# Temporal modulation transfer function for efficient assessment of auditory temporal resolution

Yi Shen[a)] and Virginia M. Richards
*Department of Cognitive Sciences, University of California, Irvine, California 92697-5100*

(Received 22 August 2012; revised 30 November 2012; accepted 5 December 2012)

Two common measures of auditory temporal resolution are the temporal modulation transfer function (TMTF) and the gap detection threshold (GDT). The current study addresses the lack of efficient psychophysical procedures for collecting TMTFs and the lack of literature on the comparisons of TMTF and GDT. Two procedures for efficient measurements of the TMTF are proposed: (1) A Bayesian procedure that adaptively chooses the stimulus modulation rate and depth to maximize the information gain from each trial and (2) a procedure that reduces the data collection to two adaptive staircase tracks. Results from experiments I and II showed that, for broadband carriers, these approaches provided similar results compared to TMTFs measured using traditional methods despite taking less than 10 min for data collection. Using these efficient procedures, TMTFs were measured from a large number of naive listeners and were compared to the gap detection thresholds collected from the same ears in experiment III. Results showed that the sensitivity parameter estimated from the TMTF measurements correlated well with the GDTs, whereas the cutoff rate is either uncorrelated or positively correlated with the gap detection threshold. These results suggest caution in interpreting a lower GDT as evidence for less sluggish temporal processing.
 [http://dx.doi.org/10.1121/1.4773271]

PACS number(s): 43.66.Yw, 43.66.Mk, 43.66.Ba [EB] Pages: 1031–1042

## I. INTRODUCTION

Sensitivity to fluctuations in the temporal envelopes of sounds is not only a fundamental property of our auditory system, it is also an important predictor of speech recognition performance in populations such as elderly listeners (Gordon-Salant and Fitzgibbons, 1993), hearing impaired listeners (Takahashi and Bacon, 1992), cochlear implant users (Fu, 2002; Won *et al.*, 2011), and patients with auditory neuropathy (Zeng *et al.*, 1999; Rance *et al.*, 2004). Therefore, developing efficient and reliable procedures to assess temporal acuity is of clinical interest. Toward this goal, the current study proposes two new procedures for the efficient measurement of temporal resolution.

In psychoacoustics, two dominant paradigms for measuring temporal resolution of the auditory system are the temporal modulation transfer function (TMTF) paradigm and the gap detection paradigm (see, e.g., Reed *et al.*, 2009). A temporal modulation transfer function is a function relating a listener's threshold for detecting sinusoidal amplitude modulation to modulation rate (see, e.g., Viemeister, 1979). For broadband-noise carriers, the TMTF typically exhibits a low-pass characteristic. That is, as the modulation rate ($f_m$) increases, the modulation detection threshold is initially stable, then increases after the modulation rate exceeds a cutoff rate ($f_c$). This cutoff rate has been traditionally described as reflecting the sluggishness of auditory temporal processing.[1] Although the behavioral estimation of the TMTF is of theoretical and clinical interest, it is usually very time-consuming. Typically, it involves measuring modulation detection thresholds at multiple modulation rates and then fitting the measured modulation detection thresholds with a temporal-integration model to estimate the parameters of the TMTF (e.g., the sensitivity $S$ and the cutoff rate $f_c$). This traditional approach usually requires more than 1000 trials of data collection and more than 2 h of listening time, which limits the wider application of the TMTF paradigm.

Besides the TMTF paradigm, an alternative way of assessing temporal resolution is the gap detection paradigm (see, e.g., Zwicker, 1965; Buunen and van Valkenburg, 1979; Fitzgibbons and Wightman, 1982; Shailer and Moore, 1983). Gap detection measures listeners' sensitivity to the presence of a silent temporal gap in a stimulus. The gap detection thresholds (GDTs) are thought to reflect temporal acuity because when the gap duration is short compared to temporal resolution of the auditory system, the gap could be partially or fully filled by the temporal integration process, making the gap harder to detect (see, e.g., Buunen and van Valkenburg, 1979). Because GDTs can be measured using a relatively small number of trials, it has been adopted widely in clinical research (see, e.g., Roberts and Lister, 2004; Lister and Roberts, 2005; Musiek *et al.*, 2005; Heinrich and Schneider, 2006; Pichora-Fuller *et al.*, 2006).

Although the TMTF and GDT can both be modeled using a leaky-integrator model of auditory temporal processing (see, e.g., Forrest and Green, 1987), it is not clear these two measures of temporal resolution are equivalent to one another. Formby and Muir (1988) reported a study where the bandwidths of the noise carriers were manipulated in both gap detection and TMTF measurements. This manipulation introduced within-subject variability in the GDT and the TMTF parameters. Studying within-subject variability, a negative correlation was observed between $S$ and GDT whereas

[a)]Author to whom correspondence should be addressed. Electronic address: shen.yi@uci.edu

$f_c$ was found to be relatively invariant with GDT. These observations suggest that the GDT is predictive of $S$, but not $f_c$. Therefore, it is likely that the TMTF provides more information regarding temporal processing than the GDT. Similar to this argument, it has been suggested that the GDT reflects both temporal and intensity resolution, whereas the TMTF paradigm allows separate assessments of temporal and intensity resolution (Strickland and Viemeister, 1997). Therefore, in many cases, the TMTF paradigm might be preferred over the gap detection paradigm.

In the current study, two new procedures for efficiently estimating the TMTF are described. In experiments I and II, TMTF estimates were collected using the new and traditional approaches to evaluate the usefulness of the new procedures. In experiment III, the TMTF estimates and GDTs were compared to provide a guideline for relating results from these two paradigms.

## II. EXPERIMENT I: A BAYESIAN PROCEDURE FOR ESTIMATING THE TMTF

### A. An entropy-based Bayesian adaptive algorithm

In this experiment, a Bayesian adaptive procedure is developed and used for TMTF measurements. This procedure does not estimate modulation detection thresholds, rather, it estimates the parameters of the TMTF directly. For each experimental trial, the modulation depth and rate for stimulus presentation are chosen so as to minimize the entropy in the estimates of these parameters. By doing so, the procedure adaptively optimizes the sampling strategy, leading to fast convergence of the posterior parameter distributions.

Formby and Muir (1988) and Eddins (1993) modeled the TMTF using a function of the form

$$\phi(f_m; S, f_c) = -S - 10\log_{10}\left[\frac{1}{1+(f_m/f_c)^2}\right], \tag{1}$$

where $\phi(f_m)$ is the modeled TMTF, $S$ is the sensitivity to amplitude modulation at low modulation rates, and $f_c$ is the cutoff rate of the TMTF. Let the threshold $\phi(f_m)$ be defined at the center of the dynamical range of the psychometric function, and assume the psychometric function takes the form of a logistic function. Then, one may express the probability of correctly detecting amplitude modulation as

$$p(x, f_m; S, f_c; \beta, \gamma) = \gamma + \frac{1-\gamma}{1+e^{-\beta[x-\phi(f_m; S, f_c)]}}, \tag{2}$$

where $x$ is the signal strength in the units of $20\log_{10} m$, with $m$ being the modulation depth; $\gamma$ represents chance performance, and is 0.5 for a two-alternative, forced-choice task; and $\beta$ gives the slope of the psychometric function. Although it has been found that the slope of the psychometric function for modulation detection experiments tends to increase slightly as $f_m$ increases (Eddins, 1993), this dependence is relatively small compared to the shift of threshold as a function of $f_m$. Therefore, we set the parameter $\beta$ to a fixed value of 1 throughout the following discussion. With the parameters $\gamma$ and $\beta$ being treated as constants, the proportion correct given in Eq. (2) may be written as $p(x, f_m; S, f_c)$ for simplicity.

The above-mentioned formulation reduces the problem of estimating the TMTF into finding the posterior distributions of the model parameters $\{S, f_c\}$, given the stimulus parameters $\{x, f_m\}$ and the listener's response on every trial. To solve this problem, a Bayesian adaptive procedure developed to estimate the psychometric function (Kontsevich and Tyler, 1999) was adopted and modified.

