking experiments.

In the following section, the computational and experimental details of the updated qAF procedure will be described. Then, three experiments that evaluate the qAF procedure will be presented.

II. THE qAF PROCEDURE

A. The computational architecture of the qAF procedure

In the current implementation of the qAF procedure, the auditory filter is composed of rounded-exponential (roex) functions (e.g., Patterson *et al.*, 1982; Rosen *et al.*, 1998). On the high frequency side, the filter shape took the form of a single roex function (Patterson *et al.*, 1982)

$$W_h(g) = (1 + p_u g) e^{-p_u g}, \quad (1)$$

where g is normalized frequency ($g = |f - f_0|/f_0$, with f being frequency and f_0 being the center frequency of the

filter) and p_u is a parameter indicating the slope of filter's upper skirt. On the low-frequency side, the filter shape took the form (Patterson *et al.*, 1982)

$$W_i(g) = (1 - w)(1 + p_l g) e^{-p_l g} + w(1 + t g) e^{-t g}. \quad (2)$$

Equation (2) includes a tip filter and a tail filter, each of which is roex in shape. The parameters p_l and t indicate the slopes of the tip and tail filters, respectively. The parameter w is the weight of the tail filter relative to the total output of the filter (on the low-frequency side).

According to the power spectrum model of masking (e.g., Patterson, 1976), the auditory-filter shape directly relates to the detection threshold for a pure-tone target and a given masker. The model assumes that listeners detect the target through an auditory filter centered at the target's frequency and only the masker energy that is within the filter's pass band contributes to masking. The relationship among the signal power at threshold, P_s , the masker spectrum power N_0 , and the auditory-filter shape can be expressed as (Patterson *et al.*, 1982)

$$P_s(\Delta f_l/f_0, \Delta f_h/f_0) = K \left[\int_{\Delta f_l/f_0}^{\Delta f_l + BW/f_0} N_0 f_0 W_i(g) dg + \int_{\Delta f_h/f_0}^{\Delta f_h + BW/f_0} N_0 f_0 W_h(g) dg \right], \quad (3)$$

where BW is the bandwidth of each noise band; K is the ratio between the signal and masker power at threshold, indicating the efficiency of detection; and Δf_l and Δf_h are the distances from the target frequency to the nearest edges of the lower and higher masker bands, respectively. The two integrals in Eq. (3) calculate the total masker power that is within the filter's pass band.

By traditional methods, the auditory filter is estimated by fixing the masker spectrum level and repeatedly estimating the threshold target level for various combinations of $\Delta f_l/f_0$ and $\Delta f_h/f_0$. The estimated thresholds would then be used to fit the model in Eq. (3) and provide estimates of all five parameters (p_u , p_l , t , w , and K), four of which (p_u , p_l , t , and w) define the shape of the auditory filter. Such procedures can be very time-consuming because behavioral thresholds are estimated for several $\Delta f_l/f_0$ and $\Delta f_h/f_0$ pairs, leading to testing time of two hours or more.

The qAF procedure does not require the collection of thresholds, rather, it aims to estimate the model parameters directly in a single experimental track. Assuming a fixed target level, the qAF procedure adaptively chooses the appropriate masker spectrum level (x) and notch settings ($\Delta f_l/f_0$ and $\Delta f_h/f_0$) for stimulus presentation. For each trial, the probability of a correct response (PC) is modeled as a logistic psychometric function

$$PC(x, \Delta f_l/f_0, \Delta f_h/f_0; p_h, p_l, t, w, K) = \gamma + (1 - \gamma) [1 + e^{\beta(x - 10 \log N_0)}]^{-1}, \quad (4)$$

where γ is the chance performance level (i.e., 1/3 for a three-alternative forced-choice task) and β is related to the slope of the psychometric function. At the outset of an experiment,

a prior distribution is set for the parameters. Following each of the listener's responses, the likelihood function is evaluated, which is used to update the posterior parameter distribution in the five-dimensional parameter space $\{p_w, p_t, t, w, K\}$. Once data collection is complete, the final parameter estimate is derived as the mean of the posterior distribution. The core computational algorithm of the qAF procedure yields the selection of the stimulus before each trial. This is done in two steps (see also [Kontsevich and Tyler, 1999](#); [Shen and Richards, 2013](#)). First, the expected entropy of the posterior parameter distribution is calculated for all possible stimulus choices. Second, the combination of x , $\Delta f_i/f_0$, and $\Delta f_h/f_0$ associated with the minimum expected entropy is selected for stimulus presentation for the following trial.

In a previous study ([Shen and Richards, 2013](#)), the qAF procedure was implemented such that the posterior parameter distribution was evaluated in a discrete parameter space. Various technical difficulties make this previous implementation of the qAF procedure unfit for estimating auditory filters with asymmetric shapes (i.e., more parameters than estimated by [Shen and Richards, 2013](#)). The novel approach adopted here was to approximate the posterior parameter distribution as a multivariate Gaussian distribution. This means that the posterior probability density in a high-dimensional parameter space could be represented analytically, which alleviates the computational effort significantly compared to the numerical methods previously used. An additional advantage of the multivariate-Gaussian assumption is that the parameter space is *continuous* rather than discrete even though the parameter space is high dimensional.

With this modification, the qAF procedure proceeded as follows. Before data collection began, priors were set for the parameters. This involved initializing a column vector $\boldsymbol{\phi}_0$ that stored the mean for each parameters and a covariance matrix \mathbf{P}_0 . This mean vector and covariance matrix defined a multivariate Gaussian distribution that represents the experimenter's *a priori* knowledge about the model parameters. During the experiment, an extended Kalman filter algorithm ([Fahrmeir, 1992](#)) was used to update the posterior parameter distribution after each trial. Let r represent the potential responses with $r=1$ indicating a correct response and $r=0$ indicating an incorrect response. For our current purposes, the mean ($\boldsymbol{\phi}_i$) and covariance matrix (\mathbf{P}_i) of the posterior distribution following the i th trial can be derived as

$$\boldsymbol{\phi}_i = \boldsymbol{\phi}_{i-1} + \mathbf{K}(r - \mu) \quad (5)$$

and

$$\mathbf{P}_i = \tilde{\mathbf{P}}_{i-1} - \mathbf{K} \cdot \mathbf{J} \cdot \tilde{\mathbf{P}}_{i-1}, \quad (6)$$

where

$$\tilde{\mathbf{P}}_{i-1} = \mathbf{P}_{i-1} + \mathbf{Q} \quad (7)$$

and

$$\mathbf{K} = \tilde{\mathbf{P}}_{i-1} \cdot \mathbf{J}' \cdot [\mathbf{J} \cdot \tilde{\mathbf{P}}_{i-1} \cdot \mathbf{J}' + \sigma]^2. \quad (8)$$

The diagonal matrix \mathbf{Q} describes the expected diffusion of the parameter distribution across trials. Each element in the

matrix \mathbf{Q} is a positive value that defines, for a single model parameter, the amount of increase in variance expected for each trial. The purpose/advantage of including this diffusion term will be discussed below. The parameters μ and σ are the expected response and response variance according to the particular combination of stimulus parameters (x , $\Delta f_i/f_0$, and $\Delta f_h/f_0$),

$$\mu = \text{PC}(x, \Delta f_i/f_0, \Delta f_h/f_0; \boldsymbol{\phi}_{i-1}), \quad (9)$$

$$\sigma = \mu(1 - \mu). \quad (10)$$

The row vector \mathbf{J} is the Jacobian of the psychometric function,

$$\mathbf{J} = \left\{ \frac{\partial \text{PC}}{\partial p_h} \frac{\partial \text{PC}}{\partial p_l} \frac{\partial \text{PC}}{\partial t} \frac{\partial \text{PC}}{\partial w} \frac{\partial \text{PC}}{\partial K} \right\}. \quad (11)$$

The extended Kalman filtering procedure in Eqs. (5) and (6) iteratively updated the posterior parameter distribution according to the stimulus (x , $\Delta f_i/f_0$, and $\Delta f_h/f_0$) and response (r) from every trial. Because the posterior distribution was approximated as a five-dimensional Gaussian multivariate distribution, the parameter estimates after the i th trial was simply the vector $\boldsymbol{\phi}_i$.

