Binaural processing model based on contralateral inhibition. II. Dependence on spectral parameters

Jeroen Breebaart[a)]
IPO, Center for User–System Interaction, P.O. Box 513, NL-5600 MB Eindhoven, The Netherlands

Steven van de Par and Armin Kohlrausch
IPO, Center for User–System Interaction, P.O. Box 513, NL-5600 MB Eindhoven, The Netherlands and Philips Research Laboratories Eindhoven, Prof. Holstlaan 4, NL-5656 AA Eindhoven, The Netherlands

(Received 23 May 2000; revised 18 December 2000; accepted 3 May 2001)

This and two accompanying articles [Breebaart *et al.*, J. Acoust. Soc. Am. **110**, 1074–1088 (2001); **110**, 1105–1117 (2001)] describe a computational model for the signal processing in the binaural auditory system. The model consists of several stages of monaural and binaural preprocessing combined with an optimal detector. In the present article the model is tested and validated by comparing its predictions with experimental data for binaural discrimination and masking conditions as a function of the spectral parameters of both masker and signal. For this purpose, the model is used as an artificial observer in a three-interval, forced-choice adaptive procedure. All model parameters were kept constant for all simulations described in this and the subsequent article. The effects of the following experimental parameters were investigated: center frequency of both masker and target, bandwidth of masker and target, the interaural phase relations of masker and target, and the level of the masker. Several phenomena that occur in binaural listening conditions can be accounted for. These include the wider effective binaural critical bandwidth observed in band-widening NoSπ conditions, the different masker-level dependence of binaural detection thresholds for narrow- and for wide-band maskers, the unification of IID and ITD sensitivity with binaural detection data, and the dependence of binaural thresholds on frequency. [DOI: 10.1121/1.1383298]

PACS numbers: 43.66.Pn, 43.66.Ba, 43.66.Dc [DWG]

I. INTRODUCTION

This article describes and discusses simulations of binaural detection tasks with a binaural processing model which is described in detail in the preceding article (Breebaart *et al.*, 2001a). This model basically consists of three stages. The first stage simulates the effective signal processing of the basilar membrane and the inner haircells and includes adaptation by means of adaptation loops (Dau *et al.*, 1996a). Binaural interaction is modeled in the second stage by means of a contralateral inhibition mechanism: the model computes the squared difference signal between the left and right ears as a function of time, frequency channel, internal interaural delay (τ in seconds), and internal interaural level adjustment (α in dB). These binaural signals are corrupted by internal noise and subsequently analyzed by the third stage in the model, the central processor. The model is used as an artificial observer in a three-interval, forced-choice procedure, in which the central processor matches the representations of the presented stimuli to templates (derived during previous presentations), and on this basis the model indicates which interval contains the signal.

The scope of the simulations of this article is restricted to binaurally ''stationary'' stimuli. For these cases, the spectral parameters and the interaural phase relations of the stimuli were not varied as a function of time, and the stimuli had a duration which was long in comparison to the temporal resolution of the auditory system. First, the ability of the model to capture some basic properties of binaural hearing is demonstrated. These include interaural intensity difference (IID) and interaural time difference (ITD) sensitivity and binaural detection performance of tones in noise as a function of the bandwidth, center frequency, and interaural phase relationships of the stimuli.

The second and major focus of this article is on the apparently wider critical bandwidth in binaural conditions that has been found in a number of studies. If an interaurally out-of-phase signal (Sπ) must be detected against an No masker of variable bandwidth, the estimate of the critical bandwidth is two to three times the estimate which is found in monaural experiments (Sever and Small, 1979; Zurek and Durlach, 1987; van de Par and Kohlrausch, 1999). Hall *et al.* (1983) found that this descrepancy between monaural and binaural estimates is largest at high masker levels. Furthermore, the wider effective bandwidth is only observed if the interaural cross correlation of the masker is very close to +1 (van der Heijden and Trahiotis, 1998). On the other hand, experiments that use a masker with frequency-dependent interaural phase relations reveal a critical band estimate that basically agrees with the monaural estimate (Sondhi and Guttman, 1966; Kohlrausch, 1988; Kollmeier and Holube, 1992; Holube *et al.*, 1998). Furthermore, data that were obtained as a function of the bandwidth of the *signal* also show a monaural bandwidth behavior (Langhans and Kohlrausch,

[a)]Now at: Philips Research Laboratories Eindhoven, Prof. Holstlaan 4, NL-5656 AA Eindhoven, The Netherlands. Electronic mail: jeroen.breebaart@philips.com

1992; Breebaart *et al.*, 1999). In this article, it is demonstrated that the current model, which includes peripheral filters with a bandwidth based on monaural estimates, can account for all observations described earlier. Depending on the experimental paradigm, the model shows different bandwidth dependencies. It is also explained that the wider effective critical bandwidth does not result from a poorer spectral resolution of the binaural auditory system compared to the monaural system, but is related to the ability of the model to *integrate* information across filters.

In a third article (Breebaart *et al.*, 2001b), temporal properties of the model are discussed, which include the effect of signal and masker duration, phase transitions in the time domain, and forward masking.

II. METHOD

A. Relevant stages of the model

In the Introduction, a coarse description of the general model setup was given. In this section, the stages of the model that are relevant for the simulations described in this article (i.e., spectral behavior) are discussed in more detail. For a detailed description of the complete model, see Breebaart *et al.* (2001a).

(i) Filtering of the gammatone filterbank. The filterbank present in the peripheral processing stage determines the spectral resolution of the model, in line with the ERB estimates published by Glasberg and Moore (1990).

(ii) Inner hair cell model. This stage consists of a half-wave rectifier followed by a fifth-order low-pass filter with a cutoff frequency (-3 dB) of 770 Hz. Hence below 1000 Hz, both the ITDs and IIDs are preserved at the output of this stage. However, above 2 kHz, the output approximates the envelope of the incoming signals and hence only IIDs and ITDs present in the envelope are preserved. Between 1 and 2 kHz, the ITD in the fine structure waveforms is gradually lost.

(iii) Adaptation loops. The chain of adaptation loops in the peripheral processor has an almost logarithmic input–output characteristic in steady state and is a nonlinear device. These properties have two consequences. First of all, it has been shown frequently that for both monaural and binaural detection of signals added to a wideband masker with a variable level, the threshold *signal-to-masker* ratio is approximately constant, as long as the masker level is well above the absolute threshold (cf. McFadden, 1968; Hall and Harvey, 1984). If it is assumed that a certain constant *change* at the output of the adaptation loops is needed to detect a signal, the signal must be equal to a certain *fraction* of the masker level due to the logarithmic compression. Hence the signal-to-masker ratio will be approximately constant at threshold. Second, due to the nonlinear behavior, large interaural intensity differences at the input cannot be canceled completely by a linear level adjustment at the output.

(iv) Compressive input–output characteristic of EI-type elements. The temporally smoothed difference signal of the EI-type elements is compressed logarithmically. In combination with an additive internal noise, this stage results in thresholds of interaural differences that depend on the interaural cross correlation of the reference stimuli.

(v) Weighting as a function of the internal delay. The model includes a weighting function that decreases with the internal delay of the EI-type elements. Consequently, the relative amount of internal noise increases with internal delay.

(vi) Optimal detector in the central processor. The EI-type element outputs are corrupted by an additive internal noise. Subsequently, the internal representations of the external stimuli are compared to a template that consists of the average masker-alone representation. The differences between the actual stimulus and the template are weighted and integrated both in the time and the frequency domain according to an optimal criterion. This enables the optimal detector to reduce the internal signal-to-noise ratio for stimuli that have a valuable detection cue in more than one auditory filter.

B. Procedure

Masked thresholds were simulated using an adaptive three-interval forced-choice (3IFC) procedure. The masker was presented in three consecutive intervals. One of the intervals contained the signal. The model's task was to indicate which interval contained the signal. The level of the signal was adjusted according to a two-down one-up algorithm (Levitt, 1971). The initial step size for adjusting the level was 8 dB. The step size was halved after every second track reversal until it reached 1 dB. The run was then continued for another eight reversals. The median level at these last eight reversals was used as the threshold value. At least five repetitions were performed for each parameter value. All thresholds are plotted as mean values of all repetitions, and the error bars denote the standard deviation of the repetitions.

C. Stimuli

All stimuli were generated digitally at a sampling frequency of 32 kHz. The maskers used in the different experiments had a duration of 400 ms unless stated differently. They were presented interaurally in-phase (No), interaurally out-of-phase ($N\pi$), or with a specific interaural correlation ($N\rho$) obtained by combining No and $N\pi$ noises (cf. Breebaart and Kohlrausch, 2001). The signals were presented interaurally in-phase (So), interaurally out-of-phase ($S\pi$), or to one ear only (Sm). Bandpass Gaussian noises were generated by computing the Fourier transform of white noise and setting the amplitude coefficients outside the desired frequency range to zero. After an inverse Fourier transform, the bandpass noise was obtained.

In all simulations, the level, bandwidth, on- and offset ramps, and onset delay of both the maskers and signals were set to the values used in the corresponding experiments with human subjects. If more data sets from various authors with different experimental settings were used, the experimental

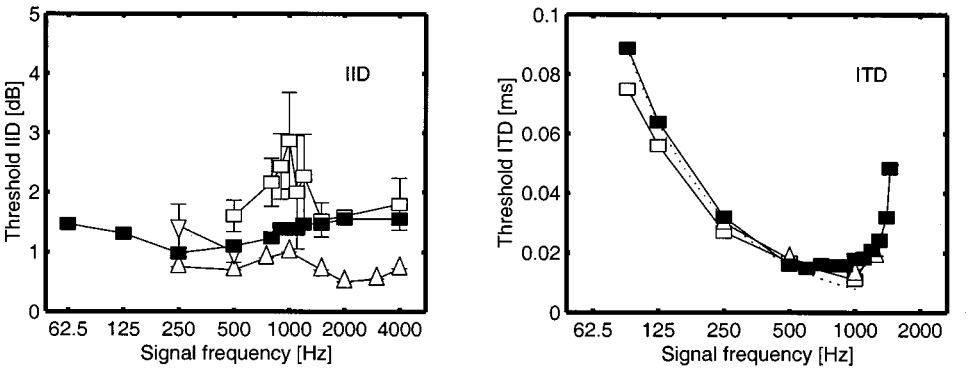

FIG. 1. IID thresholds (left panel) and ITD thresholds (right panel) for tones as a function of frequency. The open symbols denote data adapted from literature, the filled symbols are model predictions. The dotted line in the right panel denotes a constant interaural phase difference of 0.05 rad. Legend left panel: squares, Grantham (1984); upward triangles, Mills (1960); downward triangles, McFadden *et al.* (1971). Legend right panel: squares, Klumpp and Eady (1956); upward triangles, Zwislocki and Feldman (1956).

settings from one of these studies were used for determining the model simulations. Comparison with the other data sets was possible because in such conditions, the binaural masking level differences (BMLDs) were calculated or thresholds were expressed relative to the spectrum level of the masking noise.