The procedure involves the following steps: First, a discrete parameter space is set up; the ranges of all parameters ($S, f_c, x,$ and $f_m$) and their gradations are set. Then, prior distributions for the TMTF parameters $p_1(S, f_c)$ are assigned. Following the $n$th trial, the expected posterior parameter distribution given a correct response being collected on the upcoming trial is

$$p_T(S, f_c | x, f_m) = \frac{p(x, f_m; S, f_c)p_n(S, f_c)}{\sum_{S, f_c} p(x, f_m; S, f_c)p_n(S, f_c)} \tag{3}$$

and the expected posterior parameter distribution given an incorrect response being collected on the upcoming trial is

$$p_F(S, f_c | x, f_m) = \frac{(1-p(x, f_m; S, f_c))p_n(S, f_c)}{\sum_{S, f_c}(1-p(x, f_m; S, f_c))p_n(S, f_c)}. \tag{4}$$

In Eqs. (3) and (4), $p_n(S, f_c)$ is the posterior distribution derived from the data collected up to the $n$th trial. If a correct response is measured in the upcoming trial, the entropy over the parameter space would be

$$H_T(x, f_m) = -\sum_{S, f_c} p_T(S, f_c | x, f_m)\log p_T(S, f_c | x, f_m), \tag{5}$$

but if the response is incorrect, the entropy would be

$$H_F(x, f_m) = -\sum_{S, f_c} p_F(S, f_c | x, f_m)\log p_F(S, f_c | x, f_m). \tag{6}$$

The overall expected entropy is then

$$H(x, f_m) = H_T(x, f_m)p(x, f_m; S, f_c) + H_F(x, f_m)(1 - p(x, f_m; S, f_c)). \tag{7}$$

In this approach the goal is to minimize the expected entropy $H(x, f_m)$. Therefore, the stimulus parameters tested on the upcoming $(n+1)$th trial are chosen as

$$\{x, f_m\} = \arg\min_{x, f_m} H(x, f_m). \tag{8}$$

If on the $(n+1)$th trial, a correct response is indeed measured, then the posterior distribution is updated as

$$p_{n+1}(S, f_c) = p_T(S, f_c | x, f_m) \tag{9}$$

otherwise,

$$p_{n+1}(S, f_c) = p_F(S, f_c | x, f_m). \tag{10}$$

Finally, parameter estimates after the $(n+1)$th trial are

$$\hat{S}_{n+1} = \sum_{S,f_c} S p_{n+1}(S,f_c), \quad \hat{f}_{c,n+1} = \sum_{S,f_c} f_c p_{n+1}(S,f_c). \tag{11}$$

After obtaining these parameter estimates, the whole process described previously is repeated until a prespecified total number of trials is reached.

Figure 1 illustrates the parameter estimates for a single track of the Bayesian procedure.[2] In this example, a virtual listener was used to generate simulated responses. The performance of the virtual listener was defined by Eq. (2) and a set of true TMTF parameters $\{S_0, f_{c,0}\} = \{20, 60\}$ (marked by a plus sign in the left-hand panel). The estimates of the $S$ and $f_c$ parameters converged fairly quickly during the first 60 trials or so (right-hand panels) and remained stable for the remainder of the trials. After 100 trials, the posterior distribution of $\{S, f_c\}$ was highly concentrated near $\{S_0, f_{c,0}\}$ (left-hand panel), indicating that an accurate estimate of the TMTF was achieved.

The Bayesian procedure described previously optimizes the stimulus placement strategy using an entropy-based criterion. That is, stimuli are presented in the regions of the stimulus parameter space that exhibit the lowest expected entropy. Figure 2 plots the expected entropies [Eq. (7)] as functions of modulation depth and modulation rate for a range of virtual listeners (see the true TMTF parameters $S_0$ and $f_{c,0}$ listed in each panels). For each of the virtual listeners, a low-entropy region was formed following the first 20 trials, which congregated close to the true TMTF (dashed curves). The minima in the entropy function occurred near the intersection between the true TMTF and the boundaries of the stimulus parameter space. Typically, one minimum was at the lowest modulation rate, whereas the other was at the highest modulation rate (8 and 512 Hz, respectively). Although the entropy function varied slightly as the number of trials increased, the locations of the two minima were always near the upper and lower "handles" of the TMTF. It seems that these two locations provide the most information regarding the parameters $S$ and $f_c$. Therefore, during a track of the Bayesian procedure, stimuli are mostly concentrated near these two handles of the TMTF.

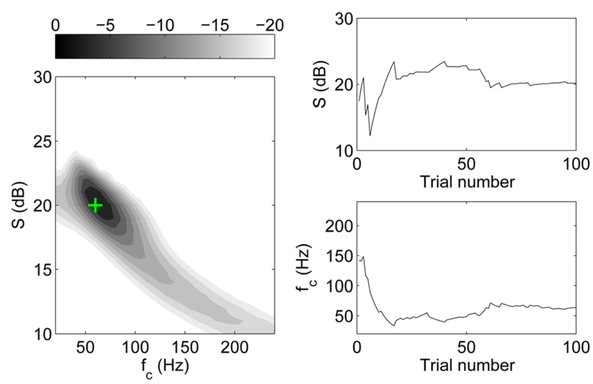


FIG. 1. (Color online) Estimates of $S$ and $f_c$ during a simulated track of the Bayesian adaptive procedure. The virtual listener used for the simulation is defined by $\{S_0, f_{c,0}\} = \{20, 60\}$ (marked by a cross symbol). The left-hand panel shows the posterior parameter distribution after 100 trials. Darker shading in this panel indicates higher probabilities. The numerical values for the natural logarithm of the posterior probability are indicated by the scale bar. The upper and lower right-hand panels plot the $S$ and $f_c$ estimates as a function of the trial number, respectively.

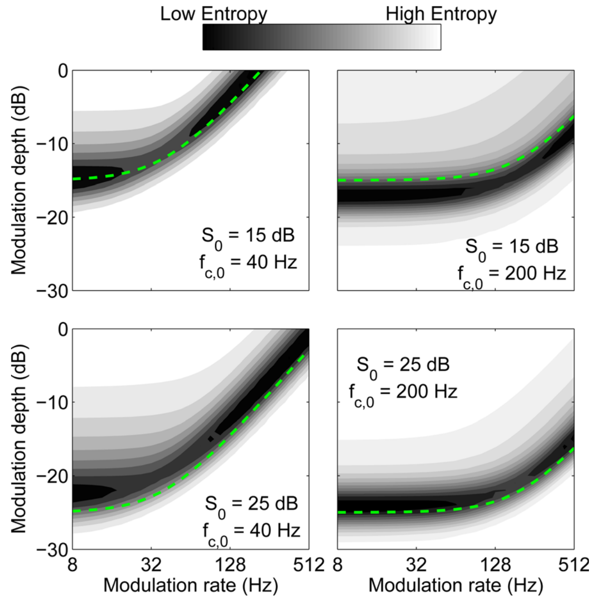


FIG. 2. (Color online) Expected entropies after the first 20 trials in the Bayesian adaptive procedure. The entropy functions are plotted as functions of modulation depth and modulation rate. In separate panels, results are shown for four different virtual listeners. Darker areas in each panel indicate lower entropies. Stimulus presentations are concentrated in these regions. The true TMTF is indicated using dashed lines.

Table I lists the parameters estimated using the Bayesian adaptive procedure for a range of virtual listeners. For each virtual listener, each estimate of the TMTF involved a track of 100 simulated trials, and the estimation was repeated 100 times ($N = 100$). Let $\{\hat{S}_i, \hat{f}_{c,i}\}$ denote the parameter estimate from the $i$th Bayesian adaptive track. The goodness of the parameter estimates was quantified by the root-mean-squared (rms) deviation from the estimated to the true parameters as follows:

TABLE I. Root-mean-squared deviations from the estimated to the true parameters for nine virtual listeners. $\{S_0, f_{c,0}\}$ are the true parameters for each virtual listener, $\{\hat{S}_1, \hat{f}_{c,1}\}$ are the parameter estimates from the first Bayesian adaptive track, and $D_S$ and $D_{fc}$ are root-mean-squared (rms) deviations from the estimated to the true parameters over 100 tracks as defined in Eq. (12).

| $\{S_0, f_{c,0}\}$ (dB, Hz) | $\{\hat{S}_1, \hat{f}_{c,1}\}$ (dB, Hz) | $D_S$ (dB) | $D_{fc}$ (Hz) |
|---|---|---|---|
| {10, 40} | {10.17, 38.57} | 0.41 | 2.99 |
| {15, 40} | {14.10, 44.69} | 0.56 | 3.76 |
| {20, 40} | {20.34, 35.97} | 0.56 | 4.28 |
| {10, 60} | {10.17, 57.14} | 0.48 | 4.96 |
| {15, 60} | {14.47, 67.19} | 0.54 | 5.51 |
| {20, 60} | {20.08, 59.19} | 0.59 | 6.82 |
| {10, 200} | {10.34, 183.18} | 0.54 | 16.14 |
| {15, 200} | {15.27, 194.68} | 0.51 | 17.52 |
| {20, 200} | {19.51, 189.51} | 0.53 | 18.44 |

$$D_S = \sqrt{\frac{\sum_{i=1}^{N}(\hat{S}_i - S_0)^2}{N}}, \quad D_{f_c} = \sqrt{\frac{\sum_{i=1}^{N}(\hat{f}_{c,i} - f_{c,0})^2}{N}}. \tag{12}$$

Even with only 100 trials, the Bayesian procedure provided very close estimates of the true TMTF parameters (i.e., the estimates $\{\hat{S}_i, \hat{f}_{c,i}\}$ listed in the second column of Table I are very close to the true parameters $\{S_0, f_{c,0}\}$ in the first column). The rms deviations $D_S$ were small, typically between 0.5 and 0.6 dB. The values of $D_{f_c}$ showed dependencies on both $S_0$ and $f_{c,0}$. When either the underlying true sensitivity $S_0$ or cutoff rate $f_{c,0}$ increased, $D_{f_c}$ increased. Nonetheless, the values of $D_{f_c}$ were typically below 10% of the $f_{c,0}$ value, therefore, the procedure provided fairly accurate TMTF estimates in despite of the dependencies of $D_{f_c}$ on $S_0$ and $f_{c,0}$.