Before each trial, the optimal stimulus that maximized the information gain was determined using a the one-step-ahead search algorithm described by [Kontsevich and Tyler \(1999\)](#). This algorithm can be broken into the following steps. First, the posterior distribution following the next trial was predicted for all possible stimuli, separately for each potential response (correct or incorrect). Second, the entropy of the predicted posterior was calculated for each potential stimulus. Let $\mathbf{P}_{i,\text{predict}}(x, \Delta f_i/f_0, \Delta f_h/f_0, r)$ be the covariance matrix for the predicted posterior, the predicted entropy $H(x, \Delta f_i/f_0, \Delta f_h/f_0, r)$ was given by

$$H(x, \Delta f_i/f_0, \Delta f_h/f_0, r) = \frac{N}{2} [1 + \ln(2\pi)] + \frac{1}{2} \ln |\mathbf{P}_{i,\text{predict}}|, \quad (12)$$

where $N=5$ is the total number of model parameters. Finally, the expected value of the total expected entropy for the i th trial, across both correct and incorrect responses, was calculated as

$$\begin{aligned} H_{\text{tot}}(x, \Delta f_i/f_0, \Delta f_h/f_0) = & \text{PC}(x, \Delta f_i/f_0, \Delta f_h/f_0, \boldsymbol{\phi}_{i-1}) \\ & \times H(x, \Delta f_i/f_0, \Delta f_h/f_0, r = 1) \\ & + [1 - \text{PC}(x, \Delta f_i/f_0, \Delta f_h/f_0, \boldsymbol{\phi}_{i-1})] \\ & \times H(x, \Delta f_i/f_0, \Delta f_h/f_0, r = 0). \end{aligned} \quad (13)$$

The procedure described above was applied to all potential stimuli (e.g., potential combinations of x , $\Delta f_i/f_0$, and $\Delta f_h/f_0$), and the stimulus with the smallest value of H_{tot} was used for the next trial. Note that while the parameter space was continuous, the stimulus space was discrete.

In a previous study, [Shen and Richards \(2013\)](#) found that early response errors could introduce instabilities in the qAF procedure. To alleviate that possibility, two novel

components were implemented/integrated into the qAF procedure.

First, the parameter distribution was assumed to diffuse over time. That is, the diagonal elements of the matrix \mathbf{Q} in Eq. (7) were set so that the variance of the posterior distribution increased by 1% for each trial. This had the effect of weighting the more recently collected data as more important so that inconsistent responses during the first few trials would not have sustained influences on the qAF track. The diffusion process was equivalent to assuming that the parameters to be estimated were random variables that, over the course of the experiment, had a path that was a random walk. This is a plausible hypothesis for psychoacoustic data collection, where the assumption of a truly stationary process over time seems unlikely.

The second component for improving the stability of the qAF procedure was the introduction of an upper and a lower limit for each parameter, so that the interim estimates of all parameters were kept within their limits. If, following the i th trial, the mean of the posterior distribution ϕ_i was expected to be outside of the limits according to Eq. (5), then the update of the posterior distribution would be skipped ($\phi_i = \phi_{i-1}$ and $\mathbf{P}_i = \mathbf{P}_{i-1} + \mathbf{Q}$). This constraint improved the stability of the qAF procedure by preventing extreme values in the interim parameter estimates.

B. The experimental modifications for testing naive listeners

The ultimate goal of the qAF procedure developed here is to provide a useful tool for assessing frequency selectivity routinely. To achieve this goal, the estimates provided by the qAF procedure would have to be reliable for listeners who are naive to masking experiments. Implementing efficient computational algorithms does not guarantee efficiency when naive listeners are tested. One reason is that it often takes a naive listener a few trials at the beginning of an experiment to fully understand the experimental task. This learning process is not incorporated into the qAF procedure, and may undermine the stability of the procedure. Moreover, since the qAF procedure governs the stimulus selection, the stimuli presented on the early trials are often poorly suited for a naive listener to learn the task quickly. For example, large changes in masker spectrum level (x) and values of $\Delta f_n/f_0$, and $\Delta f_i/f_0$ from one trial to the next have the potential of discouraging efficient learning of the task.

Based on substantial pilot work, the experimental design was modified so that the qAF procedure was not immediately activated. The ultimate procedure was as follows. Data collection was broken into blocks of 50 trials. Before the initial experimental trial, the one-step-ahead search algorithm in the qAF procedure was run based on the prior distributions, which determined the masker (including its spectrum level and notch settings) for the first qAF trial. However, before this initial qAF trial was tested, a transformed up-down procedure (Levitt, 1971) applied to the target level was used to familiarize the listener with the experiment task. During this initial “warm-up” period, the masker characteristics were fixed and equal to what would be presented on the first qAF

trial. The target level began at a high sensation level and was decreased after two consecutive correct responses and increased after a single incorrect response using a 10 dB step size. The qAF trials began after the second reversal. Once the qAF procedure was activated, the target level remained fixed and the masker spectrum level and notch bandwidths were adaptively varied. After 50 trials the listener had the opportunity to take a short break before the next experimental block was tested. The new experimental block used the same configuration as the first block, beginning with an up-down track and a fixed masker. The fixed masker was identical to the one that would be tested on the first qAF trial of that block, which resumed from the previous block.

By starting each block with a warm-up period, the listener was repeatedly reminded about the experimental task and about what the target tone sounded like. Although this warm-up period made up approximately a quarter of trials, pilot testing suggested that it significantly improved the test-retest reliability of the parameter estimates from the qAF procedure.

III. EXPERIMENT I: TEST-RETEST RELIABILITY OF THE qAF PROCEDURE

A. Methods

1. Listeners

Ten listeners were recruited from the School of Social Sciences Human Research Lab Pool at the University of California, Irvine. Listeners’ ages ranged from 18 to 28, and none had previous experience in psychoacoustic testing. For each listener, the ear with lower pure-tone average (PTA) threshold across 1, 2, and 4 kHz was tested in the experiment. In the event the PTA thresholds were the same in both ears, the left ear was tested. Because the current experiment focused on the test-retest reliability of the qAF procedure rather than the dependency of the estimated auditory filter on hearing threshold, listeners were included in the experiment regardless of their hearing status. Nevertheless, all listeners had audiometric thresholds less than or equal to 30 dB hearing level (HL) at frequencies from 250 to 8000 Hz.

2. Stimuli

Listeners detected the presence of a pure-tone target in a three-interval, three-alternative, forced choice task (3I-3AFC). Each interval contained a noise masker, which consisted of two 500-Hz-wide noise bands. The two noise bands formed a spectral notch between them, and the 2-kHz target tone was presented within the notch during one of the three intervals, drawn at random. The masker and target were both 300 ms in duration, including 10-ms onset/offset raised-cosine ramps. During the qAF procedure, the target level was fixed at 30 dB sound pressure level (SPL) while the masker notch and spectrum level were adaptively varied. On the scale of normalized frequency [see g in Eq. (1)], the potential values for the upper notch width [from the lower edge of the upper masker band to the target frequency, $\Delta f_n/f_0$ in Eq. (4)] were 0, 0.25, and 0.5; the potential values for the lower notch width [from the upper edge of the lower masker band to the target frequency, $\Delta f_i/f_0$ in

Eq. (4)] were 0, 0.075, 0.15, 0.225, 0.3, 0.375, 0.45, 0.525, and 0.6. The potential values for the masker spectrum level [x in Eq. (4)] were linearly spaced between -10 to 50 dB SPL using 15 values, inclusively. The qAF procedure selected the combination of x , $\Delta f_{il}/f_0$, and $\Delta f_{ih}/f_0$ (one of the 405 potential stimuli) that maximized the expected information gain before each trial.