D. Model calibration

As described in the preceding article (Breebaart *et al.*, 2001a), the model effectively calculates the difference signal between the left and right ears as a function of an internal delay and internal level adjustment. The subtraction is performed by so-called EI (excitation-inhibition) elements. The sensitivity to interaural differences of these EI-type elements is determined by two (fixed) model parameters *a* and *b*. By changing these parameters, the EI output is scaled relative to the internal noise which has a fixed level. These sensitivity parameters were determined as follows. Detection thresholds were simulated for a 500-Hz interaurally out-of-phase sinusoid in a Gaussian low-pass masker (cutoff frequency of 1 kHz) which had an interaural correlation of 0.64[1] and an overall level of 65 dB SPL. In this condition only the parameter *a* determines the detectability of the tone. This parameter was adjusted to reach a threshold of 46 dB SPL. Subsequently, the interaural correlation of the masker was set to $+1$ and the parameter *b* was adjusted in order to reach a threshold of 38 dB SPL. These thresholds were adapted from frozen-noise $N\rho S\pi$ data given by Breebaart and Kohlrausch (2001). The resulting values of *a* and *b* are 0.1 and 0.000 02, respectively. The monaural sensitivity of the model was adjusted such that the just noticeable difference in intensity of a 500-Hz, 400-ms sinusoid with a level of 65 dB SPL was 1 dB.

Note that during all simulations, all model parameters and procedures were kept constant. This restriction has the consequence that sometimes overall differences between model predictions and experimental data sets occur. However, similar differences exist between experimental data sets from different publications. For such conditions, much better predictions could have been obtained by calibrating the model's BMLD separately for each experiment. Nevertheless, all parameters were kept constant in order to demonstrate to what extent the current model can account for different experimental findings.

III. SIMULATIONS

A. Detection of static interaural differences

The first simulations comprised the detection of static interaural intensity and time differences. Hence in these experiments only pure tones were presented to the model in the absence of any noise masker. The pure tones had a duration of 400 ms and were gated with 50-ms Hanning ramps. The presentation level was 65 dB SPL. The reference stimuli were presented diotically and the target interval contained an IID in the first set of simulations and an ITD in the second set. The size of the ITD or IID was varied adaptively, similar to the procedure described in Sec. II A. For IID discrimination simulations, various frequencies between 62.5 and 4000 Hz were tested. For ITD discrimination simulations, frequencies with octave spacing were used below 500 Hz and a linear spacing of 100 Hz was used above 500 Hz. In the left panel of Fig. 1, the IID thresholds (filled symbols) of the model are presented as a function of the frequency of the tone together with experimental data (open symbols). The experimental data were adapted from literature: squares denote Grantham (1984), upward triangles Mills (1960), and downward triangles denote McFadden *et al.* (1971). The predicted IID thresholds do not depend systematically on the frequency and lie between 1 and 1.6 dB. This is well in the range of the experimental data. The remarkable bump at 1 kHz seen in one set of the experimental data is, however, lacking in the predictions.

The right panel of Fig. 1 shows ITD thresholds as a function of frequency. The open squares are data adapted from Klumpp and Eady (1956), the upward triangles from Zwislocki and Feldman (1956), and the filled symbols are model predictions. For frequencies up to 500 Hz, the ITD threshold decreases with increasing frequency. This ITD threshold curve can be characterized by a constant phase sensitivity, as shown by the dotted line. This line represents a constant phase difference of 0.05 rad. For frequencies above 1 kHz, the ITD threshold increases sharply due to the decrease of phase locking in the inner hair cell stage. Above 1.5 kHz, the model is not sensitive to static ITDs in the fine structure of the presented waveforms.

The predictions for IID thresholds and ITD thresholds below 1000 Hz only depend on the model parameters *a* and *b* which were derived from $N\rho S\pi$ and $NoS\pi$ detection experiments (see Sec. II D). In these experiments, both IIDs

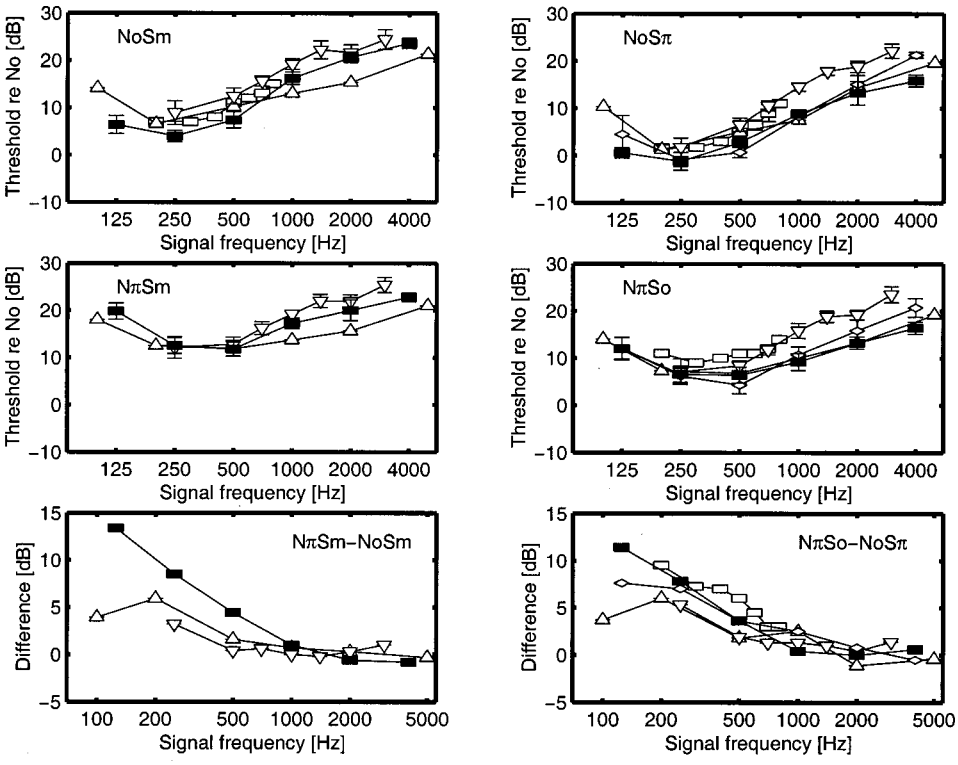

FIG. 2. Masked thresholds for wideband NoSm (upper left panel), NoSπ (upper right panel), NπSm (middle left panel), and NπSo (middle right panel) conditions as a function of the frequency of the signal. The lower left panel represents the difference in threshold between the NπSm and NoSm conditions, the lower right panel between NπSo and NoSπ. The open symbols denote experimental data adapted from literature, the filled symbols are model predictions: Squares, Kohlrausch (1988); upward triangles, Hirsh (1948); downward triangles, Hirsh and Burgeat (1958); diamonds, van de Par and Kohlrausch (1999).

and ITDs are present in the stimulus, which fluctuate as a function of time. In the simulations shown in Fig. 1, the stimuli contained only static IIDs or static ITDs. Thus, this model is able to unify IID and ITD sensitivity with binaural detection data. The use of a cancellation mechanism to achieve the correct sensitivity for both IIDs and ITDs was already suggested by Breebaart *et al.* (1999). They found that for their stimuli containing several different probability distributions of the IIDs or ITDs, a model based on subtraction may be favored over models based on the interaural cross correlation or models based on the direct evaluation of the interaural differences.

B. Dependence on frequency and interaural phase relationships in wideband detection conditions

A low-pass noise with a cutoff frequency of 8 kHz and a spectral level of 40 dB/Hz was used as masker. The following binaural conditions were tested: NoSπ, NπSo, NoSm, and NπSm. Thresholds were determined as a function of the frequency of the signal (125, 250, 500, 1000, 2000, and 4000 Hz). The upper and middle panels in Fig. 2 represent the results of the four different conditions that were tested. The open symbols are experimental data extracted from different studies (see figure caption for a description), and the filled symbols are the model predictions. The thresholds are plotted relative to the spectral level of the masker to compensate for differences in masker spectral level.

For the NoSm and the NoSπ conditions (upper panels), the thresholds decrease slightly with frequency between 125 and 500 Hz and they increase towards higher frequencies with a slope of 4 to 5 dB/oct. In the model, this slope partly results from the increased bandwidth of the auditory filters towards high frequencies because thresholds are expressed relative to the masker spectral level. This fact predicts a threshold increase by about 3 dB/oct.[2]

A second reason why thresholds increase above 1 kHz center frequency is related to the loss of phase locking in the inner hair-cell model. In the stimuli that are considered here, both interaural intensity and time differences are present (cf. Zurek, 1991). As described in Sec. II A, the model is insensitive to interaural time differences within the fine structure of the waveforms for frequencies above 1.4 kHz, because of the loss of phase locking in the inner haircell stage. Hence a part of the cues that are available at low frequencies are lost at high frequencies resulting in higher thresholds.

A third reason for a threshold increase with frequency results from peripheral compression. Above 1.4 kHz, only the envelope of the incoming waveforms is present at the output of the inner haircell stage. These envelopes are compressed by the nonlinear adaptation loops that follow the inner haircell stage. Such a compression results in a decrease of the stimulus IID (van de Par, 1998; van de Par and Kohlrausch, 1998). Furthermore, compression makes the model less sensitive to interaural time differences present in the envelopes, because the envelopes are flattened. Hence compression results in higher thresholds at frequencies above 1.4 kHz. Low-frequency detection is not affected much by compression because the model can use ITDs in the fine structure waveforms which are not affected by compression.

The major difference between the NoSπ and NoSm conditions is an overall difference of 6 dB. This difference can be understood by considering the fact that an Sπ signal results in twice as much increase in the EI-type element output compared to an Sm signal. To compensate for this difference, the Sm signal level must be raised by 6 dB.