Extending the promising simulation results presented previously, the current experiment tested the Bayesian adaptive tracking procedure in behavioral experiments. Temporal modulation transfer functions were collected from young, normal-hearing listeners using both the Bayesian and traditional procedures. The aim of the experiment is to determine whether the Bayesian procedure achieves results similar to those obtained using the traditional procedure, whereas immensely shortening the data-collection time.

### B. Subjects

Five young, normal-hearing listeners (S1–S6) participated in the current experiment. All listeners were between 18 and 35 years of age and had audiometric thresholds equal or better than 15 dB hearing level (HL) between 250 and 8000 Hz in both ears. The left ears of the listeners were tested in the experiment. None of the subjects received training before the data collection began, except for the first author (S3). Listeners were paid for their participation. The experiment was conducted in 2 h sessions. For each listener, no more than one session was run on a single day.

### C. Experimental procedures

Temporal modulation transfer functions were estimated using a four-interval, two-alternative forced-choice task. That is, on each trial, four sound intervals were presented, separated by 500-ms inter-stimulus intervals. Each sound interval contained a broadband noise, presented at an overall level of 70 dB sound pressure level (SPL). The duration of the noise was 500 ms, including 10 ms onset/offset raised-cosine ramps. In one of the intervals, the broadband noise was sinusoidally amplitude modulated:

$$s(t) = A[1 + m\sin(2\pi f_m t + \theta_m)]n(t), \tag{13}$$

where $n(t)$ is the noise carrier, $s(t)$ is the modulated stimulus, $t$ is time, $A$ is the amplitude, $m$ is the modulation depth, $f_m$ is the modulation rate, and $\theta_m$ is the initial phase of the modulation.

The amplitude $A$ was chosen to calibrate the stimulus to 70 dB SPL and compensate for the potential intensity cues following the introduction of the amplitude modulation. The initial phase $\theta_m$ was drawn from a uniform distribution spanning from 0 to $2\pi$ on a trial-by-trial basis. The signal interval always occurred on the second or the third interval, and listeners were instructed to indicate which one of the two intervals was the signal interval. This design was used to make the experiment intuitive for naive listeners.

Both the Bayesian adaptive tracking procedure and the traditional procedure were used for the estimation of the TMTF. Two tracks of the Bayesian procedure were run, before and after the data collection for the traditional procedure. Each Bayesian adaptive track contained 100 trials, which were divided into two blocks of 50 trials, and listeners were able to take a short break between the two blocks. Before each run of the Bayesian procedure, the parameter space was set up as follows. The values of $S$ ranged from 10 to 30 dB with 1 dB spacing, and the values of $f_c$ ranged from 30 to 200 Hz with 5 Hz spacing. The parameter $\beta$, which is related to the slope of the psychometric function was fixed at a value of 1. The values of $f_m$ took 25 logarithmically spaced values between 8 and 512 Hz, and the values of $x$ [as $20\log_{10}(m)$] ranged from $-30$ to 0 dB with 1 dB spacing. Uninformative prior distributions were implemented so that all potential values of $S$ and $f_c$ were assumed to be equally likely. That is, the prior probabilities for all possible combination of the $S$ and $f_c$ values were equal and summed to unity.

For the traditional method, the modulation detection threshold was measured at each of seven modulation rates (8, 16, 32, 64, 128, 256, and 512 Hz) using two-down, one-up adaptive tracks of 50 trials. On the first trial of each up–down track, the modulation depth was 0 dB. The modulation depth was decreased after two consecutive correct responses or increased after a single incorrect response with a constraint that the modulation depth was not allowed to exceed 0 dB. The initial step size for the changes in the modulation depth was 8 dB, and decreased to 5 dB after the first two reversals of the track, and decreased further to 2 dB after the first four reversals. After the 50 trials were collected, the modulation depths at the last four reversals were averaged to form a threshold estimate at the 70.7% correct point on the psychometric function (Levitt, 1971). In all instances more than eight reversals occurred. For each listener, the seven modulation rates were tested in random order, and then repeated three more times. The reported modulation thresholds were the average across the four threshold estimates from the four up–down tracks.

All stimuli were generated digitally at a sampling frequency of 44 100 Hz, and were presented monaurally to the left ears of listeners via a 24-bit soundcard installed in the experimental computer (Envy24 PCI audio controller, VIA Technologies, Inc., Fremont, CA), a programmable attenuator (PA4, Tucker-Davis Technologies, Inc., Alachua, FL), a headphone amplifier (HB6, Tucker-Davis Technologies, Inc.), and a Sennheiser HD410 SL headphone (Sennheiser Electronic Corp., Old Lyme, CT). Each stimulus presentation was followed by visual feedback indicating the correct response. During the experiment, listeners were seated in a double-walled, sound-attenuated booth.

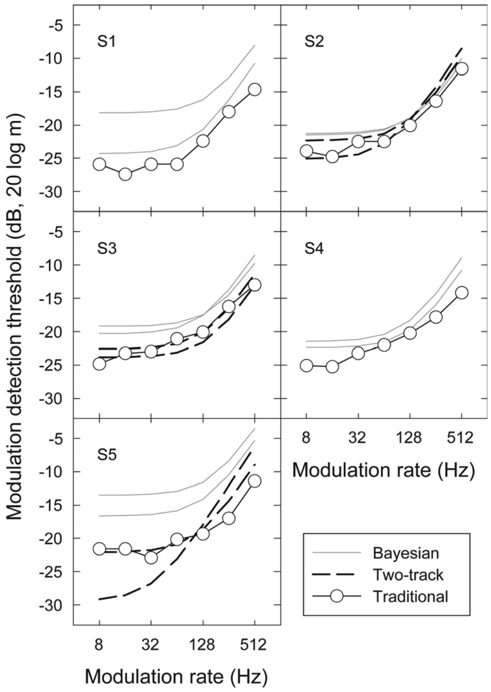


FIG. 3. TMTFs of individual listeners measured using the traditional procedure in experiment I, shown as unfilled symbols. Light solid curves indicate threshold estimates using the Bayesian procedure in experiment I. Dark dashed curves indicate threshold estimates using the two-track procedure in experiment II.

## D. Results and discussion

Modulation detection thresholds as a function of modulation rate, i.e., the TMTFs measured using the traditional procedure, are shown as open symbols in Fig. 3 for individual listeners. Across the five listeners, the TMTFs were similar; the standard errors of the mean of the threshold estimates were less than 1 dB at all modulation rates. These TMTFs also agreed well with previously published results measured using broadband carriers (see, e.g., Viemeister, 1979).

For each listener, the modulation detection thresholds were fitted with a first-order low-pass function [Eq. (1)], from which the sensitivity ($S$) and cutoff rate ($f_c$) were estimated. The estimated parameters are listed in Table II (in the rightmost two columns), together with the values of $S$ and $f_c$ estimated from the two runs of the Bayesian procedure. In general, the Bayesian procedure provided parameter estimates that were close to those estimated using the traditional procedure, though the traditional procedure tended to provide higher estimates of $S$. Using the parameter estimates with the traditional procedure as the reference ($S_0$ and $f_{c,0}$), the rms deviations across repetitions and listeners [Eq. (12) with $N=10$)] were 4.93 dB and 27.14 Hz for the $S$ and $f_c$ estimates, respectively. It should be noted, however, that some of the individual $f_c$ estimates were near the upper limit of the $f_c$ parameter space (200 Hz). It is possible that if higher values of $f_c$ were included in the parameter space, some of the $f_c$ estimates could have been higher and the rms deviation for $f_c$ could have been larger.