The stimuli were digitally generated using a sampling frequency of 44.1 kHz on a PC, which also controlled the experimental procedure and data collection through custom-written software in MATLAB (The MathWorks, Inc.). The stimuli were presented to the test ear via a 24-bit soundcard (Envy24 PCI audio controller, VIA technologies, Inc.), a programmable attenuator and headphone buffer (PA4 and HB6, Tucker-Davis Technologies, Inc.), and a Sennheiser HD410 SL headset. Each stimulus presentation was followed by visual feedback indicating the correct response.

3. Procedure

Each listener was tested in a single session of one hour. At the beginning of the session, the audiometric characteristics of the listener were determined. Then, two estimates of the listener's auditory filter at 2 kHz were obtained sequentially, each estimate being based on one qAF track. A qAF track was comprised of four blocks of 50 trials, with the warm-up periods described in Sec. II starting each block. The initial target level during the warm-up periods was 45 dB SPL.

The qAF procedure had the following configurations. First, the γ parameter of the psychometric function in Eq. (4) was set to $1/3$ according to the three-alternative, forced-choice task used in the current experiment. The slope of the psychometric function β was set to unity (Shen and Richards, 2013). The prior distribution for each of the model parameters was specified as follows. The parameter p_u and p_l were allowed to take values between 10 and 70 . The prior distributions for these parameters had a mean of 40 and a standard deviation of 40 . The parameter t ranged between 0 and 20 , with a prior mean of 5 and standard deviation of 5 . The value for $10 \log(w)$ was constrained between -60 and 0 dB and its prior had a mean of -30 dB and a standard deviation of 40 dB. Finally, the value for $10 \log(K)$ was between -10 and 20 dB and its prior distribution had a mean of 5 dB and a standard deviation of 20 dB. No prior covariance was assumed among the model parameters. The prior distributions were spread broadly, and together with the wide range of parameters available, the effect was to minimize the influence of the prior distributions on the parameter estimates.

B. Results

In a typical qAF track, the estimate for each of the parameters [p_u , p_l , t , $10 \log(w)$, and $10 \log(K)$] during the first few (<10) trials reflected the mean of the parameter's prior distribution. Beyond these early trials, the estimate either gradually or somewhat abruptly, shifted toward a final estimate. Visual inspections suggest that it typically took 100 qAF trials (approximately 130 trials including the warm-up period) for the parameter estimate to reach asymptote.

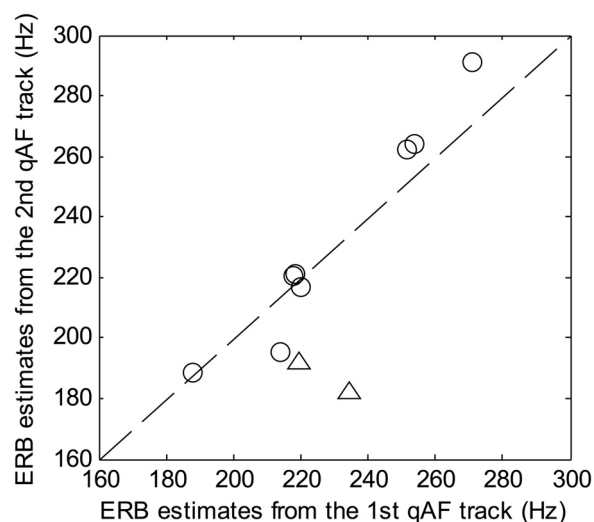


FIG. 1. The ERB estimates from two qAF tracks are shown, with the second estimate being plotted as a function of the first. Two “outlier” estimates are represented by triangles while the results for the remaining eight listeners are plotted as circles.

Figure 1 summarizes the two estimates of the auditory-filter bandwidths from all listeners. The auditory-filter bandwidth was quantified as the equivalent rectangular bandwidth (ERB) (e.g., Glasberg and Moore, 1990). For the current study, the ERB of the auditory filter was approximated from the parameter p_u and p_l by (Patterson *et al.*, 1982)

$$\text{ERB} \approx f_0 \left(\frac{2}{p_u} + \frac{2}{p_l} \right). \quad (14)$$

Note that the above equation approximated the auditory-filter bandwidth at low sound intensities. At high intensities, it was expected the tail portion of the filter's low-frequency skirt might contribute to the overall filter shape to a larger degree [a higher value for $10 \log(w)$], and Eq. (14) would under estimate the filter bandwidth.

There was a significant correlation between the ERB estimates from the first and second qAF tracks ($r=0.84$; $p < 0.01$). The mean and median distances between the two estimates (absolute value of the difference) were 15.0 and 10.3 Hz, respectively. Except for two listeners (triangles), the results showed excellent agreement between the first and second estimate of the auditory-filter ERB, within 20 Hz of each other (circles). For the remaining two listeners, the ERB estimates were narrower for the second qAF track.

The average ERB estimate, pooled across listeners and replicates, was 226.1 Hz, which was close to the ERB estimates for normal-hearing listeners from a number of previous studies. For example, Dubno and Dirks (1989) measured the auditory-filter shapes from normal-hearing as well as hearing-impaired listeners. The ERB estimates at 2 kHz obtained from the nine young, normal-hearing listeners ranged between 220 and 320 Hz with a mean of 258 Hz. Sommers and Humes (1993) estimated auditory filters for young and elderly listeners. Their young, normal-hearing group showed ERB estimates ranging between 282 and 330 Hz with a mean of 307 Hz. Glasberg and Moore (1990)

provided a standard formula to compute the expected ERB for a young, normal-hearing listener as a function of the auditory filter's center frequency, which predicts an ERB of 241 Hz at 2 kHz. In a more recent study, Baker and Rosen (2006) also studied the change of the auditory-filter shape as a function of center frequency. Their fitted model predicted an ERB of 260 Hz for young, normal-hearing listeners at low sound intensities. In summary, the ERB estimates from the current experiment were reasonably consistent with published results. The agreement between the estimates of the auditory-filter shapes using the qAF procedure and estimates based on traditional results were investigated further in experiment II.

Figure 2 plots the mean (solid curves) and median (dashed curves) distances between the two qAF tracks for all model parameters and the ERB estimate as functions of trial number. These plots excluded the warm-up trials at the beginning of each experimental block. For the ERB estimate (bottom right), the mean and median distances rose sharply during early trials and began to decrease after 20 qAF trials. Beyond about 60 trials, the mean distance plateaued at about 15 Hz while the median distance continued to decrease with more trials. The agreement between the two qAF tracks on the first few trials reflected the fact that the same prior distribution was used for both tracks. After several trials, the two tracks diverged, owing to differences in responses, to the diffusion term, and to differences in the stimuli tested. After 20

or so trials, the differences in parameter estimates reduced, and then reached an asymptote. This is indicative of the ultimate convergence of the procedure. Overall, to ensure that the expected test-retest differences ERB will be on average 20 Hz or less, it is recommended that each experimental participant run at least two blocks, approximately 75 to 80 qAF trials.

Inspections of the top two panels of Fig. 2 suggested that the ERB result better mirrored the parameter p_u than p_l . In contrast to ERB and p_u , the test-retest distance for p_l remained small, between values of 2 and 4 across trials even for the first few trials. The test-retest distance for the parameters t , $10 \log(w)$, and $10 \log(K)$ shared a similar dependency on trial number: following a rapid rise during the first 20 trials or so, the test-retest distance gradually decreased with increasing trial number.

For most parameters, the median test-retest distance (dashed curves) was lower than the mean (solid curves). This was most evident for the ERB and $10 \log(w)$. This difference between the mean and median values indicated that for a majority of listeners the test-retest reliability was excellent, although there were occasional outliers with relatively large test-retest distances. This was consistent with the apparent "outliers" in test-retest values obtained for the ERB estimate (triangles in Fig. 1).¹

In summary, experiment I demonstrated the satisfactory test-retest reliability of the qAF procedure for a five-parameter roex model of the auditory filter and for a 30-dB SPL, 2-kHz tonal target. For the key parameter of the auditory filter, the ERB, 200 trials of data collection led to a test-retest difference within 20 Hz for eight out of ten listeners. On average, it is expected that a within-20-Hz test-retest difference would be achieved using as few as two blocks of 50 trials for the experimental condition tested here.