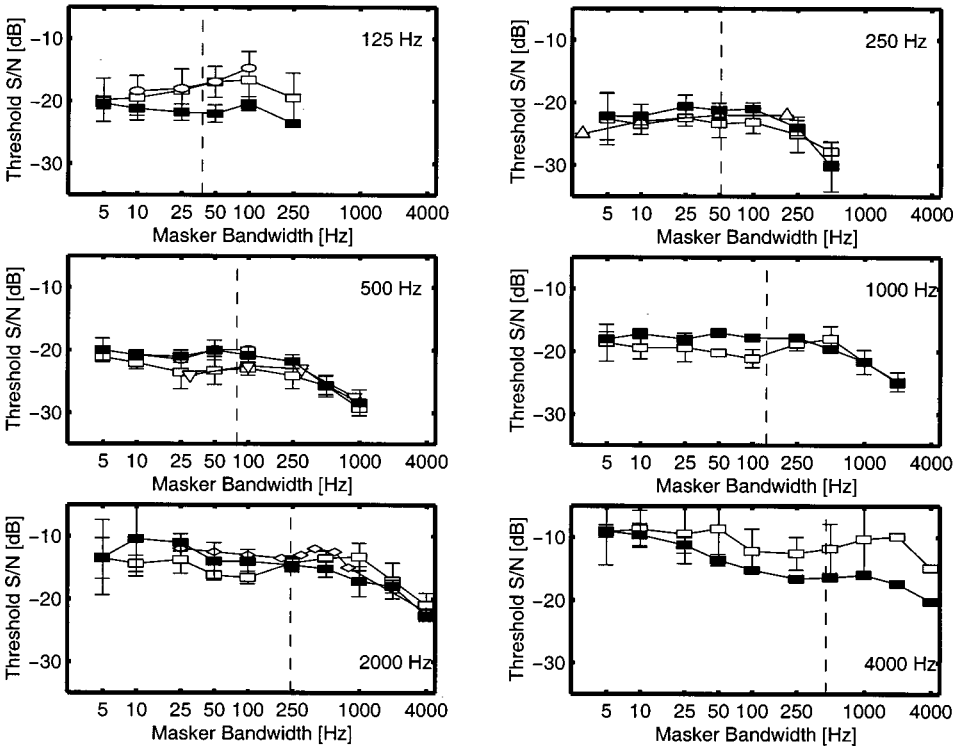

FIG. 3. NoSπ thresholds as a function of the masker bandwidth for a constant overall level of the masker. The six panels represent center frequencies of 125, 250, 500, 1000, 2000, and 4000 Hz, respectively. The dashed line indicates the ERB value at the signal frequency. The filled symbols are model predictions. The open symbols are data adapted from literature: squares, van de Par and Kohlrausch (1999); upward triangles, Wightman (1971); downward triangles, van der Heijden and Trahiotis (1998); diamonds, Sever and Small (1979); circles, Breebaart *et al.* (1998).

The NπSo and NπSm conditions show similar thresholds as the NoSπ and NoSm conditions for frequencies beyond 1 kHz, while for frequencies below 1 kHz, the interaurally out-of-phase maskers result in higher thresholds than the in-phase maskers. This difference is depicted in the lower panels of Fig. 2. The left panel shows the difference in thresholds between the NπSm and NoSm conditions, the right panel between NπSo and NoSπ. The model predictions in these two panels are very similar for the monaural and dichotic signal, having differences of 12 to 14 dB at 125 Hz center frequency which decreases to 0 dB around 1 or 2 kHz. These frequency effects are the result of two model properties which are included in most EC-like models (cf. Durlach, 1963; Rabiner *et al.*, 1966; Metz *et al.*, 1968) and in models based on coincidence detectors (Colburn, 1977; Stern and Shear, 1996). The first comprises a limited repertoire of internal delays. A lower center frequency corresponds to a larger internal delay necessary to compensate for the phase shift of the masking noise and hence less sensitivity to changes in the EI-type element output. The second is the fact that the phase shift can only be compensated by an internal delay, resulting in imperfect cancellation of the noise masker due to damping of the autocorrelation function of the noise.

C. NoSπ masker-bandwidth dependence

If an interaurally out-of-phase signal is masked by an interaurally in-phase noise of variable bandwidth, a remarkable phenomenon is observed which is usually referred to as the wider effective binaural critical bandwidth: the critical bandwidth estimate from binaural band-widening experiments is often a factor 2 to 3 higher than monaural estimates (cf. Bourbon and Jeffress, 1965; Sever and Small, 1979; Hall *et al.*, 1983; Zurek and Durlach, 1987; van de Par and Kohlrausch, 1999). In order to show that the model can account for this phenomenon, NoSπ thresholds were determined as a function of the bandwidth of the masker at center frequencies of 250, 500, 1000, 2000, and 4000 Hz. The bandwidth was varied between 5 Hz and twice the center frequency. The overall masker level was kept constant at 65 dB SPL. Both model predictions (filled symbols) and experimental data (open symbols) are shown in Fig. 3, where the 6 panels correspond to center frequencies of 125, 250, 500, 1000, 2000, and 4000 Hz, respectively. The data can be characterized as staying fairly constant up to a certain bandwidth and then declining with 3 dB/oct. The bandwidth at which this decline starts is much larger than what is expected from monaural notched-noise experiments (Glasberg and Moore, 1990), as shown by the dashed lines in Fig. 3. The estimated critical bandwidths of the model are also larger by a factor up to 2 or 3, in line with the experimental data.

At 4 kHz, the simulated thresholds decrease as a function of bandwidth for bandwidths between 10 and 250 Hz, an effect which is not observed in the experimental data. The threshold decrease with bandwidth at 4 kHz of the model can be understood by considering the properties of the adaptation loops in the peripheral processor. Due to the loss of phase locking at high frequencies, the outputs of the inner haircell model only contain IIDs and some ITD information present in the envelope of the waveforms. The IIDs are reduced in magnitude due to the compressive nature of the adaptation loops. The amount of compression depends on the bandwidth of the stimulus. If the bandwidth is small (i.e., 10 Hz), the envelope of the masker varies slowly with time. Hence the adaptation loops which adapt with certain time constants easily follow these envelope fluctuations. Consequently, the envelopes are compressed logarithmically. If the masker bandwidth is increased, the speed of fluctuation in the envelopes

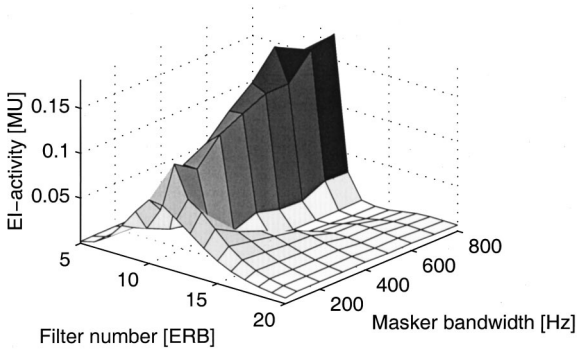

FIG. 4. EI activity without internal noise as a function of masker bandwidth and peripheral filter number for an NoSπ condition with a signal-to-masker ratio of −25 dB and a fixed masker level of 65 dB.

increases accordingly. As described in the preceding article (Breebaart *et al.*, 2001a), fast fluctuations are processed more linearly, while slow fluctuations in the envelope are processed logarithmically. Hence the magnitude of the IID after the adaptation loops depends on the bandwidth of the stimulus: a very large bandwidth results in larger IID after adaptation and hence lower NoSπ thresholds.[3]

At all center frequencies the increased effective bandwidth is captured by the model. The explanation for this result in this specific experimental paradigm relies on the across-frequency integration of information according to an optimal detector, as described by van de Par and Kohlrausch (1999). This spectral integration is an integral part of the central processor in this model (cf. Breebaart *et al.*, 2001a, Sec. VI). For a narrow-band masker (i.e., below the monaural critical bandwidth), the on-frequency filter has the largest stimulus power. For off-frequency channels, the entire stimulus resides at the skirts of the filters and is therefore reduced in its power. However, the relative amount of masker and signal energy is hardly changed. Since the signal-to-masker ratio within an auditory channel determines the detectability, information about the presence of the signal is not only available in the on-frequency channel, but also in several off-frequency channels. The only limitation for this extended availability is the absolute threshold: if the stimulus is attenuated too much it becomes undetectable in that channel. This extended availability of information is depicted in Fig. 4. In this figure, the output for a 500-Hz NoSπ stimulus of an EI-type element with $\alpha=\tau=0$ without internal noise is shown as a function of the masker bandwidth and of the auditory-filter number. The masker had a fixed level of 65 dB while the signal-to-masker ratio was −25 dB.

If a masker alone were presented, the output of these EI-type elements would be zero (neglecting the internal noise) since the masker can be canceled completely. Thus, any increase in the activity can be used as a cue for the presence of the signal. This increase by the presence of the Sπ signal is shown in Fig. 4. If the masker bandwidth is very small (10 Hz wide), a whole range of EI-type elements shows a considerable amount of activity. The fact that in this condition the cue for detection is available in several channels enables reduction of the internal error in the following way. The internal error which is added to the EI-type elements is *independent* across elements. Thus, individual noise sources add up by their intensities. On the other hand, the increase in the EI-type elements by the addition of the signal is available in several filters and is *correlated* and therefore adds up linearly. Thus, if the model uses the sum of activities across several elements instead of using the output of only one element, the internal signal-to-noise ratio is increased. This results in lower thresholds for narrow-band maskers.

For bandwidths that just exceed the critical bandwidth, the masker energy in the on-frequency channel starts to be reduced by the bandpass filter. This results in an increase in the signal-to-masker ratio in the on-frequency channel. This can be observed from Fig. 4 by the increasing EI activity with increasing bandwidth for filter 10. However, the signal-to-masker ratio in the *off-frequency* channels starts to be *reduced*. This is clearly visible for filters 13 to 20; the activity *decreases* with increasing masker bandwidth. Therefore, the ability of the model to reduce the internal error by integrating across filters is diminished. Both processes influence the internal error about equally but in opposite directions, resulting in constant thresholds for bandwidths between the critical bandwidth and two to three times that value. For even larger bandwidths, all off-frequency channels are masked. Only the on-frequency channel provides useful information and due to filtering thresholds decrease with 3 dB/oct of masker bandwidth. In summary, the wider critical bandwidth in the model is the result of an uncorrelated noise source in each auditory filter combined with an optimal detector. Other implementation issues are relatively unimportant, a notion which is supported by the results of Zerbs (2000). He developed a binaural signal detection model which is also based on an EC-like process but with a different implementation of the binaural processing stage. This model does account in a very similar way for the wider effective critical bandwidth.

Another set of data that can be explained by this across-channel integration hypothesis has been published by Hall *et al.* (1983), who also performed band-limiting NoSπ measurements. In their study, the spectral *energy* density of the masker was kept constant and they estimated the binaural critical bandwidth at three noise levels, namely, 10, 30, and 50 dB/Hz. Their critical-bandwidth estimate increased as the noise level increased. To test whether the model can account for this observation, the same experiment was simulated with masker bandwidths of 10, 50, 100, 200, 400, 600, and 800 Hz and a center frequency of 500 Hz. Both model predictions and data adapted from Hall *et al.* (1983) are shown in the left panel of Fig. 5.