TABLE II. Parameter estimates from experiment I.

| | Bayesian (first run) | | Bayesian (second run) | | Traditional | |
|---|---|---|---|---|---|---|
| Listener | $S$ (dB) | $f_c$ (Hz) | $S$ (dB) | $f_c$ (Hz) | $S$ (dB) | $f_c$ (Hz) |
| S1 | 17.8 | 168.1 | 24.0 | 110.2 | 26.5 | 117.6 |
| S2 | 20.9 | 159.3 | 21.2 | 143.1 | 23.7 | 123.8 |
| S3 | 20.0 | 136.5 | 18.8 | 184.3 | 23.3 | 142.7 |
| S4 | 22.1 | 139.2 | 21.1 | 124.8 | 24.0 | 152.1 |
| S5 | 16.3 | 145.8 | 13.2 | 171.9 | 21.8 | 166.4 |

To study the agreement between the traditional and Bayesian procedures further, the parameter estimates obtained using the Bayesian procedure were used to estimate modulation detection thresholds at 70.7% correct by using Eq. (2). The resulting threshold estimates are plotted as light curves in Fig. 3. For all listeners, the threshold estimates from the Bayesian procedure resembled the shape of the TMTF measured using the traditional method, however, the Bayesian procedure tended to estimate thresholds at higher modulation depths than the traditional procedures (especially for listeners S1 and S5). This is consistent with the higher estimated values of $S$ (Table II).

Combining data from all individual listeners, the left-hand panel of Fig. 4 plot threshold estimates from the Bayesian procedure against those obtained using the traditional procedure. The threshold estimates based on the first and second tracks of the Bayesian procedure largely overlapped and were significantly correlated ($r=0.86$, $p<0.01$) with each other. The regression coefficients were 1.00 for the slope and $-0.05$ for the intercept, suggesting good repeatability of the Bayesian procedure. Because the first run of the Bayesian procedure was conducted without any previous training on the modulation detection task, the correlation between the two tracks suggests that the similar results are expected regardless of whether a listener is experienced or naive to the task. Comparing the threshold estimates using the two procedures, the Bayesian procedure always estimated thresholds that were higher than those provided by the traditional procedure. In some cases, this difference in threshold estimates was as large as 10 dB. Nonetheless, high correlations were found between the thresholds estimated using the two procedures for both the first ($r=0.84$, $p<0.01$) and the second tracks ($r=0.88$, $p<0.01$) of the

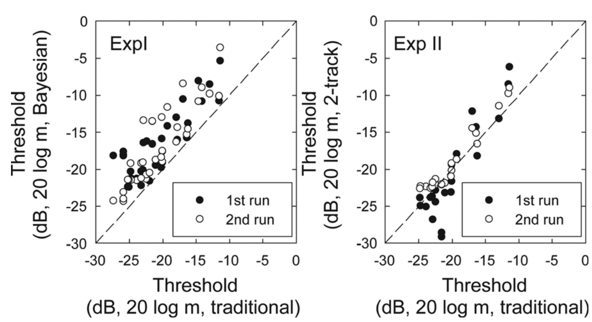


FIG. 4. Estimated modulation detection thresholds from the first (filled circles) and second (unfilled circles) tracks of the Bayesian procedure in experiment I (left-hand panel) and the two-track procedure in experiment II (right-hand panel). Thresholds estimated using the Bayesian and two-track procedures are plotted as functions of the thresholds obtained using the traditional procedure in the corresponding conditions and for the corresponding listener.

Bayesian procedure, indicating a consistent relationship between the two procedures.

Figure 5 shows the stimulus placement for all trials during experiment I as a scatter plot. Consistent with expectations (e.g., Fig. 2), the entropy-based search algorithm presented almost all stimuli at the highest (512 Hz) and the lowest (8 Hz) modulation rates. Only about 1% of the trials occurred at the modulation rates between 16 and 256 Hz. When overlaying the estimated TMTFs onto this scatter plot, it is quite clear that the stimuli were concentrated near the upper and lower handles of the TMTFs. It appears that these two locations on the TMTF provided the most information regarding the $S$ and $f_c$ parameters, which was consistent with the simulation results shown in Fig. 2.

An evaluation of the Bayesian adaptive tracks indicated that the algorithm sampled the stimulus space in a relatively random manner. Within an adaptive track, the stimulus presentations frequently switched between the two handles of the TMTF, more or less at random. Because the stimuli presented at high and low modulation rates sound very different, the leaps in modulation rate occurring unexpectedly may have confused listeners, introducing increases in internal decision noise. As a result, even for the same stimulus parameters ($x$ and $f_m$), the percent correct could be lower in the Bayesian procedure than in the traditional procedure. This might explain why the Bayesian procedure under-estimated the sensitivity $S$ and predicted higher modulation detection thresholds.

## III. EXPERIMENT II: A TWO-TRACK PROCEDURE FOR ESTIMATING THE TMTF

According to Figs. 2 and 5, the optimal places to sample the stimuli are at the upper and lower handles of the TMTF. If the biases in the estimates of $S$ obtained using the Bayesian procedure were associated with random changes in modulation rate, blocking the sampling of the two regions of interest (the two handles of the TMTF) into separate tracks should substantially reduce these biases. This consideration suggests a second approach, one in which traditional procedures are reduced to two adaptive staircase tracks, each associated with one handle. The viability of this two-track procedure was examined in the current experiment.

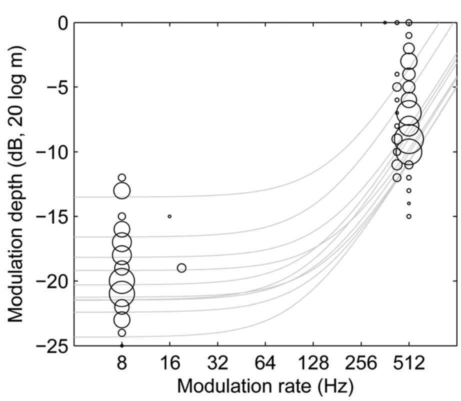


FIG. 5. Stimulus placement in two tracks of the Bayesian procedure in experiment I. Each circle indicates the occurrence of a particular combination of modulation rate and modulation depth in all trials, pooling across listeners and tracks. The size of the circle indicates the number of occurrences. The light curves are the estimated TMTF (as a 70.7% correct contour) from the five listeners for each of the two tracks, replotted from Fig. 3.

### A. Methods

Three of the five listeners from experiment I (S2, S3, and S5) participated in this experiment, which was conducted approximately 1 month after experiment I. The task and the stimuli used in experiment II were identical to those in experiment I. However, a new experimental procedure was used to collect the data.

Two runs of the two-track procedure were conducted. Each run consisted of two adaptive staircase tracks (a low-rate track and a high-rate track). Each of the two staircase tracks contained 50 trials. The low-rate track was always tested before the high-rate track. For the low-rate track, the modulation rate was fixed at 16 Hz,[3] and the modulation depth was varied adaptively to search for the performance threshold at 70.7% correct using a two-down, one-up algorithm (Levitt, 1971). The initial modulation depth was 0 dB, and the initial step size was 8 dB. The step size was reduced to 5 dB after the first two reversals, and was reduced further to 2 dB after the first four reversals. For the high-rate track, following the work of Formby (1985) and Strickland (2000), the modulation depth was fixed at −10 dB, and the modulation rate was varied adaptively to search for the threshold according to a two-down, one-up algorithm.[4] Variations in the modulation rate were on a logarithmic scale. That is, the modulation rate was decreased by a step factor following two consecutive correct responses and was increased by a step factor following a single incorrect response. The initial modulation rate was 32 Hz, and the initial step factor was 2.5. The step factor was reduced to 1.8 after the first two reversals, and was reduced further to 1.25 after the first four reversals.

For each run of the two-track procedure, after the stimuli and responses from the 100 trials were recorded (50 trials in the low-rate track and 50 trials in the high-rate track), the parameters $S$ and $f_c$ were estimated using a maximum-likelihood algorithm. The likelihood function following the $i$th trial $l_i(S,f_c)$ was $p_T(S,f_c|x,f_m)$ [Eq. (3)] if a correct response was recorded or $p_F(S,f_c|x,f_m)$ [Eq. (4)] if an incorrect response was obtained on that trial. Combining all 100 trials, the total likelihood function was

$$l(S,f_c) = \prod_i l_i(S,f_c). \tag{14}$$

The $S$ and $f_c$ that maximized the value of $l(S,f_c)$ were taken as the final parameter estimates. The search ranges of the parameters were from 10 to 30 for $S$ and from 25 to 250 for $f_c$.