IV. EXPERIMENT II: AUDITORY-FILTER SHAPE AS FUNCTIONS OF CENTER FREQUENCY AND INTENSITY

To evaluate the generalizability of the qAF procedure to other frequencies and target levels, auditory-filters were estimated at two frequencies and two target levels in experiment II. The resulting estimates, from naive listeners, are then compared to data available from the literature.

A. Methods

Sixteen listeners were recruited from the School of Social Sciences Human Research Lab Pool at the University of California, Irvine. The listeners were aged from 18 to 24 years and all were naive concerning psychoacoustic experiments and had hearing thresholds equal or less than 15 dB HL for all audiometric frequency between 250 and 8000 Hz in both ears. For each listener, the ear with better PTA threshold was tested, and, in the case of equal PTA threshold across ears, the left ear was tested.

Listeners detected a target tone in a 3I-3AFC task that was configured identically compared to experiment I, except

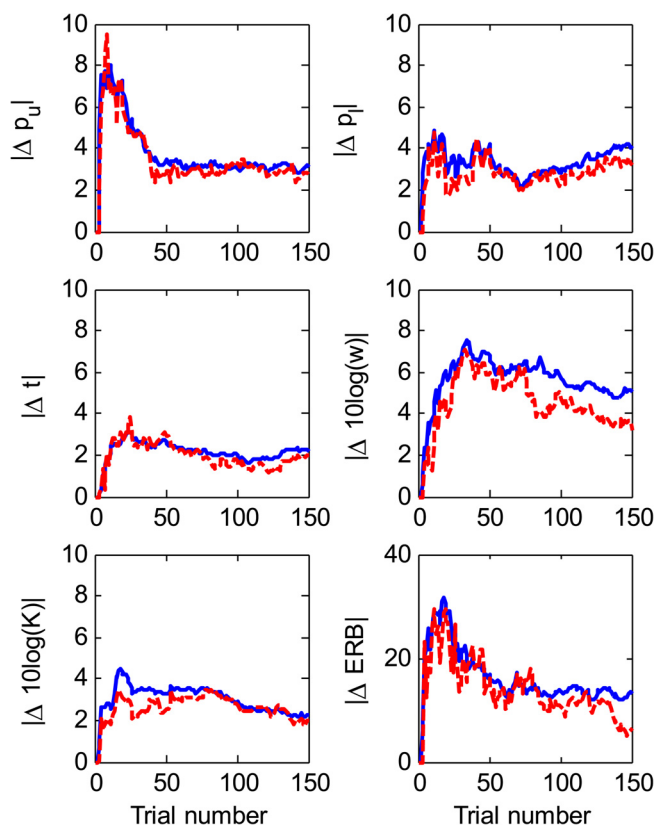


FIG. 2. (Color online) The differences in parameter estimates between two repetitions of the qAF procedure (absolute values shown) as functions of trial number. The solid and dashed curves plot the mean and median results across ten listeners. Warm-up trials, which initiated each block, are excluded.

that the target frequency was 1 kHz for eight listeners and was 4 kHz for the remaining eight listeners. For each listener, the target level (P_s) was either 30 or 50 dB SPL in separate conditions.

Experiment II was conducted in a single one-hour session. For each listener, the two target levels were tested in random order. One qAF track with four blocks of 50 trials was run for each target-level condition. The qAF procedure was configured as in experiment I, except that the initial target level for the warm-up period was set to 15 dB above the target level to be tested.

B. Results

Figure 3 plots the estimated filter shapes from experiments I and II with the three target frequencies arranged in rows and the two target levels arranged in columns. In each panel, gray curves plot the estimated filter shape for individual listeners, while the black solid curve plots the filter shapes derived from the averaged parameter estimates.² The filter fits reported by Baker and Rosen (2006) are plotted as dashed curves for comparison. The filter shapes based on the averaged parameter estimates were within 10 dB of the results from the study of Baker and Rosen, except for the 1-kHz, 50-dB target condition. In this condition (upper right panel), the low-frequency tail section of the estimated

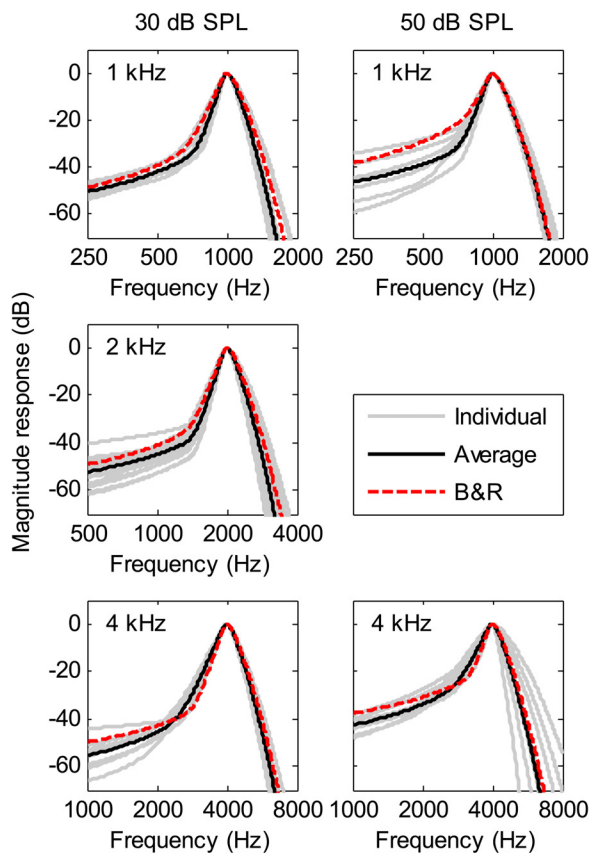


FIG. 3. (Color online) The estimated auditory-filter shapes from experiments I and II for two target levels (columns) and three target frequencies (rows) are plotted. In each panel, the gray solid curves are for individual listeners, while the black solid curves indicate the filter shape derived from the averaged parameter estimates across individual listeners. The dashed curves indicate the auditory-filter shapes reported by Baker and Rosen (2006).

auditory-filter shape exhibited large individual differences and an overall lower gain than the fit of Baker and Rosen.

The top panel of Fig. 4 plots the estimated ERB's as functions of the target frequency. Target levels of 30 and 50 dB SPL are indicated using squares and diamonds, respectively. Data collected from experiment I with a 2-kHz, 30-dB SPL target are also plotted in the figure. The standard formula for the ERB (Glasberg and Moore, 1990) and the fitted function describing the frequency dependency of the ERB from a more recent study (Baker and Rosen, 2006) are shown as solid and dashed lines, respectively.

The obtained ERB estimates replicated the previously published results well. At 1 and 2 kHz, the ERB fits from the two previous studies were within the range of ERB estimates from individual listeners. For the 50-dB SPL target at 4 kHz, the ERB estimates revealed large individual differences such that for some listeners the ERB estimate differed from the standard ERB formula by nearly as much as one octave.