The squares denote a spectral energy density of 10 dB/Hz, the upward triangles of 30 dB/Hz, and the downward triangles 50 dB/Hz. At bandwidths below the monaural critical bandwidth, the thresholds increase with increasing bandwidth. This is the result of the increasing amount of masker energy within the auditory filter. At a certain wider bandwidth the thresholds remain constant. Hall *et al.* used the bandwidth that corresponds to a threshold 3 dB below this constant threshold level as an estimate of the critical bandwidth. For the 10 dB/Hz condition, the estimate was 58 Hz, very close to the monaural estimate of 79 Hz at 500 Hz (Glasberg and Moore, 1990). At 50 dB/Hz, however, the estimate was 220 Hz, which was close to three times the mon-

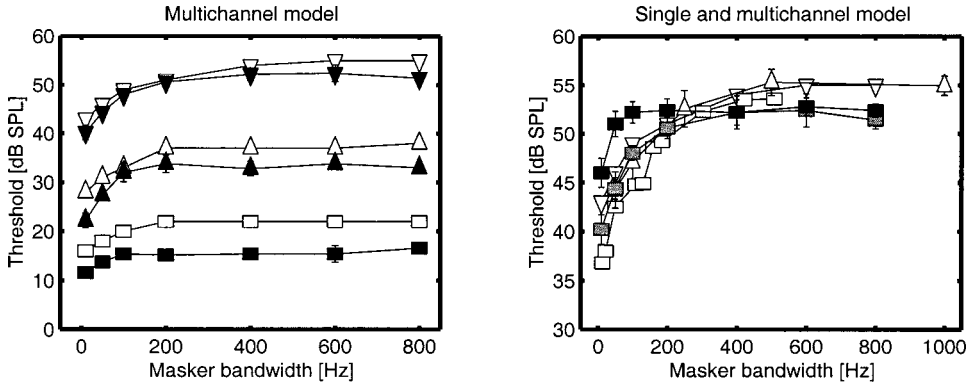

FIG. 5. NoSπ thresholds as a function of masker bandwidth. Left panel: predictions of a multichannel model. In all conditions, the spectral energy density of the masker was kept constant at 10 dB/Hz (squares), 30 dB/Hz (upward triangles), and 50 dB/Hz (downward triangles). The open symbols are data adapted from Hall *et al.* (1983), the filled symbols are model predictions. Right panel: similar to the left panel for a masker level of 50 dB/Hz. The squares denote data adapted from Bourbon (1966), the upward triangles denote Cokely and Hall (1991), and the downward triangles Hall *et al.* (1983). The gray symbols represent the multichannel model, the black symbols the single-channel model.

aural estimate. Our explanation for this level dependence of the critical band estimate also relies on across-frequency integration. Consider the experiment with a narrow-band masker with a spectral level of 10 dB/Hz. In this case, the excitation pattern across auditory channels is very narrow due to the low stimulus level. Off-frequency channels do not provide useful information since the stimulus level in these channels is below absolute threshold. Therefore, the bandwidth dependence of thresholds will depend only on processing of the on-frequency channel and consequently reflects the critical bandwidth of this on-frequency channel. For high stimulus levels (i.e., 50 dB/Hz) the same argument for the wider critical bandwidth can be given as described earlier.

The differences in predictions that are obtained between a model that uses across-frequency integration and a model with only single-channel processing are shown in the right panel of Fig. 5. Here, experimental data for a spectral level of 50 dB/Hz are shown (open symbols) combined with model simulations for a single-channel model (on-frequency channel only, filled symbols) and for the multi-channel model (gray symbols). Clearly, the bandwidth dependence of the single-channel model resembles monaural behavior instead of the wider binaural bandwidth. Furthermore, the multi-channel model has lower thresholds for bandwidths up to 400 Hz. The difference is as large as 6 dB. Thus, below a bandwidth of 400 Hz, the model can improve its detection performance by integrating information across filters. At larger bandwidths, all off-frequency channels are masked and the performance for both the single- and multi-channel models is equal. This demonstrates that in the model, the wider critical bandwidth results from across-frequency integration.

A general observation regarding the experiments from Hall *et al.* (1983) is that the model predictions are up to 5 dB lower than the experimental data. One reason for this difference may be the fact that the model used a two-down, one-up procedure to vary the signal level, while in the original experiment, a three-down, one-up procedure was used. Hence the thresholds for the experimental data are somewhat higher than for the model predictions.

In summary, critical aspects of the model necessary to explain the wider effective critical bandwidth are the uncorrelated internal noise in each auditory filter combined with an optimal integration of information across frequency.

D. NπSo masker-bandwidth dependence

In this section, interaurally out-of-phase maskers of variable bandwidth combined with a diotic signal (i.e., NπSo) are discussed. Thresholds for different center frequencies (125, 250, 500, and 1000 Hz) and bandwidths (from 5 Hz up to twice the center frequency) were simulated. The overall masker level was kept constant at 65 dB SPL. The experimental data for the four different center frequencies (adapted from van de Par and Kohlrausch, 1999) and model simulations are shown in Fig. 6.

If the NπSo thresholds are compared to the NoSπ thresholds shown in Fig. 3, one can see that at 125 and 250 Hz, the slope relating bandwidth to threshold at subcritical bandwidths is significantly different, while at higher center frequencies, no difference is observed. This observation holds for both model predictions and experimental data. In our model, the differences between the two experimental conditions are the result of the same model properties that were mentioned in Sec. III B: the fact that the interaural phase shift in the masker can only be compensated by an internal delay. Thus both a *lower* center frequency and a *wider* masker bandwidth result in more masker energy that cannot be canceled and a relative increase in the internal noise. In addition, the limited range of internal delays becomes more important for lower frequencies. This is supported by the data: the slope relating threshold to bandwidth is steeper at 125 Hz center frequency than at higher frequencies. Moreover, this slope is practically zero for the 500- and 1000-Hz conditions. At these frequencies, the effect of damping of the autocorrelation function of the noise is so small that the thresholds are not influenced by this effect. Hence the thresholds approach the NoSπ thresholds (see Fig. 3). Furthermore, overall differences in thresholds occur across center frequencies, which are most clearly visible at 125 and 250 Hz.

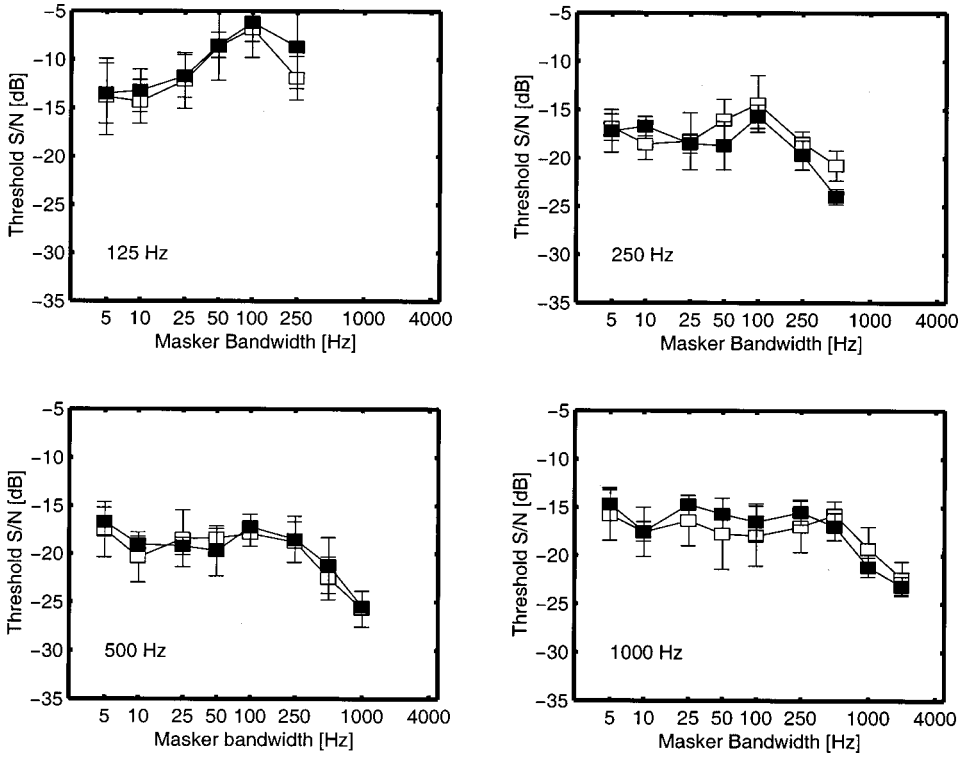

FIG. 6. NπSo thresholds as a function of masker bandwidth for 125-Hz (upper-left panel), 250-Hz (upper-right panel), 500-Hz (lower-left panel), and 1000-Hz center frequency (lower-right panel). The open symbols are data adapted from van de Par and Kohlrausch (1999), the filled symbols are model predictions. The masker had the same overall level at all bandwidths.

E. NρSπ masker-bandwidth dependence

NρSπ detection thresholds were measured as a function of the bandwidth of the masker by van der Heijden and Trahiotis (1998). The overall masker level was kept constant. They used several values of interaural correlation (ρ) ranging from -1 to $+1$. Their results show that the wider effective critical bandwidth is only observed for interaural masker correlations >0.97 and that for smaller correlations, the effective critical bandwidth is similar to the monaural estimate. Their results and the model simulations are shown in Fig. 7. The left panel shows the experimental data, the right panel shows model predictions.

The empirical curves for the largest interaural correlations (i.e., 1 and 0.997) show flat thresholds for masker bandwidths between 30 and 300 Hz, indicating a wider critical bandwidth. In these conditions, the *internal* noise limits the detection and an internal error reduction scheme is applied as described in the NoSπ band-widening condition (see Sec. III C). For an interaural masker correlation of -1, however, the thresholds are similar to the monaural thresholds since no binaural advantage is present in an NπSπ condition. These monaural thresholds show a bandwidth dependence well in line with the view that the decision variable that is used in the detection process has the same statistics as the energy at the output of an auditory filter. For bandwidths below the critical bandwidth, thresholds decrease with 1.5 dB per octave of masker bandwidth. This slope results from the sample-by-sample variability in the masker energy (Bos and de Boer, 1966). Thus, in these conditions, the *external stimulus variability* limits the detection of the tone instead of the internal noise. Since this external variability (or masker energy fluctuation) is highly correlated across auditory channels, the model *cannot reduce this variability* by combining information across frequency. Breebaart and Kohlrausch

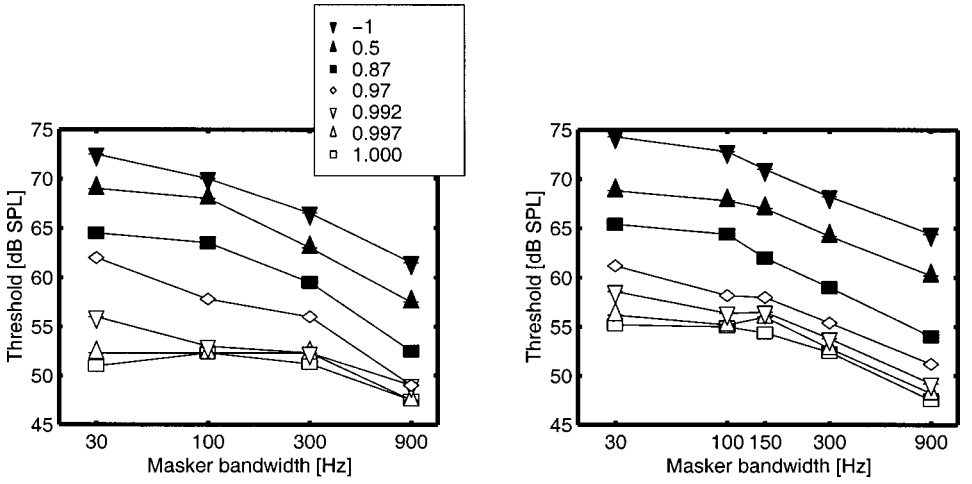

FIG. 7. NρSπ thresholds as a function of masker bandwidth for a constant overall level of the masker. The left panel shows experimental data adapted from van der Heijden and Trahiotis (1998); the right panel shows model predictions. Different curves denote different interaural masker correlations.