### B. Results and discussion

Table III lists the estimated $S$ and $f_c$ values for the two runs of the two-track procedure and for individual listeners, together with the corresponding parameter estimates from experiment I using the traditional procedure. The estimated $S$ and $f_c$ were comparable between the two-track and the traditional procedures. Using the parameter estimates from the traditional procedure as the reference ($S_0$ and $f_{c,0}$), the rms

TABLE III. Parameter estimates from experiment II.

| | Two-track (first run) | | Two-track (second run) | | Traditional | |
|---|---|---|---|---|---|---|
| Listener | $S$ (dB) | $f_c$ (Hz) | $S$ (dB) | $f_c$ (Hz) | $S$ (dB) | $f_c$ (Hz) |
| S2 | 24.8 | 76.8 | 22.0 | 123.1 | 23.7 | 123.8 |
| S3 | 23.5 | 155.9 | 22.3 | 147.0 | 23.3 | 142.7 |
| S5 | 29.0 | 35.6 | 21.8 | 116.0 | 21.8 | 166.4 |

deviations [Eq. (12) with $N = 6$] were 2.40 dB and 46.97 Hz for the $S$ and $f_c$ estimates, respectively.

The estimated thresholds from the two-track procedure are shown as dashed curved in Fig. 3. For listeners S2 and S3, the two-track procedure, in both runs, provided very close resemblance of the TMTF measured using the traditional procedure. However, for listener S5, the first run of the two-track procedure failed to match the TMTF obtained using the traditional procedure. In this case, the largest deviation between the two procedures occurred at the lowest modulation rate, where a threshold difference of 8 dB was found.

The right-hand panel of Fig. 4 plots the threshold estimates using the two-track procedure against those obtained using the traditional procedure. These threshold estimates are plotted against those obtained using the traditional procedure in experiment I. The first and second runs of the two-track procedure showed a good agreement with a high correlation ($r = 0.95$, $p < 0.01$), suggesting good repeatability of the two-track procedure. The threshold estimates obtained using the two-track and traditional procedures were very close, and high correlations were found for both the first run ($r = 0.90$, $p < 0.01$) and the second run ($r = 0.97$, $p < 0.01$) of the two-track procedure.

One major difference between the Bayesian and two-track procedures was that the two-track procedure separated the high- and low-modulation rates into different tracks so that no leap of modulation rate would occur. Comparing the results from experiments I and II, the $S$ estimates and predicted threshold using the two-track procedure appear closer to those obtained using the traditional procedure than the Bayesian procedure. These results supported the hypothesis that the biases in the $S$ estimates using the Bayesian procedure were related to the leaps of modulation rate.

## IV. EXPERIMENT III: COMPARISONS BETWEEN TMTF AND GAP DETECTION MEASUREMENTS

Experiments I and II demonstrated that the two new approaches to estimate the TMTF, the Bayesian and two-track procedures, provide comparable results to those obtained using the traditional procedures while reducing the time for data collection from a few hours to ~10 min. As described previously, gap detection experiments provide an alternative paradigm to probe temporal acuity of the auditory system. With the procedures developed in experiments I and II, the estimation of the TMTF can be carried out as rapidly as gap detection measurements. However, there is an absence of data suggesting how the TMTF and gap detection results compare to each other. The current experiment was conducted to establish the relationship between the TMTF and GDT estimates among young, normal-hearing listeners. Experiment III was composed of two sub-experiments. In experiment IIIa, TMTF estimates using the Bayesian procedure were compared to GDTs; in experiments IIIb and IIIc, TMTF estimates obtained using the two-track procedure were compared to GDTs.

### A. Subjects

Forty-one normal-hearing undergraduate college students (S6–S46) participated in this experiment. All listeners were between 18 and 25 years of age and had audiometric thresholds equal to or better than 15 dB HL between 250 and 8000 Hz in both ears. Left ears of the listeners were tested in the current experiment. The 41 listeners were divided into three cohorts. Nineteen listeners (S6–24) participated in experiment IIIa, eight listeners (S25–S32) participated in experiment IIIb, and 14 listeners (S33–S46) participated in experiment IIIc. None of the subjects received training before the data collection began. Listeners received course extra credit for their participation. For each participant, the experiment was conducted in a single 1 h session.

### B. Stimuli

For the TMTF measurements in experiments IIIa and IIIb, the stimuli and procedures were identical to those described for experiments I and II, respectively. For the gap detection measurements, the detection of a silent gap in a broadband noise carrier was measured using a four-interval, two-alternative forced choice task. On each trial, four intervals were presented, separated by 500 ms inter-stimulus intervals. Each interval contained a 500 ms broadband noise presented at an overall level of 70 dB SPL. The noise was gated on and off using 5 ms raised-cosine ramps. In either the second or third interval, determined at random, a silent gap was imposed at the temporal center of the noise carrier. The gap was turned on and off using 5-ms ramps. The duration of the gap was defined from the half-amplitude point of its onset to that of its offset. The listeners were instructed to select the interval that contained the gap with the knowledge that it would appear only in the middle two of the four intervals. For both the TMTF and gap detection measurements in experiment III, the same system (as in experiments I and II) was used for stimulus generation and presentation. The stimuli used in experiment IIIc were identical to those in experiment IIIb, except that in the gap detection measurements, the gap was gated on and off using a ramp duration of 0.5 ms instead of 5 ms. The 5 ms ramp duration used in experiments IIIa and IIIb was much long than those used in previous studies of gap detection (see, e.g., Formby and Muir, 1988). To be consistent to previous gap detection studies, experiment IIIc was conducted using a much shorter ramp duration after the completion of experiments IIIa and IIIb.

### C. Procedure

In all experiments, listeners with odd code numbers (S7, S9,…, S41) ran two blocks of the TMTF measurements before two blocks of the gap detection measurements. Then, the process was repeated in reverse order. For listeners with even code numbers (S6, S8,…, S40), the gap detection was measured in the first two blocks followed by two blocks of the TMTF measurements. Then, the measurements were repeated once in reverse order. Each block contained 50 experimental trials.

For gap detection measurements, the gap duration was varied according to a two-down, one-up staircase procedure (Levitt, 1971). At the beginning of each block, the initial gap duration was 50 ms. The gap duration was reduced by a step factor after two consecutive correct responses and increased by a step factor after a single incorrect response. The initial step factor was 2.5, which was reduced to 1.8 after the first two reversals and reduced further to 1.25 after the first four reversals. The geometric mean of the gap durations of the last four reversals formed a threshold estimate. For all data collection, the total number of reversals exceeded eight. The ultimate GDTs were the geometric means across the threshold estimates from the four blocks.

For the TMTF measurements in experiment IIIa, two runs of the Bayesian procedure were tested, each of which consisted of 100 trials (or two blocks as in experiment I). The data collected from all 200 trials were pooled together and used to fit the model shown in Eq. (2) using the maximum-likelihood criterion [Eq. (14)]. For the TMTF measurements in experiments IIIb and IIIc, two runs of the two-track procedure were tested. Each run consisted of two blocks, the first of which contained a low-rate track and the second contained a high-rate track. As for experiment IIIa, all 200 trials of data were pooled, and the parameters $S$ and $f_c$ were estimated using Eq. (14).

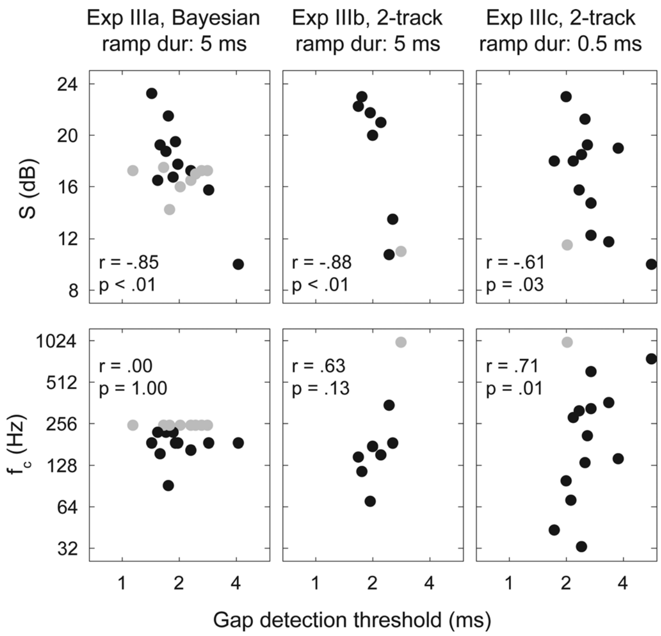


FIG. 6. Estimated sensitivity $S$ (upper panels) and cutoff rate $f_c$ (lower panels) as functions of the GDT for individual listeners in experiments IIIa (left), IIIb (middle), and IIIc (right). The upper limit for $f_c$ was 250 Hz for experiment IIIa, and was 1000 Hz for experiments IIIb and IIIc. Data from listeners with $f_c$ estimated at the upper limit are plotted in gray. Data from the rest of the listeners are shown as dark symbols, from which correlations between the TMTF parameters and GDTs were calculated. The label in each panel indicates the result from the correlational analysis in the corresponding experiment.