The mean parameter estimates obtained from experiments I and II are summarized in Table I with their standard deviations in parentheses. The mean estimates for the four parameters describing the auditory-filter shape [p_u , p_l , t , and $10 \log(w)$] are shown in the two middle and two bottom

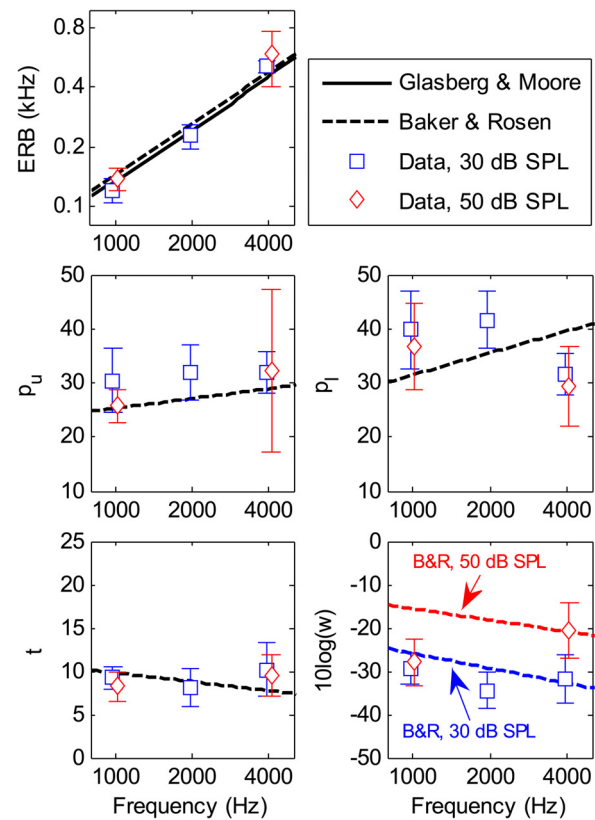


FIG. 4. (Color online) The average estimates of ERB (top panel) and auditory-filter parameters (middle and bottom panels) from experiments I and II as functions of target frequency. Target levels of 30 and 50 dB SPL are indicated using squares and diamonds, respectively. Error bars indicate \pm one standard deviation. For comparison, standard formulas describing the frequency dependency of the auditory-filter bandwidth in normal-hearing listeners from Glasberg and Moore (1990) (solid curve) and Baker and Rosen (2006) (dashed curve) are shown. For the parameter $10 \log(w)$, the fits of Baker and Rosen were level dependent and consequently are shown using two dashed lines (lower right panel).

TABLE I. The averaged auditory filter parameters and ERB estimates from experiments I and II, for two target levels and three target frequencies. The standard deviations are shown in parentheses.

Frequency (kHz)	p_u	p_l	t	$10 \log(w)$	$10 \log(K)$	ERB (Hz)
30 dB SPL						
1	30.4 (5.9)	39.8 (7.1)	9.3 (1.3)	-29.4 (3.5)	3.8 (3.3)	120.0 (17.1)
2	31.9 (5.1)	41.7 (5.2)	8.0 (2.2)	-34.5 (4.2)	3.0 (4.1)	226.1 (30.7)
4	31.8 (3.9)	31.6 (3.9)	10.1 (3.1)	-31.8 (5.5)	3.6 (2.9)	510.7 (32.9)
50 dB SPL						
1	25.7 (3.1)	36.8 (8.0)	8.3 (1.7)	-28.0 (5.3)	1.3 (3.2)	135.2 (18.0)
4	32.3 (15.0)	29.4 (7.5)	9.5 (2.4)	-20.7 (6.4)	1.4 (2.8)	580.9 (187.3)

panels of Fig. 4. For each parameter, the estimated value is plotted as a function of target frequency using squares and diamonds for target levels of 30 and 50 dB SPL, respectively. The auditory-filter fit by Baker and Rosen (2006) is shown in Fig. 4 as dashed lines (labeled as “B&R”) and will be referred to as the “reference fit” in the following discussion. Since their model included level dependency for the parameter $10 \log(w)$, the predicted values for $10 \log(w)$ for target levels of 30 and 50 dB SPL are plotted as two separate dashed lines.

For a target level of 30 dB SPL, close agreement between the current parameter estimates and the reference fit by Baker and Rosen was observed for the parameters p_u , t , and $10 \log(w)$ (squares in the left panels as well as in the bottom right panel). For the parameter p_l , the estimates seemed to be too high at 1 kHz and too low at 4 kHz compared to the reference fit. This difference may reflected the fact that Baker and Rosen described the auditory-filter shapes across frequency regions (from 0.25 to 4 kHz) using a single model, in which the parameters were linear functions of the logarithmically transformed target frequency. In contrast, the auditory-filter shapes for various target frequencies and levels were estimated separately from different individual listeners in the present study. Although the origins for the discrepancies between the current and predicted p_l estimates were not clear, the degree to which these discrepancies affected the ERB estimates (top panel) and the filter shape (Fig. 3) at low target levels was limited.

For a target level of 50 dB SPL, the mean estimates for the parameters p_u , p_l , and t (diamonds in the left panels) were similar to those obtained at 30 dB SPL. Moreover, the standard deviations for p_u at 4 kHz were much larger than those at 30 dB SPL, reflecting an increase in individual differences. This was also apparent for the filter shapes derived using a 50-dB SPL, 4-kHz target (Fig. 3, light curves in the lower right panel) where high-frequency skirts showed relatively large individual differences. Overall, satisfactory agreements between the mean parameter estimates to the reference fit were observed with some exceptions; the p_l estimate at 4 kHz and the $10 \log(w)$ estimate at 1 kHz were lower than the corresponding estimates in the reference fit.

V. EXPERIMENT III: PROBING COCHLEAR NONLINEARITY

Auditory-filter shapes were estimated for both simultaneous and forward masking conditions. Past work has demonstrated that ERBs estimated with simultaneous maskers

are approximately 1.3 times those observed for forward maskers (Moore *et al.*, 1987; Sommers and Gehr, 1998; Oxenham and Shera, 2003; Unoki *et al.*, 2007; Unoki and Tan, 2005). Low-level targets were tested to estimate filters when they were sharply tuned (Oxenham and Shera, 2003). The goal was to determine whether the qAF procedure could reliably detect the known differences in filter shapes between simultaneous and forward masking conditions.

A. Methods

Seven listeners naive with respect to psychoacoustic experimentation were recruited using the same protocol as experiment II. For each listener, the ear with lower hearing threshold at 2 kHz was tested, and, in case of identical thresholds across ears at 2 kHz, the left ear was tested.

Listeners detected a brief, 2-kHz tone pip in a 3I-3AFC task. The 15 ms-long target was 15 dB above its threshold (15 dB SL) and turned on and off using 7.5-ms raised cosine onset/offset ramps. This target was similar to the one tested by Oxenham and Shera (2003), but was somewhat more intense (15 versus 10 dB SL) and somewhat longer (15 versus 10 ms). This slight modification was implemented to ensure that naive listeners could learn the task quickly and reliably. For the simultaneous- and forward-masking conditions, the target was either at the temporal center of the masker or immediately following the masker, respectively. The potential masker configurations were as in experiment I except that its duration was 400 ms including 10-ms raised-cosine onset and offset ramps. The masker spectrum levels that could be tested were 15 values linearly spaced between -10 and 60 dB SPL.

Four listeners first ran the simultaneous-masking condition, and the other three began with the forward-masking condition. In other regards, the procedures were as described for experiment I. Before data collection began, each listener’s absolute threshold for a 15-ms, 2-kHz tone-pip target was first determined using a 3I-3AFC paradigm with a 2-down, 1-up adaptive procedure (Levitt, 1971).³ Two estimates of the absolute threshold were obtained sequentially, and the average of the two estimates was used to set the target level at 15 dB SL.