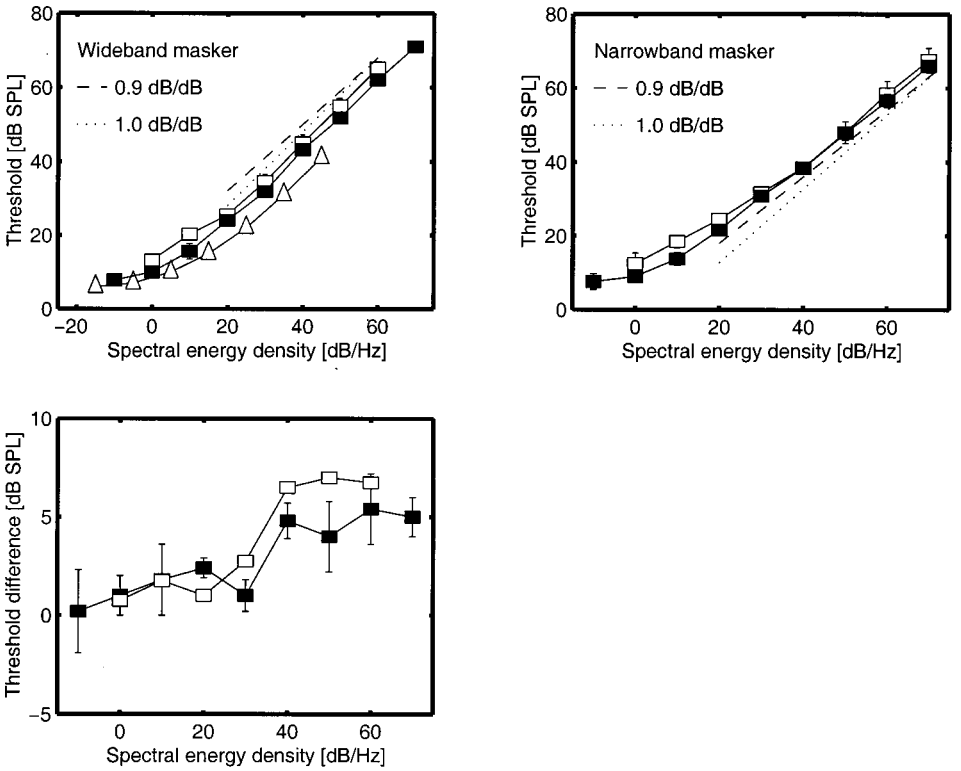

FIG. 8. NoSπ thresholds as a function of the spectral energy density of the masker for broadband noise (upper-left panel) and narrowband noise (upper-right panel). Open symbols are data adapted from literature: squares from Hall and Harvey (1984), upward triangles from McFadden (1968). The filled symbols are model predictions. The dashed lines have a slope of 0.9 dB/dB, the dotted lines have a slope of 1 dB/dB. The lower panel contains the difference in thresholds between broadband and narrow-band conditions in the same format as the upper panels.

(1999) suggested that a similar argument holds for NρSπ conditions. They measured NρSπ thresholds as a function of both the masker correlation and the masker bandwidth. They showed that variability of the interaural differences has a strong effect on the binaural performance and that reduction of this variability results in a decrease in thresholds. The current model supports this hypothesis in a qualitative way. For interaural correlations below 0.97, the fluctuations in the EI-element output are dominated by external stimulus fluctuations. The amount of fluctuation increases with masker correlation reduction or masker bandwidth reduction. This is also seen in the model predictions: the data decrease with both a bandwidth increase or a correlation increase. This experiment shows that our model can account both for external stimulus fluctuations and for internal errors as limitations for detection.

F. NoSπ masker-level dependence

If the across-frequency hypothesis as stated earlier is correct, different influences of masker level on NoSπ are expected for a narrow-band masker compared to a broadband masker. For a broadband masker, off-frequency channels cannot contribute to the detection of the signal because these channels are masked by the broadband noise. On the other hand, off-frequency channels can contribute in the case of a narrow-band masker as long as the stimulus level in these channels is above the absolute threshold. The number of off-frequency channels that can be used depends on the stimulus level: at a higher level the excitation pattern along the frequency axis is larger and hence more auditory filters can contribute to the reduction of the internal error. Thus, the increase in the NoSπ threshold with masker level should be shallower for a narrow-band masker than for a broadband masker.

For the broadband condition, a broadband (0–8000 Hz) Gaussian noise served as masker. Its spectral energy density varied between −15 and 70 dB/Hz. The signal was a 500-Hz interaurally out-of-phase tone. The upper-left panel of Fig. 8 shows the experimental data (open symbols) and the model predictions (filled symbols) as a function of the noise level.

For spectral levels below 0 dB/Hz, the thresholds are approximately constant. For these conditions, the masker energy within an auditory filter is too low to influence the detectability of the signal: the threshold is determined by the absolute hearing threshold. If the spectral level of the masker is increased, the amount of masker energy within the on-frequency auditory filter increases also, resulting in higher thresholds. The slope relating masker level and threshold equals 1 dB/dB (indicated by the dotted line) for masker spectral levels above about 20 dB/Hz.

The upper-right panel of Fig. 8 shows NoSπ thresholds for a narrow-band masker (50 Hz wide) spectrally centered around the signal. The data are very similar to the data for the wideband condition except for the fact that the slope relating signal threshold to spectral noise level is lower. This is indicated by the dashed line, which has a slope of 0.9 dB/dB. As expected, the availability of off-frequency channels in the narrow-band condition results in a shallower slope compared to the broadband condition where off-frequency channels do not provide useful information about the presence of the signal. The difference in thresholds between broadband and narrow-band conditions is depicted in the lower panel of Fig. 8, for both the experimental data (open symbols) and model predictions (filled symbols).

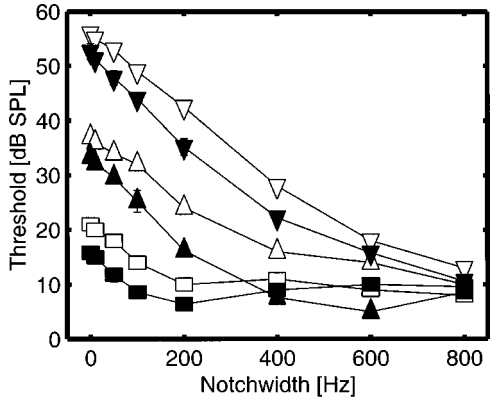

FIG. 9. NoSπ thresholds as a function of masker notchwidth. For each set of connected points, the spectral energy density of the masker was kept constant at 10 dB/Hz (squares), 30 dB/Hz (upward triangles), and 50 dB/Hz (downward triangles). The open symbols are data adapted from Hall *et al.* (1983), the filled symbols are model predictions.

G. NoSπ notchwidth dependence

In the context of across-frequency integration it is interesting to consider additional conditions where the use of off-frequency channels is disabled. Such a paradigm was presented by Hall *et al.* (1983). Instead of a band-limited masker, they used a notched noise and varied the notchwidth. The out-of-phase signal was spectrally centered in the gap. This experiment was performed at three spectral noise levels, 10, 30, and 50 dB/Hz. The data (open symbols) and the model predictions (filled symbols) are shown in Fig. 9. The data show a decrease with increasing notchwidth until the absolute threshold is reached. Hall *et al.* (1983) used the notchwidth corresponding to a 3-dB threshold improvement compared to thresholds at a notchwidth of 0 Hz as an estimate of the critical bandwidth. These estimates for the three masker levels were close to the monaural estimate and did not depend on the masker level. This finding is also supported by the model predictions. The slope relating threshold to notchwidth is very similar and hence the 3-dB estimates of the critical bandwidth are similar, too. Thus, as expected, the wider effective critical bandwidth is not observed for these stimuli. An overall difference of up to 5 dB between the model predictions and the experimental data adapted from Hall *et al.* (1983) is observed in Fig. 9. This difference was already discussed in Sec. III C, where for a different data set from the same study the same systematic difference was observed.

H. Maskers with phase transitions in the spectral domain

Another experimental paradigm to measure the spectral resolution of the binaural auditory system was used by Kohlrausch (1988). He measured the detectability of an interaurally out-of-phase signal in a masker which had a frequency-dependent phase difference: for frequencies below 500 Hz, the masker was in phase, while above 500 Hz the masker was interaurally out-of-phase. This condition is denoted NoπSπ. The spectral energy density of the masker was 43 dB/Hz. When the signal frequency was sufficiently below 500 Hz, the effective stimulus configuration was NoSπ and a large BMLD was observed. For frequencies well above 500 Hz, the stimulus corresponded to NπSπ, and no BMLD could be measured. For frequencies near 500 Hz, a gradual increase in thresholds was observed, indicating a limited spectral resolution of the auditory system. The data combined with model predictions are shown in the left panel of Fig. 10. The right panel shows thresholds for an NπoSπ condition, where the masker is interaurally out-of-phase below 500 Hz and in-phase above 500 Hz.

The gradual change in the thresholds near 500 Hz results from the limited spectral resolution of the filterbank. For example, if the signal has a frequency of 200 Hz in the NoπSπ condition, the channel tuned to 200 Hz effectively contains an NoSπ condition. The cue for detection is most salient for an EI-type element with $\alpha=\tau=0$. If the signal frequency is increased, an increasing amount of the antiphasic masker energy is present in a channel tuned to the signal frequency. This part of masker energy cannot be canceled by the EI-type element. Hence the input level of the EI-type element for a masker alone increases with increasing signal frequency. Therefore, a gradual increase of the thresholds is observed near the frequency of the masker phase transition. This gradual increase reflects the monaural critical bandwidth because spectral integration cannot occur.