## D. Results

Results from experiments IIIa, IIIb, and IIIc are shown in the left-hand, middle, and right-hand panels of Fig. 6, respectively. For each experiment, the upper panel shows a scatter plot of the $S$ estimates against the GDTs, and the lower panel plots the $f_c$ estimates against the GDTs. Each circle indicates data for an individual listener.

For experiment IIIa, the $S$ estimated using the Bayesian procedure had a mean of 17.3 dB and a standard deviation of 2.7 dB. Most of the individual estimates of $S$ were clustered between 15 and 20 dB. On the other hand, the $f_c$ estimates exhibited large individual differences with a mean of 211.0 Hz and a standard deviation of 44.1 Hz. For 8 of the 27 listeners, the $f_c$ was estimated to be 250 Hz, which was the upper limit of the parameter space for $f_c$.[5]

For these eight listeners (referred to as group A, gray circles in the left-hand panels of Fig. 6), additional checks of the trial-by-trial data suggested that at modulation rates lower than 200 Hz, the average percent correct across the eight listeners was 61%, which was much lower than the rest of the listeners (Group B, dark circles in the left-hand panels of Fig. 6) who had an average percent correct of 78%. At high modulation rates, the average percent correct was 91% for group A and was 89% for group B. It seemed as though the listeners in group A concentrated on high modulation rates and discounted those trials with stimuli presented at low modulation rates. As mentioned earlier, the Bayesian adaptive track contained random leaps of modulation rates. One strategy to overcome such uncertainty in the stimuli would be to weight low modulation rates less than high modulation rates in the decision-making process, taking advantage of the fact that the stimuli were presented at the low modulation rates only 37% of the time. By this strategy, although chance performance would be expected at the low modulation rates, listeners could still maintain an overall performance level (combining trials at low and high rates) of approximately 70% correct. This unexpected strategy violated the assumption that the psychometric function was stationary with a slope of 1 with no lapses of attention regardless of modulation rate. Therefore, the data appear to have been inappropriately fitted, and the parameter estimates were unreliable (e.g., $f_c$ was estimated to be 250 Hz).

The estimated GDTs from experiment IIIa had a mean of 2.09 ms and a standard deviation of 0.68 ms. Potential correlations between the TMTF parameters and GDTs were investigated. Because the TMTF estimates for the listeners in group A were potentially unreliable, the correlation analysis was calculated for group B only. For this subgroup of listeners, $S$ and GDT were found to be significantly correlated ($r = -0.85$, $p < 0.01$), whereas no significant correlation was found between $f_c$ and GDT ($r = 0.00$, $p = 1.00$).

For experiment IIIb (middle panels of Fig. 6), the $S$ estimated using the two-track procedure had a mean of 17.9 dB and a standard deviation of 5.2 dB; the $f_c$ estimates had a mean of 274.7 Hz and a standard deviation of 304.0 Hz. Compared to experiment IIIa, the model [Eq. (2)] was fitted to the data using a wider $f_c$ parameter space with an upper limit of 1000 Hz.[6] Even with this high upper limit, $f_c$ was estimated to be 1000 Hz in one of the eight listeners (gray circles in the middle panels of Fig. 6). The estimated GDTs in experiment IIIb had a mean of 2.2 ms and a standard deviation of 0.4 ms. The correlations between the TMTF parameters and GDTs were calculated based on the seven listeners with estimated $f_c$'s less than 1000 Hz (dark circles in the middle panels of Fig. 6). Sensitivity $S$ and GDT were found to be negatively correlated ($r = -0.88$, $p < 0.01$), and no significant correlation was found between $f_c$ and GDT ($r = 0.63$, $p = 0.13$).

For experiment IIIc (right-hand panels of Fig. 6), the $S$ estimated using the two-track procedure had a mean of 17.1 dB and a standard deviation of 4.7 dB; the $f_c$ estimates had a mean of 311.9 Hz and a standard deviation of 283.0 Hz. The model parameters were estimated using identical procedures as in experiment IIIb. For one of the listeners, $f_c$ was estimated at 1000 Hz (gray circle in the right-hand panels of Fig. 6). The estimated GDTs in experiment IIIc had a mean of 2.7 ms and a standard deviation of 1.0 ms. The correlations between the TMTF parameters and GDTs were calculated based on the 13 listeners with estimated $f_c$'s less than 1000 Hz (dark circles in the right-hand panels of Fig. 6). Sensitivity $S$ and GDT were found to be negatively correlated ($r = -0.61$, $p = 0.03$). A positive correlation was found between $f_c$ and GDT ($r = 0.71$, $p = 0.01$).

## E. Discussion

Formby and Muir (1988) studied the within-subject correlations between the GDT and the TMTF parameters. A negative correlation was observed between $S$ and GDT, whereas $f_c$ was found to be relatively invariant with GDT. Consistent with this result, in experiments IIIa, IIIb, and IIIc, a negative correlation between $S$ and GDT was observed when the correlation analysis was based on between-subject variability.

The positive correlation between $f_c$ and GDT found in experiment IIIc, however, was unexpected. It has been argued that both the cutoff rate of the TMTF and the GDT are related to the time constant of a leaky-integrator model of temporal processing (see, e.g., Forrest and Green, 1987). According to the model, a longer time constant corresponds to a more sluggish auditory system, which would consequently lead to lower cutoff rates and longer GDTs. Therefore, a negative correlation between $f_c$ and GDT is expected, which was not observed in experiments IIIa and IIIb and was opposite to the findings of experiment IIIc. Due to the unexpected results of experiment IIIc, it is of interest to verify whether a leaky-integrator model would indeed fail to predict the relationship between $f_c$ estimated from the TMTF measurements and GDT estimated from the gap detection measurements.

To investigate expectations from the leaky-integrator model of temporal processing regarding the relationship between $f_c$ and GDT, a model was implemented following the work of Viemeister (1979) and Forrest and Green (1987). The resulting model consisted of an initial bandpass filter (with its passband spanning from 4 to 6 kHz), a half-wave rectifier, and a first-order low-pass filter with a cutoff rate of $f_c$. The model was used to predict GDTs using the following procedure. First, an array of gap durations was chosen (1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, and 5 ms). At each gap duration, 400 experimental trials were presented to the model. For each trial, a noise stimulus that did not contain the gap (in the no-signal interval) and a noise stimuli that had the gap (in the signal interval) were presented to the model. A max/min decision variable was then calculated based on the output for each of the two stimuli (Forrest and Green, 1987). Across the 400 trials, the means and variances for the two decision variables were obtained, from which a $d'$ statistic was calculated at each gap duration. Defining the GDT at $d' = 1$, predictions were calculated through linear interpolation. The above-described procedure was repeated at $f_c$ values of 16, 32, 64, 128, 256, 512, and 1024 Hz.

Figure 7 plots the predicted GDTs as functions of $f_c$ as unfilled symbols. The two panels of the figure correspond to experiments IIIb (left-hand panel) and IIIc (right-hand panel). Comparing across the two panels, the leaky-integrator model predicted a strong interaction between ramp duration and $f_c$ on GDT. When the ramp duration was 5 ms (in experiment IIIb, left-hand panel), the predicted GDT decreased from 2.62 ms at 16 Hz to 2.00 ms at 128 Hz, then returned to 2.99 ms at 1024 Hz.[7] On the other hand, when the ramp duration was 0.5 ms (in experiment IIIc, right-hand panel), the GDT, in general, decreased monotonically.