B. Results

Figure 5 plots ERB estimates from the forward masking condition as a function of the estimates from the

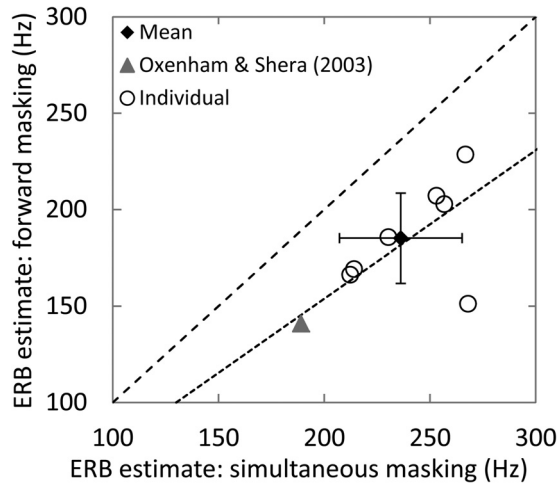


FIG. 5. ERB estimates for the forward masking condition are plotted as a function of the ERB estimates for the simultaneous masking conditions for each subject (unfilled circles). The averaged value is indicated with a filled diamond and error bars indicate \pm one standard deviation of the estimate. Results from Oxenham and Shera (2003) for a similar condition are shown using a gray triangle. The dashed line indicates equivalence and the dotted line indicates the expected relationship between the ERB estimates from the two masking conditions based on previous studies (see text for details).

simultaneous masking condition for the individual listeners (circles) and their averaged results (black diamond). The filled triangle indicates the average ERB estimates from a similar experiment reported by Oxenham and Shera (2003). The dotted line indicates a ratio of 1.3, the expected ratio of the ERB estimates for forward relative to simultaneous maskers. The dashed line indicates equivalence.

For all seven listeners tested, the ERB estimate from the simultaneous masking condition was larger than that from the forward masking condition. Moreover, the ratio between the ERB estimates from the two conditions was similar to the expected value of 1.3, except for one listener. The estimated ERBs were somewhat larger than those of Oxenham and Shera (2003), who used the traditional method and similar stimuli (filled triangle). It is possible that this difference in ERB magnitude reflected the relatively less energetic signal tested by Oxenham and Shera (10 dB SL versus 15 dB SL, 10 vs 15 ms duration), although differences in the experience/practice of the listeners might have also contributed to the difference.

The top panel of Fig. 6 shows the difference in ERB estimates, simultaneous masker minus forward masker, as a function of qAF trial number for the listeners (gray). The mean difference as a function of trial number is indicated using a black curve. As the number of trials increased, the difference in ERB estimates increased such that after 30 qAF trials, the difference in ERBs was consistently greater than zero and after approximately 75 qAF trials, the difference remained approximately constant.

Using the data shown in the top panel of Fig. 6, the bottom panel plots the expected probability of failing to detect a positive difference in ERBs (simultaneous minus forward masker) as a function of trial number using a logarithmic ordinate. This is essentially a graph of the expected probability of a type I error as a function of trial number. Beyond

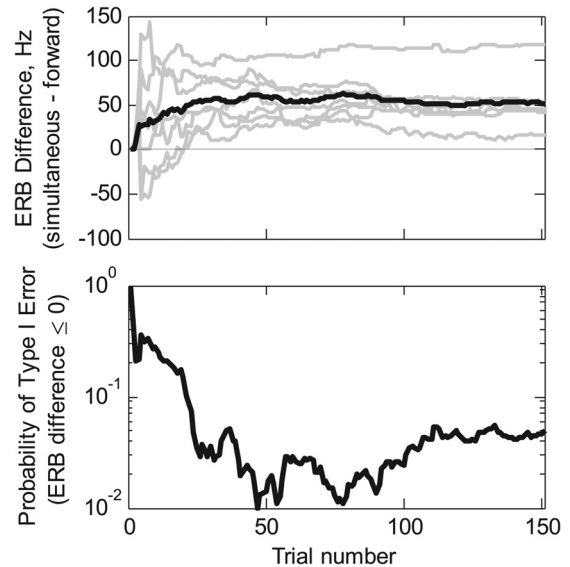


FIG. 6. Top: The difference in ERB estimates, the simultaneous-minus forward-masked conditions, as a function of trial number. Results from individual listeners are shown in gray, and the averaged difference is shown using a thick black curve. Bottom: The probability of a type I error (failure to detect a positive difference in ERB, simultaneous minus forward maskers) is plotted as a function of the number of trials.

110 or so trials, the type I error rate was relatively stable at approximately 0.045, indicating reasonable sensitivity for the detection of the expected changes in the ERBs for simultaneous versus forward masked experiments.

Table II shows the averaged parameter estimates, and the standard deviations of those estimates, for the simultaneous and forward maskers. First, consider the parameter estimates for the simultaneous masking condition from experiment III and the estimates from experiment I (see the 30-dB SPL, 2-kHz condition in Table I). The estimates of all the parameters, except for $10 \log(K)$, are similar to one another. The value of $10 \log(K)$ is larger for the current experiment, presumably because the target duration is short (increased thresholds at 30 ms relative to 200 ms, e.g., Wright and Dai, 1994). Overall, the estimated auditory-filter shapes were consistent across groups as well as across target durations (see also Wright and Dai, 1994).

Next, consider the difference in the parameter estimates for forward versus simultaneous maskers (Table II). The largest difference in parameter estimates, as a ratio, was for the parameter p_u . Moore *et al.* (1987) also noted this result and it has been a consistent result in additional reports (Sommers and Gehr, 1998; Oxenham and Shera, 2003; Unoki and Tan, 2005). In contrast, estimates of p_l were variable across studies, sometimes larger and sometimes smaller for the simultaneous than the forward masking conditions (Sommers and Gehr, 1998; Oxenham and Shera, 2003; Unoki and Tan, 2005), potentially due to the differences in the form of the auditory filters fitted. With regard to the pattern of changes in the magnitude of p_u and p_l , as found by Sommers and Gehr (1998), the correlation between the estimate of p_u with the forward masker and the difference in p_u estimates between the forward- and simultaneous-masking

TABLE II. The averaged auditory filter parameters and ERB estimates from experiment III, for simultaneous and forward maskers. The standard deviations of the mean are shown in parentheses.

	p_u	p_l	t	$10 \log(w)$	$10 \log(K)$	ERB (Hz)
Forward	41.2 (8.2)	47.4 (7.2)	9.5 (1.4)	-39.9 (7.3)	9.3 (4.8)	187.4 (27.1)
Simultaneous	29.0 (3.4)	40.0 (7.2)	8.1 (3.4)	-33.7 (6.8)	12.1 (3.5)	243.0 (23.9)

conditions was highly significant ($r=0.96$, $p<0.0005$). Unlike Sommers and Gehr (1998), the parallel correlation for p_l did not reach significance, but the correlation coefficient was large and approached significance ($r=0.73$, $p \sim 0.06$). For the forward masker the range of the values of p_u across listeners was larger than the range for simultaneous maskers (note the change in the standard deviations of the mean in Table II), whereas for p_l the ranges of values were approximately the same.

To summarize, these data indicate that after approximately 130 trials (counting about 20 “warm-up” trials), the known difference in the ERB estimates for simultaneous and forward masking conditions was obtained for a five-parameter auditory-filter model with the qAF procedure. Moreover, the likelihood of failing to detect this difference was found to be small, approximately 5%. Also important, the magnitude of the differences in ERBs, as a ratio of estimates from simultaneous versus forward masking, was consistent with past data at a value of 1.3 and the change in the value of p_u across masker conditions replicated earlier studies. The current results were consistent with the literature, and indicated that for one use of auditory filter estimation, the comparison of ERB estimates for simultaneous and forward maskers with young, normal-hearing listeners, the qAF procedure would be expected to be at least as sensitive as traditional procedures even though as few as 130 qAF trials sufficed for each estimate.

VI. GENERAL DISCUSSION

The current study demonstrated that the quick-auditory-filter (qAF) procedure, updated from the one tested in an earlier study (Shen and Richards, 2013), was able to produce estimates of auditory-filter shapes with significantly reduced testing time relative to traditional methods. Compared to its prototype, the current implementation of the qAF procedure included an updated computational algorithm (i.e., extended Kalman filtering) to allow simultaneous estimation of five model parameters and modifications to the experimental procedure (e.g., by introducing the warm-up periods) to allow reliable measurements from naive listeners.

Although the qAF procedure was successful for a population of young listeners, the application to other populations would require an understanding of the basic processes of the qAF procedure and the interaction between the procedure and the stimulus space. Two aspects of the process that experimenters should consider when applying the qAF procedure to a novel population are (1) the stimulus configuration and (2) the choice of auditory-filter models to be tested.