A modified version of such a measurement of the spectral resolution of the auditory system uses a masker that has an inverted interaural phase relationship within a passband of the masker. Basically four conditions can be used in this way which are referred to as NoπoSπ, NπoπSπ, NoπoSo, and NπoπSo. The S denotes the interaural phase of the signal, the N denotes the frequency-dependent interaural phase relationship of the masker. Thus, an Noπo masker is interaurally in-phase except for a certain inner passband which is interaurally out-of-phase. This passband is centered around the

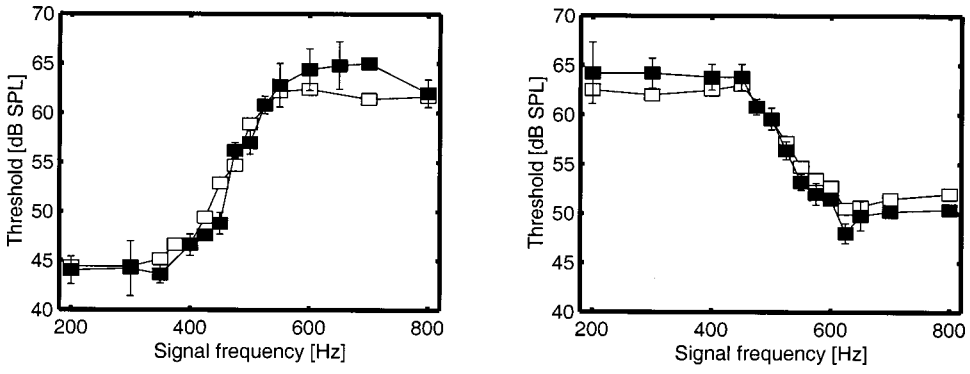

FIG. 10. Frequency dependence of Sπ thresholds for a masker with a frequency-dependent interaural phase difference. The left panel shows thresholds for an NoπSπ stimulus configuration (masker interaurally in-phase below 500 Hz and out-of-phase above 500 Hz), the right panel for NπoSπ (masker interaurally out-of-phase below 500 Hz and in-phase above 500 Hz). The open symbols are data adapted from Kohlrausch (1988), the filled symbols are model predictions.

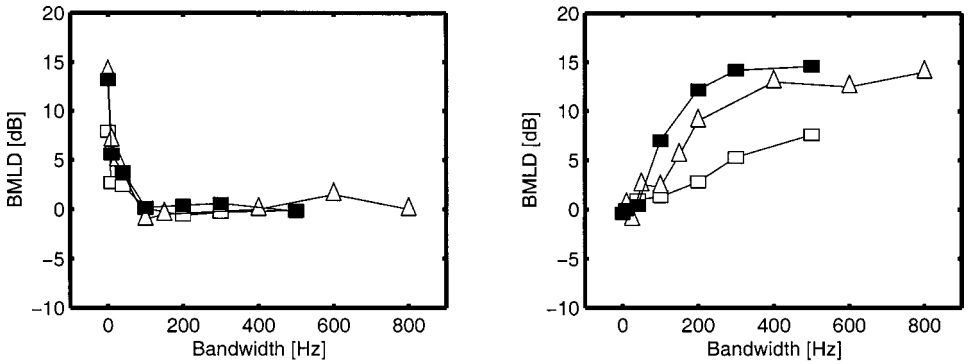

FIG. 11. NoπoSπ MLDs (left panel) and NπoπSπ MLDs (right panel) as a function of the bandwidth of the central band (see text for details). The open squares are experimental data adapted from Sondhi and Guttman (1966), the triangles are adapted from Holube *et al.* (1998). Filled symbols are model predictions.

signal. Thus, in case of NoπoSπ, the condition is effectively NπSπ if the passband is wider than the critical band, while it is NoSπ if the passband has a bandwidth of 0 Hz. By studying the bandwidth dependence of such conditions, the frequency resolution of the binaural auditory system can be estimated. Such an experiment was performed by Sondhi and Guttman (1966) and also by Holube *et al.* (1998). Their most striking result was that NoπoSπ and NπoπSπ thresholds reveal a completely different bandwidth dependence. This can be observed in Fig. 11. The left panel shows the BMLD as a function of the bandwidth of the inner band for the NoπoSπ condition, the right panel for NπoπSπ. The filled symbols are the model predictions, the open symbols are experimental data.

The NoπoSπ condition (left panel) has a BMLD which is large for a very small bandwidth but decreases quickly with bandwidth. This strong decrease in the BMLD is expected for the following reason. If the bandwidth of the out-of-phase passband is equal to half the equivalent rectangular bandwidth of the auditory filter, the amount of masker energy from the in-phase noise and the out-of-phase noise is approximately equal in the on-frequency filter.[4] Consequently, the effective interaural correlation of the masker in the on-frequency filter is about zero. For such a low correlation value, the BMLD is reduced to only 3 dB. Thus, if the bandwidth of the passband is 40 Hz, the BMLD should be significantly reduced. This is supported by the data in Fig. 11: for this bandwidth the BMLD is only a few dB.

In the NπoπSπ condition (right panel of Fig. 11), the bandwidth dependence of the thresholds is completely different. If the bandwidth of the passband is very large (i.e., 500 Hz), the condition is effectively NoSπ and a large BMLD is observed. If the bandwidth is decreased, the amount of masker energy in the on-frequency filter that is interaurally out-of-phase increases. Since these parts of the masker reside at the skirt of the filter, this masker energy is strongly attenuated. Consequently, the interaural masker correlation in the on-frequency channel decreases, but not very much. Only if the bandwidth is equal to about half the equivalent rectangular bandwidth,[4] the BMLD is reduced to about 3 dB for the same reason as in the NoπoSπ condition. Therefore, a much more gradual decrease in the BMLD is observed if the bandwidth of the passband is decreased. This is supported by both model predictions and experimental data. This interpretation is also supported by Fig. 12, which shows the BMLDs of the model as a function of the computed interaural cross correlation of the masker after peripheral filtering. The squares denote the NoπoSπ condition, the triangles the NπoπSπ condition. For both conditions, the BMLD for an interaural correlation of zero is very close to 3 dB. Furthermore, the BMLD as a function of the correlation is very similar for both conditions.

A substantial difference between the two data sets is observed in the maximum BMLD: the data of Sondhi and Guttman (1966) have a maximum BMLD of about 7 dB, while the data of Holube *et al.* (1998) show BMLDs of up to 14 dB. The reason for these differences is unclear, but the model could accommodate these differences by changing the parameters *a* and *b*.

I. NoSπ signal-bandwidth dependence

Langhans and Kohlrausch (1992) measured NoSπ thresholds for target signals of variable bandwidth. In this experiment, the masker consisted of a band-limited diotic noise (0–2 kHz, No=47 dB/Hz), while the signal consisted of harmonic complexes with a flat amplitude spectrum, a spectral spacing between the components of 10 Hz, and a center frequency of 400 Hz. The upper and lower frequency boundaries of the harmonics were varied in order to generate signals of different bandwidths. The total number of components in the complex was 1, 3, 5, 7, 9, 11, 15, 17, 19, or 41. The masked thresholds of the harmonic complex tones as a function of the number of components is shown in Fig. 13. The thresholds are expressed as level per component.

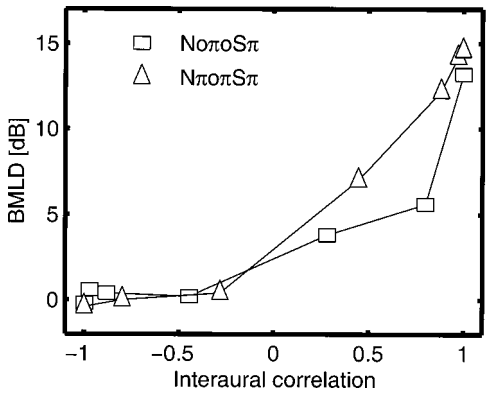

FIG. 12. BMLDs of the model as a function of the (computed) interaural correlation after peripheral filtering. The squares correspond to the NoπoSπ condition, the triangles to the NπoπSπ condition.

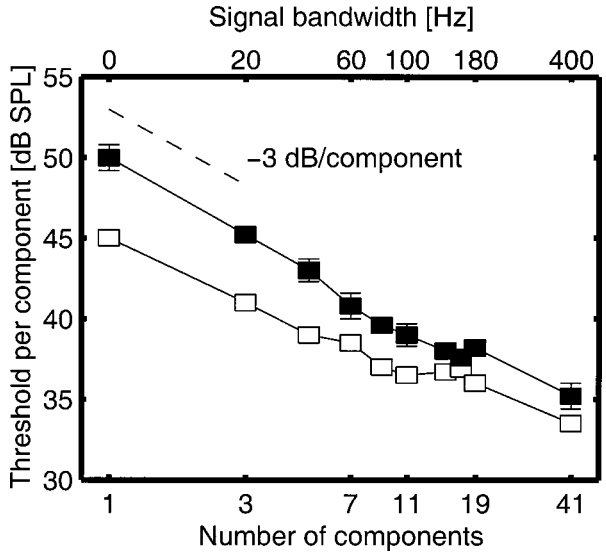

FIG. 13. Threshold level per component for an out-of-phase harmonic complex in an in-phase noise masker as a function of the number of signal components. The signal components had a spectral spacing of 10 Hz and were centered around 400 Hz. The open symbols are data adapted from Langhans and Kohlrausch (1992), the filled symbols are model predictions.

Figure 13 shows a decrease of the masked threshold with increasing number of components. This can be understood as follows. If the number of components is increased from one to three, the total signal level is 4.7 dB higher than the level of the individual components. The bandwidth of a signal which consists of three components equals 20 Hz, hence the signal has a smaller bandwidth than the auditory filter. Therefore, a threshold decrease of 4.7 dB per component is expected between one and three components in order to keep the signal power in the auditory filter constant. This is supported by the data and the model predictions. As long as the signal bandwidth is below the critical bandwidth, it is expected that an increase in the number of harmonics results in a decrease in the masked threshold level per component due to the increase in the total signal level. If the auditory filters had a rectangular shape (ideal bandpass filter), a doubling of the number of harmonics would result in a decrease of 3 dB in the level per component as long as the signal bandwidth is below the auditory filter bandwidth (nine harmonics). This is indicated by the dashed line. Of course, a rectangular filter is not a valid description of the auditory filters. Therefore, the decrease in threshold is a little bit less than 3 dB [see Langhans and Kohlrausch (1992) for a detailed analysis of the slope in their data]. A striking result is, however, that the thresholds still decrease for more than ten harmonics. In this case, the bandwidth of the signal exceeds the auditory filter bandwidth. To account for these results, it is necessary to include several filters in the detection process rather than only the center channel since the cue for detection is available in several filters. As described in Sec. III C, the availability of the cue for detection in several filters enables the improvement of the internal signal-to-noise ratio which results in lower signal thresholds for broadband signals. As can be observed from Fig. 13, the model accounts for this across-channel processing of binaural cues. However, the efficiency of this process in the model seems to be a little bit too high because the model predictions decrease stronger with increases in signal bandwidth than the experimental data.

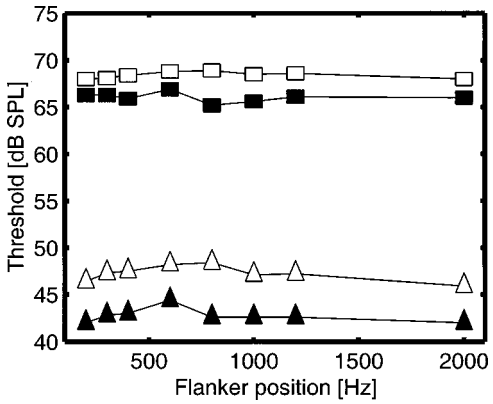

FIG. 14. Detection thresholds for a 500-Hz signal added to a 50-Hz-wide noise masker with a center frequency of 500 Hz as a function of the center frequency of a 30-Hz-wide flanking band. Squares correspond to thresholds for an interaurally in-phase signal, triangles to an out-of-phase signal. Open symbols are experimental data adapted from Cokely and Hall (1991), filled symbols are model predictions.