Also plotted in Fig. 7 (using filled symbols) are the experimental data for individual listeners obtained from experiments IIIb and IIIc. Although the predicted and measured data were visually similar for experiment IIIb (comparing filled and unfilled symbols in the left-hand panel), the leaky-integrator model failed to capture the experimental data obtained in experiment IIIc (comparing filled and unfilled symbols in the right-hand panel). Thus, the

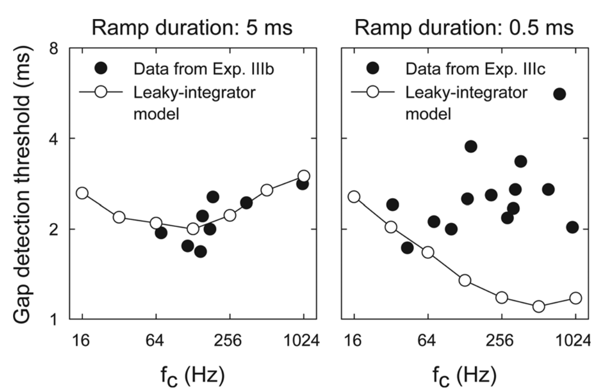


FIG. 7. Gap detection thresholds plotted against the $f_c$ estimates from the TMTF measurements (filled symbols). Results from experiments IIIb and IIIc are arranged in the left- and right-hand panels, respectively. Unfilled symbols indicate predictions from a leaky-integrator model of temporal processing in the corresponding conditions.

predicted interaction between ramp duration and $f_c$ was not observed in the results of experiments IIIb and IIIc. Therefore, the leaky integrator model, as implemented here, failed to reliably predict GDT from the $f_c$ parameter of the TMTF.

In summary, the current experiments consistently demonstrated significant negative correlations between GDT and the sensitivity parameter $S$ of the TMTF (upper panels in Fig. 6). Because $S$ primarily reflects listeners' performance at low modulation rates, the GDT is likely to be closely associated with sensitivity to low-rate amplitude modulation. On the other hand, the relationship between $f_c$ and GDT observed in the current experiments is not likely to be fully explained by the leaky-integrator model of temporal processing (right-hand panel of Fig. 7). Therefore, the current results did not support the direct interpretation of the GDT as a measure of the sluggishness of auditory temporal processing.

## V. POTENTIAL EXTENSIONS

The current study presented two psychophysical procedures for the efficient measurement of the temporal modulation transfer function. These procedures aim to provide estimates of parameters ($S$ and $f_c$) for a first-order, low-pass model of auditory temporal processing (or leaky-integrator model). Therefore, the major assumption underlying these newly developed procedures is that the leaky-integrator model provides a close description of temporal processing in the auditory system. For the current stimulus configurations (broadband stimuli), many previous published experimental results support this assumption (e.g., Viemeister, 1979; Forrest and Green, 1987; Eddins, 1993). On the other hand, studies that have used narrowband stimuli suggest that the leaky-integrator model is not always sufficient in describing modulation detection data (e.g., Dau *et al.*, 1997a,b, 1999; Ewert and Dau, 2000; Verhey, 2002).

In such cases, more sophisticated models might be required. One successful example is the modulation-filter-bank model (see, e.g., Münkner and Püschel, 1993; Dau *et al.*, 1997a,b; Chi *et al.*, 1999), which describes envelope processing as a bank of band-pass filters (modulation filters). The introduction of these modulation filters is inspired by the physiological finding that most neurons at the level of the inferior colliculus demonstrate band-pass tuning in the modulation domain (Langner and Schreiner, 1988; Krishna and Semple, 2000). Much like the auditory filters in the frequency domain (see, e.g., Fletcher, 1940), the bandwidths of the modulation filters determine the spectral resolution in the modulation domain. The model has been successfully applied to a wide range of auditory tasks proving temporal processing, including forward masking, gap detection, TMTF, and modulation masking (Münkner and Püschel, 1993; Dau *et al.*, 1997a,b; Verhey, 2002).

In the framework of modulation-filter-bank model, a TMTF stimulus would be processed mainly by the modulation filter at the modulation rate. Because different modulation rates could be processed independently, the modulation-filter-bank model allows assigning different decision weights to low versus high rates. This mechanism parallels the suggested differences in decision strategy between listener groups A and B in experiment IIIa. Moreover, a gap detection stimulus could be processed mainly by modulation filters at low modulation rates. Therefore, the modulation-filter-bank model should be able to capture the close associate between GDT and modulation processing at low rates ($S$). In all, the modulation-filter-bank model appears to have sufficient flexibility to encapsulate many features of the TMTF and gap detection data.

In principle, efficient Bayesian psychophysical procedures for the estimation of the modulation filter-bank parameters could be developed. However, the cost of introducing the modulation filter bank, is an increase in the number of free model parameters. Fitting modulation filter-bank model using experimental paradigms such as broadband modulation detection or gap detection used in the current study would likely lead to over-parameterization. That is, the experimental data might not provide sufficient information to allow the convergence of a unique solution in the high-dimensional parameter space. Given limited evidence at present, future investigations are required to reveal whether extending the current procedures to those based on the modulation-filter-bank theory is advantageous.

## VI. SUMMARY AND CONCLUSIONS

In the current study, two procedures were developed for the efficient assessment of the temporal modulation transfer function. Both procedures adjust stimuli adaptively to maximize the information gain regarding the TMTF parameters $S$ and $f_c$. One procedure (the Bayesian procedure) adopted the Bayesian adaptive algorithm described by Kontsevich and Tyler (1999), while the other procedure (the two-track procedure) utilized two up−down tracks, each of which targets a specific region in the stimulus space. The TMTF estimates were compared with results from the TMTF measurements using the traditional procedure and from gap detection measurements. The results indicated the following:

1. Although the Bayesian adaptive procedure is theoretically sound and demonstrates excellent efficiency in computational simulations, it underestimated $S$ when human listeners were tested. The estimated TMTFs predicted higher modulation detection thresholds than the experimental data using the traditional procedure.
2. In contrast to the Bayesian procedure, the estimated TMTFs using the two-track procedure predicted the modulation detection thresholds that were more consistent with the traditional procedure.
3. For both the Bayesian and two-track procedure, the estimated $S$ was negatively correlated with GDT, whereas no correlation or a positive correlation was found between $f_c$ and GDT. These results suggest that the gap detection paradigm mainly probes listeners' sensitivity to low-rate envelope fluctuations. A low GDT does not necessarily mean that the listener has a larger value of $f_c$ for his/her TMTF.

## ACKNOWLEDGMENTS

This work was supported by NIH Grant No. R21 DC010058. The authors thank the two anonymous reviewers for their insightful comments and Theodore Lin and Andrew Silva for their assistance in data collection and preparation of the manuscript.

[1]Note that other factors besides the sluggishness of the auditory system, such as the inherent envelope fluctuations in narrowband carriers, could also affect the TMTF (see, e.g., Dau *et al*., 1999; Stellmack *et al*., 2005). Similar influences of inherent fluctuations have also been found for the gap detection paradigm (e.g., Grose *et al*., 2008).

[2]For the simulations presented in this section, including data shown in Figs. 1 and 5 and Table I, the initialization of the parameter space was identical to that used in the behavioral experiment (see Sec. II C), except that the range for $S$ was between 0 and 30 dB with 1 dB spacing in the simulations.

[3]A modulation rate of 16 Hz was used here instead of 8 Hz to limit the influence of the gated stimulus presentation on the TMTF estimates. Viemeister (1979) showed that when the stimuli were gated on and off at the onset and offset of the amplitude modulation, the modulation detection threshold could be higher compared to those measured using continuous stimuli. This effect of gating mainly affected thresholds at low modulation rates [typically less than 16 Hz for a stimulus duration of 500 ms, see Fig. 6 in Viemeister (1979)]. However, based on the TMTF collected using the traditional procedure (unfilled circles in Fig. 3), the difference between the modulation detection thresholds at 8 and 16 Hz was negligibly small. Therefore, similar parameter estimates (especially for $S$) would be expected even if a modulation rate of 8 Hz was used during the low-rate track of the two-track procedure in the current experiment.

[4]The stimulus manipulations in the high-rate track varied the modulation rate instead of modulation depth based on practical considerations. If the high-rate track was based on the modulation depth and the modulation rate was fixed at a high rate (e.g., 512 Hz), the track would be limited to a very narrow dynamic range (typically from $-10$ to 0 dB). This dynamic range would be even smaller if the listener had a low value of $f_c$. On the other hand, by varying the modulation rate, one is assured that overmodulation ($x > 0$) will not occur, and the dynamic range of the adaptive track is large enough that the initial stimulus could be presented well above the performance threshold. This procedure has been successfully used in previous studies (Formby, 1985; Strickland, 2000) to improve the estimates of the TMTF at high modulation rates.

[5]For a significant proportion of the listeners in experiment IIIa (8 out of 27) the estimated $f_c$ was at the upper limit of the $f_c$ parameter space, therefore the standard deviation of the $f_c$ estimates across listeners was small (44.1 Hz). Had the experiment been repeated with a higher upper limit for the $f_c$ parameter space, a much larger standard deviation would be expected.