The stimulus configurations tested in an experiment are important for the stability of the qAF procedure. As an

example, to estimate the slope of the high-frequency skirt of the auditory filter (p_u), intuitively one would hope to sample stimuli as remote in frequency from the filter tip as possible. When estimating the auditory filter for high target levels, however, wide notch bandwidths cannot be tested because the masker spectrum level required to mask the target might exceed reasonable/safe test intensities. Using maskers with narrower notches, where the energy is near the filter tip, may slow the convergence of the p_u estimate. Similar considerations hold for the estimation of the shape of the lower-frequency portion of the auditory filter model. In either case, it is important to consider the level being tested and the ranges of masker levels and notch bandwidths.

To illustrate this point, a simulation was conducted in which a virtual listener described by a set of “true” parameters [$p_{u0}=42$, $p_{l0}=42$, $t_0=9$, $10 \log(w_0)=-35$, and $10 \log(K_0)=0$] was tested. The auditory-filter shapes were estimated using the qAF procedure configured identically to that used for experiments I and II. One hundred and fifty qAF trials were included in each simulated tracks, and 100 tracks were run. From these repetitions, the mean bias and the rms error for the ERB estimate were obtained. The resulting mean bias and rms error are shown in Fig. 7 (squares) as functions of the target level. The rms error for the ERB estimate (right panel) increased from below 10 Hz at a target level of 30 dB SPL to about 50 Hz at 60 dB SPL. At 50 and 60 dB SPL target levels, the estimated ERB’s were on average larger than the “true” ERB of the virtual observer (190.5 Hz in this simulation), showing a positive bias (left panel, squares). Therefore, the stability of the qAF procedure degrades as the target level increases due to the upper

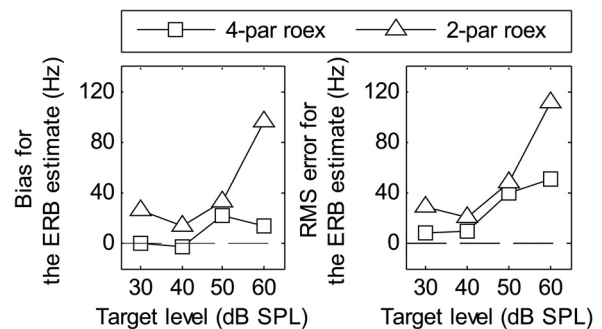


FIG. 7. The average bias (left) and rms errors (right) for the ERB estimates as functions of target level from a simulated virtual listener. The virtual listener’s auditory filter was described by a four-parameter roex function (as in the current experiments) with a set of “true” parameters. To estimate the auditory-filter parameters, the qAF procedure was implemented with either the same four-parameter roex model (squares) or a reduced, two-parameter roex model (triangles). The average bias and rms errors were derived by comparing the “true” ERB of the virtual listener to the ERB estimates obtained from 100 repeated qAF tracks, each of which contained 150 trials.

limit of the masker spectrum level set by the experimenter (50 dB SPL in the above simulation). It is worth pointing out that this limitation might be more evident when the auditory filter is estimated using a forward masking paradigm where the masker spectrum level at the detection threshold is usually far higher than for simultaneous masking.

A second important consideration for the success of the qAF procedure is the choice of the auditory filter model. In the traditional procedure, the models are fitted after all data are collected. Therefore, if a model was not successful in predicting psychophysical responses, it would only lead to an increase in error or large variability in the parameter estimates. However, the effect of an incorrect model choice might be over-emphasized for an adaptive qAF procedure. Not only would the parameter estimates be poor, but also the sampling of the stimuli might become inefficient. Therefore, it is recommended that the qAF procedure be implemented with established models and for experimental conditions where these models are known to provide satisfactory predictions of behavioral results.

Among various available models of the auditory filter (e.g., roex filter, Patterson and Nimmo-Smith, 1980; Patterson *et al.*, 1982; gammatone filter, de Boer, 1975; Allerhand *et al.*, 1992; Patterson *et al.*, 1995; gammachirp filter, Irino and Patterson, 1997), a good choice for a model is one with sufficient numbers of parameters to ensure the close encapsulation of behavioral data while not being too flexible so as to result in over-fitting. As an example, when fitting the low-frequency skirt of the auditory filter that exhibits a sharp tip section and a shallow tail section using an model with only a single slope parameter [e.g., the roex(p) function], the model cannot accurately represent the low-frequency side of the filter. In practice, the parameter estimation algorithm might settle on sampling mainly the tip section, mainly tail section, or vacillate between these two options (corresponding to large and small values of the single slope parameter p , respectively). These two modes (large or small values of p) may lead to a bimodal posterior parameter distribution, which would be poorly approximated by Gaussian distribution, potentially leading to instability in the Kalman filtering algorithm and the qAF procedure. Broadly speaking, it is difficult to achieve convergent parameter estimates when an inflexible model is used, and models with fewer parameters do not necessarily guarantee more efficient estimations of auditory-filter shapes. On the other hand, a model that is too flexible might be over-fitted because the data do not sufficiently constrain some or all parameters. This might lead to large variability for one or more of the parameter estimates, and subsequently poor test-retest reliability.

To demonstrate the effect of an inappropriately chosen model, the simulation described above was repeated using the same four-parameter roex-function virtual listener [$p_{i0} = 42$, $p_{l0} = 42$, $t_0 = 9$, $10 \log(w_0) = -35$, and $10 \log(K_0) = 0$]. The auditory-filter model implemented in the qAF procedure was a two-parameter roex function, which only consisted of two slope parameters, one on either side of the filter. One hundred qAF tracks were simulated, each containing 150 trials. The ERB estimate at the end of each track was compared to the “true” ERB of the virtual listener.

The mean bias and rms error for the ERB estimate are shown in the left and right panels of Fig. 7 (triangles), respectively. Compared to the results obtained using a matching four-parameter roex model (squares), larger positive biases and rms errors were observed for the two-parameter model. Moreover, for the mis-matched model the qAF procedure was more susceptible to errors at larger target levels. As the target level increased from 40 to 60 dB SPL, both the mean bias and rms error from the two-parameter model increased at a higher rate than for the four-parameter model.

The dip in RMS error from 30 to 40 dB SPL is highly reproducible and may reflect an interaction between the model choice and stimulus configuration. At the lowest target level, the estimates of the low-frequency filter slope would exhibit bimodal instability, mainly representing either the tip section (described by p_{l0}) or the tail section (described by t_0) of the “true” model. This bimodal instability is alleviated when the target level increases. For higher target levels, the upper limit of the masker spectrum level would lead to the sampling of notch widths within a narrow range near the filter tip, biasing the estimated model to represent the tip section of the filter. Because the ERB is determined by the tip section of the filter [Eq. (14)], frequent sampling of the stimulus near the filter tip would lead to an improvement in the ERB estimate.

In short, limitations on the stimulus search space and the use of inappropriate auditory-filter models could potentially lead to a relatively unstable qAF procedure. To take full advantage of the qAF procedure, it is to the experimenters’ advantage to configure the stimulus space and the model with care prior to data collection.

Although the current study demonstrated the potential of the qAF procedure as an efficient procedure for estimating the auditory filter in naive, normal-hearing listeners, additional development is also warranted. First, to provide an initial analysis of the convergence properties of the qAF procedure, a fixed number of trials was tested. Presumably, convergence criteria/tolerances based on one of several alternatives (e.g., the variance of the parameter estimates across recent trials, the variance of the stimuli across recent trials, or the variance of the posterior parameter distributions) will ultimately provide a more appropriate termination rule. Second, the usefulness of the qAF procedure would also benefit from a systematic investigation of the choices among alternative prior distributions. The current study used broad prior distributions for the parameters, minimizing the effects of the experimenter’s *a priori* beliefs regarding the parameter estimates. However, for narrower prior distributions, faster convergence rates would be expected. Third, for auditory filter shapes from very high or very low frequency regions, it may be advantageous to incorporate the effects of the middle-ear transfer function and to use maskers with upper and lower bands of unequal levels (e.g., Jurado and Moore, 2010). Or, one might wish to implement the qAF procedure so that the spectrum levels associated with the two masker bands are separately adjusted. Finally, the qAF procedure has not yet been tested using listeners with hearing impairment, whose

restricted dynamic ranges may pose difficulties for the qAF procedure.