J. NoSπ including spectral flanking bands

Cokely and Hall (1991) measured narrow-band NoSo and NoSπ thresholds with a fixed-frequency masker (50 Hz wide centered around 500 Hz, No=50 dB/Hz) combined with an interaurally in-phase flanking noise band (30 Hz wide, No also 50 dB/Hz) of variable frequency. They found that for monaural detection (i.e., masker and signal both interaurally in phase), the flanking band had only a small effect on the masked thresholds if presented spectrally close to the signal (thresholds increased by less than 1 dB, which is expected on the basis of the increase in the masker energy in the on-frequency channel). For the NoSπ condition, however, a larger effect was observed (up to 2.5 dB), which is difficult to understand in terms of stimulus properties within the on-frequency channel. The experimental data combined with model predictions are given in Fig. 14, where the thresholds are shown as a function of the center frequency of the flanking band. The squares denote the monaural (NoSo) condition, the triangles refer to the binaural (NoSπ) condition. The open symbols are experimental data, the filled symbols are model predictions.

Although there is an overall difference of about 5 dB between experimental data and model predictions for the NoSπ condition, the effect of the flanking band is very similar. If the flanker has a center frequency that is close to the fixed-frequency masker, the thresholds of the model increase up to 2.4 dB. As described in Sec. III C, a narrow-band NoSπ condition allows the reduction of the internal error in the model. The addition of an additive noise band at a higher or lower frequency results in (partial) masking of the off-frequency channels. Therefore, the efficiency of internal error reduction is decreased and an increase in thresholds is observed. In the NoSo condition, this effect is not present, because in this condition, the sample-by-sample variability of noise energy limits the detection. As described in Sec.

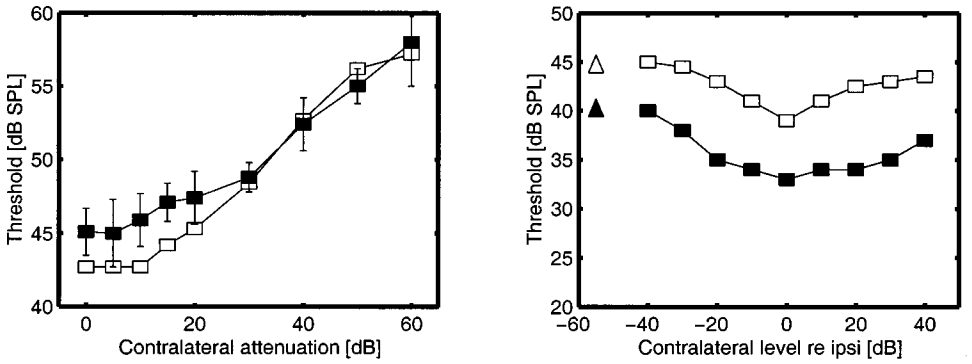

FIG. 15. Left panel: NoSπ thresholds as a function of the attenuation of the stimuli at one ear. The thresholds are expressed in terms of the level at the nonattenuated ear. The open symbols are data adapted from McFadden (1968), the filled symbols are model predictions. Right panel: NoSm thresholds as a function of the noise level in the nonsignal ear relative to the signal ear. The open symbols are adapted from Weston and Miller (1965), the filled symbols are model predictions. The squares denote the NoSm conditions, the triangles are the monaural NmSm reference conditions.

III E, this external variability is correlated across auditory channels and cannot be reduced by the model.

K. NoSπ with interaural disparities in stimulus intensity

In this experiment, performed by McFadden (1968), the masker consisted of an interaurally in-phase wideband noise with a spectral density of 45 dB/Hz; the signal was a 400-Hz interaurally out-of-phase tone. The total stimulus at one ear was attenuated by a variable amount (i.e., both masker and signal had an equal contralateral attenuation, hence the signal-to-masker ratio was the same in both ears). The results are shown in the left panel of Fig. 15. The abscissa shows the disparity in interaural stimulus intensity, the ordinate shows the signal threshold in terms of the level at the ear without attenuation.

Clearly, for a contralateral attenuation between 0 and 10 dB, the thresholds remain constant. In this region, the model can fully compensate for the externally presented IID. If the external IID is increased, however, the thresholds increase. This is the result of the nonlinear preprocessing stage. As described in the accompanying model paper, the adaptation loops which are present in the peripheral processor are highly nonlinear and show a compressive behavior. Therefore, the externally presented IID is reduced and the EI-type element that optimally compensates for the external IID has a characteristic IID that is much smaller than the external IID. Despite the availability of EI-type elements that can compensate for the *mean* level difference at the output of the adaptation loops, these cells cannot cancel the masker noise completely due to the nonlinear processing which results in different waveforms at the output of the adaptation loops from the left and right sides. Hence parts of the noise masker are present in the output of the EI-type element, resulting in increasing thresholds. Moreover, for an IID of 60 dB, the attenuated signal lies below the absolute threshold; the model cannot cancel any part of the masking noise. Hence the thresholds are determined by the monaural masker level in the nonattenuated ear.

A similar experiment was performed by Weston and Miller (1965). They measured NoSm detection thresholds at 500 Hz as a function of the noise level in the nonsignal ear. The noise level in the signal ear was 26 dB/Hz. Besides attenuating, Weston and Miller (1965) also increased the noise level in the nonsignal ear. Their results are shown in the right panel of Fig. 15 as a function of the relative noise level in the nonsignal ear compared to the signal ear. The open symbols denote their experimental data, the filled symbols are model predictions. The triangles refer to the monaural reference conditions. Interestingly, both for a decrease and an *increase* in the contralateral noise level, thresholds increase. This effect is also captured by the model. Both curves show an increase of 3 to 4 dB if the contralateral level is increased by 40 dB. In the model, this increase is caused by the nonlinear processing of the peripheral adaptation loops as described earlier.

The predictions and experimental data for NoSm differ by about 5 to 6 dB, the latter being higher. The maximum BMLD found by Weston and Miller (1965) is about 5.5 dB, while the model predicts a maximum BMLD of about 7 dB. These values are in line with other experimental data, showing BMLDs between 5 and 10 dB for NoSm (Hirsh, 1948; Hirsh and Burgeat, 1958; Kohlrausch, 1988).

L. NoSm as a function of the notchwidth and bandwidth in the nonsignal ear

Hall and Fernandes (1984) measured NoSm detection thresholds for stimuli with a variable masker bandwidth or notchwidth in the nonsignal ear. A 500-Hz pure-tone signal was presented with a 600-Hz-wide band of noise to the signal ear. Bands of noise ranging in width from 25 to 600 Hz, or notched noises (bandwidth also 600 Hz) ranging in notchwidth from 0 to 600 Hz, were presented to the nonsignal ear. The effects of varying the bandwidth were different from those of varying the notchwidth. If the bandwidth was varied, the thresholds decreased over a range of 400 Hz, while for the notched experiment, significant threshold changes only occurred for notchwidths between 0 and 50 Hz. These results are shown in Fig. 16. The left panel corresponds to thresholds as a function of the bandwidth of the masker in the nonsignal ear, the right panel to the notchwidth.

Both model and experimental data show a different behavior as a function of bandwidth for a notched or bandlimited masker. This can be understood as follows. If a

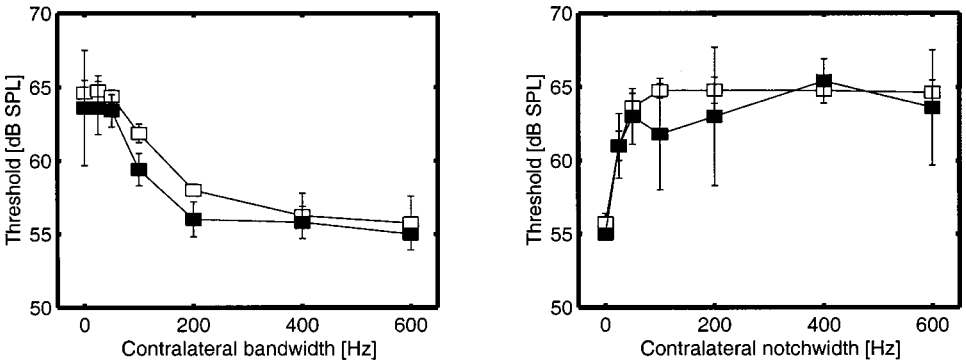

FIG. 16. NoSm thresholds as a function of the bandwidth in the nonsignal ear (left panel) and as a function of the notchwidth in the nonsignal ear (right panel). The open symbols are data adapted from Hall and Fernandes (1984), the filled symbols are model predictions.

narrow-band masker is presented in the nonsignal ear, the masker cannot be canceled by a simple subtraction, since the masker waveforms in both ears are completely different. In fact, for a bandwidth of 0 Hz, the stimulus corresponds to NmSm (i.e., masker and signal are presented to one ear only), and no binaural advantage can be achieved. If the bandwidth in the nonsignal ear is increased, the similarity (or cross correlation) between the maskers at both ears increases which enables the model to cancel the masker more efficiently. This efficiency increases even beyond the critical bandwidth. As a result of this, the signal thresholds decrease with increasing bandwidth, as observed in the data. For the notched noise, a similar process occurs. In the absence of a notch (i.e., a notchwidth of 0 Hz), the masker can be canceled completely and a large binaural advantage is observed. With increasing notchwidth, the cancellation of the masker is less successful and thresholds increase. The essential reason for the different slopes relating threshold to bandwidth or notchwidth lies in the fact that the masker cancellation is performed after peripheral filtering.

As described earlier, the nonsignal ear contains only *part* of the masker of the signal ear; some spectral components are removed. If more spectral components are removed, the amount of masker energy that remains after cancellation increases and hence the signal thresholds increase. Thus, the amount of masker energy *after peripheral filtering* that is present in the signal ear but *not* in the nonsignal ear determines the detection threshold. If the notchwidth is increased from 0 to 25 Hz, there exists a large difference between the maskers at both ears since this part of the masker is in the center of the auditory filter. Therefore, an increase in the notchwidth has a strong effect on the thresholds. Moreover, for notchwidths beyond 50 Hz, the thresholds are equal to the monaural thresholds, indicating that the binaural system cannot increase the detection performance.

In the band-limited case, however, this process is reversed. For a bandwidth of 600 Hz, the maskers in both ears are equal and a large binaural advantage is observed. If the bandwidth is decreased to 400 Hz, the thresholds do not increase since the parts of the masker that are removed in the nonsignal ear are filtered out by the bandpass filter of the inner ear. Only at a bandwidth of 300 Hz do the maskers at both ears become somewhat decorrelated after peripheral filtering and the thresholds show a slight increase. For smaller bandwidths, the correlation between the maskers after filtering decreases and thresholds show a gradual increase. Since the part of the masker that is removed in the nonsignal ear resides in the filter skirt, a reduction in the bandwidth has less effect than an increase in the notchwidth at the center of the filter.