[6]The Bayesian procedure requires the configuration of the parameter space to be set prior to data collection, whereas the two-track procedure allows setting up the parameter space after the completion of the experiment. To reduce the chance of estimating the $f_c$ parameter at its highest possible value (as occurred for listener group A in experiment IIIa), a high upper limit of 1000 Hz was used for fitting the model in experiments IIIb and IIIc.

[7]The non-monotonic dependence of the GDT on $f_c$ predicted by the leaky integrator for a ramp duration of 5 ms can be understood by considering the spectrum of the gap detection stimuli in the modulation domain. The modulation power associated with the presence of the gap occurs predominantly at modulation rates below 100 Hz, which is determined by the 5 ms ramp duration. As the $f_c$ of the leaky integrator increases from very small value toward 100 Hz, detection of the gap improves because more modulation power is allowed to pass through the low-pass filter. On the other hand, when $f_c$ is larger than 100 Hz, the modulation power associated with the gap would be integrated with noises located beyond 100 Hz, given by the inherent envelope fluctuation of the carrier. In such cases, detection of the gap degrades as $f_c$ increases further from 100 Hz.

Buunen, T. J., and van Valkenburg, D. A. (**1979**). "Auditory detection of a single gap in noise," J. Acoust. Soc. Am. **65**, 534–536.

Chi, T., Gao, Y., Guyton, M. C., Ru, P., and Shamma, S. (**1999**). "Spectro-temporal modulation transfer functions and speech intelligibility," J. Acoust. Soc. Am. **106**, 2719–2732.

Dau, T., Kollmeier, B., and Kohlrausch, A. (**1997a**). "Modeling auditory processing of amplitude modulation. I. Detection and masking with narrow-band carriers," J. Acoust. Soc. Am. **102**, 2892–2905.

Dau, T., Kollmeier, B., and Kohlrausch, A. (**1997b**). "Modeling auditory processing of amplitude modulation. II. Spectral and temporal integration," J. Acoust. Soc. Am. **102**, 2906–2919.

Dau, T., Verhey, J., and Kohlrausch, A. (**1999**). "Intrinsic envelope fluctuations and modulation-detection thresholds for narrow-band noise carriers," J. Acoust. Soc. Am. **106**, 2752–2760.

Eddins, D. A. (**1993**). "Amplitude modulation detection of narrow-band noise: Effects of absolute bandwidth and frequency region," J. Acoust. Soc. Am. **93**, 470–479.

Ewert, S. D., and Dau, T. (**2000**). "Characterizing frequency selectivity for envelope fluctuations," J. Acoust. Soc. Am. **108**, 1181–1196.

Fitzgibbons, P. J., and Wightman, F. L. (**1982**). "Gap detection in normal and hearing impaired listeners," J. Acoust. Soc. Am. **72**, 761–765.

Fletcher, H. (**1940**). "Auditory patterns," Rev. Mod. Phys. **12**, 47–61.

Formby, C. (**1985**). "Differential sensitivity to tonal frequency and to the rate of amplitude modulation of broadband noise by normally hearing listeners," J. Acoust. Soc. Am. **78**, 70–77.

Formby, C., and Muir, K. (**1988**). "Modulation and gap detection for broad-band and filtered noise signals," J. Acoust. Soc. Am. **84**, 545–550.

Forrest, T. G., and Green, D. M. (**1987**). "Detection of partially filled gaps in noise and the temporal modulation transfer function," J. Acoust. Soc. Am. **82**, 1933–1943.

Fu, Q.-J. (**2002**). "Temporal processing and speech recognition in cochlear implant users," Neuroreport **13**, 1635–1639.

Gordon-Salant, S., and Fitzgibbons, P. J. (**1993**). "Temporal factors and speech recognition performance in young and elderly listeners," J. Speech Hear. Res. **36**, 1276–1285.

Grose, J. H., Buss, E., and Hall, J. W. (**2008**). "Gap detection in modulated noise: across frequency facilitation and interference," J. Acoust. Soc. Am. **123**, 998–1007.

Heinrich, A., and Schneider, B. (**2006**). "Age-related changes in within and between-channel gap detection using sinusoidal stimuli," J. Acoust. Soc. Am. **119**, 2316–2326.

Kontsevich, L. L., and Tyler, C. W. (**1999**). "Bayesian adaptive estimation of psychometric slope and threshold," Vision Res. **39**, 2729–2737.

Krishna, B. S., and Semple, M. N. (**2000**). "Auditory temporal processing: responses to sinusoidally amplitude-modulated tones in the inferior colliculus," J. Neurophysiol. **84**, 255–273.

Langner, G., and Schreiner, C. E. (**1988**). "Periodicity coding in the inferior colliculus of the cat. I. Neuronal mechanisms," J. Neurophysiol. **60**, 1799–1822.

Levitt, H. (**1971**). "Transformed up-down methods in psychoacoustics," J. Acoust. Soc. Am. **49**(Suppl. 2), 467–477.

Lister, J. J., and Roberts, R. A. (**2005**). "Effects of age and hearing loss on gap detection and the precedence effect: narrow-band stimuli," J. Speech Lang. Hear. Res. **48**, 482–493.

Münkner, S., and Püschel, D. (**1993**). "A psychoacoustical model for the perception of non-stationary sounds," in *Contributions to Psychological Acoustics*, edited by A. Schick (Universitätsgesellschaft Oldenburg, Oldenburg), pp. 121–134.

Musiek, F. E., Shinn, J. B., Jirsa, R., Bamiou, D.-E., Baran, J. A., and Zaida, E. (**2005**). "GIN (gaps-in-noise) test performance in subjects with confirmed central auditory nervous system involvement," Ear. Hear. **26**, 608–618.

Pichora-Fuller, M. K., Schneider, B. A., Benson, N. J., Hamstra, S. J., and Storzer, E. (**2006**). "Effect of age on detection of gaps in speech and non-speech markers varying in duration and spectral symmetry," J. Acoust. Soc. Am. **119**, 1143–1155.

Rance, G., McKay, C., and Grayden, D. (**2004**). "Perceptual characterization of children with auditory neuropathy," Ear. Hear. **25**, 34–46.

Reed, C. M., Braida, L. D., and Zurek, P. M. (**2009**). "Review article: review of the literature on temporal resolution in listeners with cochlear hearing impairment: a critical assessment of the role of suprathreshold deficits," Trends Amplif. **13**, 4–43.

Roberts, R. A., and Lister, J. J. (**2004**). "Effects of age and hearing loss on gap detection and the precedence effect: broadband stimuli," J. Speech Lang. Hear. Res. **47**, 965–978.

Shailer, M. J., and Moore, B. C. (**1983**). "Gap detection as a function of frequency, bandwidth, and level," J. Acoust. Soc. Am. **74**, 467–473.

Stellmack, M. A., Viemeister, N. F., and Byrne, A. J. (**2005**). "Monaural and interaural temporal modulation transfer functions measured with 5-kHz carriers," J. Acoust. Soc. Am. **118**, 2507–2518.

Strickland, E. A. (**2000**). "The effects of frequency region and level on the temporal modulation transfer function," J. Acoust. Soc. Am. **107**, 942–952.

Strickland, E. A., and Viemeister, N. F. (**1997**). "The effects of frequency region and bandwidth on the temporal modulation transfer function," J. Acoust. Soc. Am. **102**, 1799–1810.

Takahashi, G. A., and Bacon, S. P. (**1992**). "Modulation detection, modulation masking, and speech understanding in noise in the elderly," J. Speech Hear. Res. **35**, 1410–1421.

Verhey, J. L. (**2002**). "Modeling the influence of inherent envelope fluctuations in simultaneous masking experiments," J. Acoust. Soc. Am. **111**, 1018–1025.

Viemeister, N. F. (**1979**). "Temporal modulation transfer functions based upon modulation thresholds," J. Acoust. Soc. Am. **66**, 1364–1380.

Won, J. H., Drennan, W. R., Nie, K., Jameyson, E. M., and Rubinstein, J. T. (**2011**). "Acoustic temporal modulation detection and speech perception in cochlear implant listeners," J. Acoust. Soc. Am. **130**, 376–388.

Zeng, F. G., Oba, S., Garde, S., Sininger, Y., and Starr, A. (**1999**). "Temporal and speech processing deficits in auditory neuropathy," Neuroreport **10**, 3429–3435.

Zwicker, E. (**1965**). "Temporal effects in simultaneous masking and loudness," J. Acoust. Soc. Am. **38**, 132–141.