VII. SUMMARY

A Bayesian adaptive procedure for the efficient estimation of the auditory-filter shapes, the qAF procedure, was presented and evaluated. Compared to its prototype described previously by Shen and Richards (2013), the current implementation of the qAF procedure allowed auditory-filter models with many parameters to be estimated. A series of experiments showed that excellent test-retest reliability could be achieved by the qAF procedure, even for listeners who were naive to the experimental task. At low target levels, the qAF procedure provided estimates of the auditory-filter parameters that were generally comparable to those reported from previous studies using traditional psychophysical techniques. Moreover, the qAF procedure was robust in demonstrating the known difference between the filter shapes estimated using simultaneous and forward masking paradigms. At high target levels, increases in individual differences and increases in the variability in the parameter estimates were observed. This might reflect instability in the qAF procedure that can occur when the most informative stimuli are situated outside the allowed presentation range.

ACKNOWLEDGMENTS

This work was supported by the Summer Undergraduate Research Program at the University of California, Irvine awarded to R.S. and the NIH Grant No. R21 DC013406 (co-PIs: V.M.R. and Y.S.).

¹A close inspection of the qAF tracks from the listeners with relatively poor ERB test-retest reliability did not reveal systematic patterns of stimuli or responses that might have contributed to differences in the parameter estimates.

²An alternative way of representing the group results is to average the predicted filter shapes across listeners instead of deriving the filter shapes from the averaged filter parameters. For the current experiment, these two methods did not lead to observable difference. For the ease of visualization, only the ones based on the averaged parameters are shown.

³For each up-down track, the initial level of the 2-kHz target was 60 dB SPL. The initial change in stimulus level was 8 dB, which was reduced to 5 dB after two reversals and then reduced to 2 dB following two additional reversals. The track terminated after a block of 50 trials and the average of the reversals, excluding the first four or five to ensure an even number of reversals, provided the estimate of threshold.

Allerhand, M., Butterfield, S., Cutler, A., and Patterson, R. (1992). "Assessing syllable strength via an auditory model," *Proc. Inst. Acoust.* **14**, 297.

Baker, R. J., and Rosen, S. (2006). "Auditory filter nonlinearity across frequency using simultaneous notched-noise masking," *J. Acoust. Soc. Am.* **119**, 454–462.

Charaziak, K. K., Souza, P., and Siegel, J. H. (2012). "Time-efficient measures of auditory frequency selectivity," *Int. J. Audiol.* **51**, 317–325.

de Boer, E. (1975). "Synthetic whole-nerve action potentials for the cat," *J. Acoust. Soc. Am.* **58**(5), 1030–1045.

- Dubno, J. R., and Dirks, D. D. (1989). "Auditory filter characteristics and Consonant recognition for hearing impaired listeners," *J. Acoust. Soc. Am.* **85**, 1666–1675.
- Fahrmeir, L. (1992). "Posterior mode estimation by extended Kalman filtering for multivariate dynamic generalized linear models," *J. Am. Stat. Assoc.* **87**, 501–509.
- Glasberg, B. R., and Moore, B. C. J. (1990). "Derivation of auditory filter shapes from notch-noise data," *Hear. Res.* **47**, 103–138.
- Irino, T., and Patterson, R. D. (1997). "A time-domain, level-dependent auditory filter: The gammachirp," *J. Acoust. Soc. Am.* **101**(1), 412–419.
- Jurado, C., and Moore, B. C. (2010). "Frequency selectivity for frequencies below 100 Hz: comparisons with mid-frequencies," *J. Acoust. Soc. Am.* **128**, 3585–3596.
- Kontsevich, L. L., and Tyler, C. W. (1999). "Bayesian adaptive estimation of psychometric slope and threshold," *Vision Res.* **39**, 2729–2737.
- Levitt, H. C. C. H. (1971). "Transformed up-down methods in psychoacoustics," *J. Acoust. Soc. Am.* **49**(2B), 467–477.
- Malicka, A. N., Munro, K. J., and Baker, R. J. (2009). "Fast method for psychophysical tuning curve measurement in school-age children," *Int. J. Audiol.* **48**, 546–553.
- Moore, B. C. J., Poon, P. W. F., Bacon, S. P., and Glasberg, B. R. (1987). "The temporal course of masking and the auditory filter shape," *J. Acoust. Soc. Am.* **81**, 1873–1880.
- Oxenham, A. J., and Shera, C. A. (2003). "Estimates of human cochlear tuning at low levels using forward and simultaneous masking," *J. Assoc. Res. Otolaryn.* **4**, 541–554.
- Patterson, R. D. (1976). "Auditory filter shapes derived with noise stimuli," *J. Acoust. Soc. Am.* **59**, 640–654.
- Patterson, R. D., Allerhand, M. H., and Giguere, C. (1995). "Time-domain modeling of peripheral auditory processing: A modular architecture and a software platform," *J. Acoust. Soc. Am.* **98**(4), 1890–1894.
- Patterson, R. D., and Nimmo-Smith, I. (1980). "Off-frequency listening and auditory-filter asymmetry," *J. Acoust. Soc. Am.* **67**, 229–245.
- Patterson, R. D., Nimmo-Smith, I., Weber, D. L., and Milroy, R. (1982). "The deterioration of hearing with age: frequency selectivity, the critical ratio, the audiogram, and speech threshold," *J. Acoust. Soc. Am.* **72**, 1788–1803.
- Rosen, S., Baker, R. J., and Darling, A. (1998). "Auditory filter nonlinearity at 2 kHz in normal hearing listeners," *J. Acoust. Soc. Am.* **103**, 2539–2550.
- Sek, A., Alcántara, J., Moore, B. C., Kluk, K., and Wicher, A. (2005). "Development of a fast method for determining psychophysical tuning curves," *Int. J. Audiol.* **44**, 408–420.
- Shen, Y., and Richards, V. M. (2013). "Bayesian adaptive estimation of the auditory filter," *J. Acoust. Soc. Am.* **134**, 1134–1145.
- Sommers, M. S., and Gehr, S. E. (1998). "Auditory suppression and frequency selectivity in older and younger adults," *J. Acoust. Soc. Am.* **103**, 1067–1074.
- Sommers, M. S., and Humes, L. E. (1993). "Auditory-filter shapes in normal-hearing, noise-masked normal and elderly listeners," *J. Acoust. Soc. Am.* **93**, 2903–2913.
- Unoki, M., Miyauchi, R., and Tan, C.-T. (2007). "Estimates of tuning of auditory filter using simultaneous and forward notched-noise masking," in *Hearing-From Sensory Processing to Perception*, edited by B. Kollmeier, G. Klump, V. Hohmann, U. Langemann, M. Mauermann, S. Uppenkamp, and J. Verhey (Springer, Berlin), pp. 19–26.
- Unoki, M., and Tan, C.-T. (2005). "Estimates of auditory filter shape using simultaneous and forward notched-noise masking," in *proceedings of Forum Acusticum Budapest*, 29 August–2 September, 2005, Budapest, Hungary, edited by Fulup Augusztinovic, Attila B. Bagy, and Zoltan Hunyadi (OPAKFI, Budapest), pp. 1497–1502.
- Vogten, L. L. M. (1978). "Simultaneous pure-tone masking: The dependence of masking asymmetries on intensity," *J. Acoust. Soc. Am.* **63**, 1509–1519.
- Wright, B. A., and Dai, H. (1994). "Detection of unexpected tones with short and long durations," *J. Acoust. Soc. Am.* **95**, 931–938.