IV. CONCLUSIONS

Predictions for binaural detection performance were shown as a function of the spectral parameters and interaural phase relationships of both maskers and signals. Although some overall differences exist between the model predictions and the experimental data, most of the experimental results can be accounted for with the present model. All stages that are included in the model contribute in some way to the exactness of the predictions. For example, the loss of phase locking in the inner haircells at frequencies above 770 Hz is necessary to account for the ITD thresholds shown in Fig. 1 and the increase of binaural thresholds above 1 kHz (see Fig. 2). For some simulations, almost all stages have to be taken into account to obtain a good fit. Examples are the experiments discussed in Sec. III E (i.e., NπSo as a function of the bandwidth and center frequency). The combined effect of peripheral filtering, compressive behavior of the EI-type elements, the distribution of internal delays, and the chain of adaptation loops in the peripheral preprocessing stage results in good predictions, which cannot be achieved if any of these elements is removed. The price one has to pay is a more complex model than those used and described so far.

In summary, the current model accounts for many binaural detection phenomena in a quantitative way. These include

(i) the wider effective critical bandwidth in bandwidening NoSπ conditions,
(ii) the unification of IID and ITD sensitivity with binaural detection data,
(iii) the level dependence of binaural thresholds,
(iv) the frequency dependence of binaural detection thresholds,
(v) the effect of frequency and bandwidth on the difference between NoSπ and NπSo thresholds,
(vi) the combination of both external stimulus fluctuations and internal errors into one decision variable, and
(vii) the influence of interaural level differences on binaural thresholds.

ACKNOWLEDGMENTS

The investigations were supported by the Research Council for Earth and Life-sciences (ALW) with financial aid from the Netherlands Organization for Scientific Research (NWO). We would like to thank the Associate Editor Wesley Grantham, Steve Colburn, and an anonymous reviewer for their very valuable contributions for improving the contents of the original manuscript.

[1]The value of 0.64 for the interaural correlation was chosen because this was the lowest correlation that was used in a recent study measuring NρSπ thresholds (Breebaart and Kohlrausch, 2001). Data from this study were used because these thresholds were obtained with *frozen* noise. The advantage of using frozen noise is that these thresholds are not influenced by stimulus uncertainty, but are only determined by internal noise.

[2]Note that in the peripheral preprocessing stage of the model, a chain of adaptation loops is included. For an input signal with a constant envelope, the input–output characteristic of the chain of adaptation loops is almost logarithmic. Due to this compression, the signal level must be a certain *fraction* of the masker level to produce a fixed change in an EI-type-element output which is necessary to exceed the internal noise level. Thus, if the masker energy within one auditory channel increases (which is the case towards higher frequencies), the thresholds increase by the same amount.

[3]This mismatch is another example of the same disadvantage caused by the strong overshoot of the adaptation loops [see Dau *et al.* (1996b) for a discussion]. Despite several years of simulations, neither we nor our colleagues in Oldenburg have so far found a satisfying solution for this problem which would reduce the overshoot and preserve the major advantage of this stage: its ability to predict nonsimultaneous masking.

[4]For the filter at 500 Hz in the model, a bandwidth of 42 Hz (i.e., 0.53 times the ERB) resulted in a correlation of zero.

Bos, C. E., and de Boer, E. (**1966**). ''Masking and discrimination,'' J. Acoust. Soc. Am. **39**, 708–715.

Bourbon, W. T. (**1966**). ''Effects of bandwidth and level of masking noise on detection of homophasic and antiphasic tonal signals,'' Ph.D. thesis, University of Texas, Austin, Texas.

Bourbon, W. T., and Jeffress, L. A. (**1965**). ''Effect of bandwidth of masking noise on detection of homophasic and antiphasic tonal signals,'' J. Acoust. Soc. Am. **37**, 1180–1181.

Breebaart, J., and Kohlrausch, A. (**1999**). ''Modeling the role of masker-correlation uncertainty in binaural masking experiments,'' J. Acoust. Soc. Am. **105**, 1391.

Breebaart, J., and Kohlrausch, A. (**2001**). ''The influence of interaural stimulus uncertainty on binaural signal detection,'' J. Acoust. Soc. Am. **109**, 331–345.

Breebaart, J., van de Par, S., and Kohlrausch, A. (**1998**). ''Binaural signal detection with phase-shifted and time-delayed noise maskers,'' J. Acoust. Soc. Am. **103**, 2079–2083.

Breebaart, J., van de Par, S., and Kohlrausch, A. (**1999**). ''The contribution of static and dynamically varying ITDs and IIDs to binaural detection,'' J. Acoust. Soc. Am. **106**, 979–992.

Breebaart, J., van de Par, S., and Kohlrausch, A. (**2001a**). ''Binaural processing model based on contralateral inhibition. I. Model setup,'' J. Acoust. Soc. Am. **110**, 1074–1088.

Breebaart, J., van de Par, S., and Kohlrausch, A. (**2001b**). ''Binaural processing model based on contralateral inhibition. III. Dependence on temporal parameters,'' J. Acoust. Soc. Am. **110**, 1105–1117.

Cokely, J. A., and Hall, J. W. (**1991**). ''Frequency resolution for diotic and dichotic listening conditions compared using the bandlimiting measure and a modified bandlimiting measure,'' J. Acoust. Soc. Am. **89**, 1331–1339.

Colburn, H. S. (**1977**). ''Theory of binaural interaction based on auditory-nerve data. II. Detection of tones in noise,'' J. Acoust. Soc. Am. **61**, 525–533.

Dau, T., Püschel, D., and Kohlrausch, A. (**1996a**). ''A quantitative model of the 'effective' signal processing in the auditory system: I. Model structure,'' J. Acoust. Soc. Am. **99**, 3615–3622.

Dau, T., Püschel, D., and Kohlrausch, A. (**1996b**). ''A quantitative model of the 'effective' signal processing in the auditory system: II. Simulations and measurements,'' J. Acoust. Soc. Am. **99**, 3623–3631.

Durlach, N. I. (**1963**). ''Equalization and cancellation theory of binaural masking-level differences,'' J. Acoust. Soc. Am. **35**, 1206–1218.

Glasberg, B. R., and Moore, B. C. J. (**1990**). ''Derivation of auditory filter shapes from notched-noise data,'' Hear. Res. **47**, 103–138.

Grantham, D. W. (**1984**). ''Interaural intensity discrimination: insensitivity at 1000 Hz,'' J. Acoust. Soc. Am. **75**, 1191–1194.

Hall, J. W., and Fernandes, M. A. (**1984**). ''The role of monaural frequency selectivity in binaural analysis,'' J. Acoust. Soc. Am. **76**, 435–439.

Hall, J. W., and Harvey, A. D. G. (**1984**). ''NoSo and NoSπ thresholds as a function of masker level for narrow-band and wideband masking noise,'' J. Acoust. Soc. Am. **76**, 1699–1703.

Hall, J. W., Tyler, R. S., and Fernandes, M. A. (**1983**). ''Monaural and binaural auditory frequency resolution measured using bandlimited noise and notched-noise masking,'' J. Acoust. Soc. Am. **73**, 894–898.

Hirsh, I. (**1948**). ''The influence of interaural phase on interaural summation and inhibition,'' J. Acoust. Soc. Am. **20**, 536–544.

Hirsh, I., and Burgeat, M. (**1958**). ''Binaural effects in remote masking,'' J. Acoust. Soc. Am. **30**, 827–832.

Holube, I., Kinkel, M., and Kollmeier, B. (**1998**). ''Binaural and monaural auditory filter bandwidths and time constants in probe tone detection experiments,'' J. Acoust. Soc. Am. **104**, 2412–2425.

Klumpp, R. G., and Eady, H. R. (**1956**). ''Some measurements of interaural time difference thresholds,'' J. Acoust. Soc. Am. **28**, 859–860.

Kohlrausch, A. (**1988**). ''Auditory filter shape derived from binaural masking experiments,'' J. Acoust. Soc. Am. **84**, 573–583.

Kollmeier, B., and Holube, I. (**1992**). ''Auditory filter bandwidths in binaural and monaural listening conditions,'' J. Acoust. Soc. Am. **92**, 1889–1901.

Langhans, A., and Kohlrausch, A. (**1992**). ''Spectral integration of broadband signals in diotic and dichotic masking experiments,'' J. Acoust. Soc. Am. **91**, 317–326.

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

McFadden, D. (**1968**). ''Masking-level differences determined with and without interaural disparities in masker intensity,'' J. Acoust. Soc. Am. **44**, 212–223.

McFadden, D., Jeffress, L. A., and Ermey, H. L. (**1971**). ''Difference in interaural phase and level in detection and lateralization: 250 Hz,'' J. Acoust. Soc. Am. **50**, 1484–1493.

Metz, P., Bismarck, G., and Durlach, N. (**1968**). ''Further results on binaural unmasking and the EC model. II. Noise bandwidth and interaural phase,'' J. Acoust. Soc. Am. **43**, 1085–1091.

Mills, A. (**1960**). ''Lateralization of high-frequency tones,'' J. Acoust. Soc. Am. **32**, 132–134.

Rabiner, L., Laurence, C., and Durlach, N. (**1966**). ''Further results on binaural unmasking and the EC model,'' J. Acoust. Soc. Am. **40**, 62–70.

Sever, J., and Small, A. (**1979**). ''Binaural critical masking bands,'' J. Acoust. Soc. Am. **66**, 1343–1350.

Sondhi, M. M., and Guttman, N. (**1966**). ''Width of the spectrum effective in the binaural release of masking,'' J. Acoust. Soc. Am. **40**, 600–606.

Stern, R. M., and Shear, G. D. (**1996**). ''Lateralization and detection of low-frequency binaural stimuli: Effects of distribution of internal delay,'' J. Acoust. Soc. Am. **100**, 2278–2288.

van de Par, S. (**1998**). ''A comparison of binaural detection at low and high frequencies,'' Ph.D. thesis, Eindhoven University of Technology, Eindhoven.

van de Par, S., and Kohlrausch, A. (**1998**). ''Diotic and dichotic detection using multiplied-noise maskers,'' J. Acoust. Soc. Am. **103**, 2100–2110.

van de Par, S., and Kohlrausch, A. (**1999**). ''Dependence of binaural masking level differences on center frequency, masker bandwidth and interaural parameters,'' J. Acoust. Soc. Am. **106**, 1940–1947.

van der Heijden, M., and Trahiotis, C. (**1998**). ''Binaural detection as a function of interaural correlation and bandwidth of masking noise: Implications for estimates of spectral resolution,'' J. Acoust. Soc. Am. **103**, 1609–1614.

Weston, P., and Miller, J. (**1965**). ''Use of noise to eliminate one ear from masking experiments,'' J. Acoust. Soc. Am. **37**, 638–646.

Wightman, F. (**1971**). ''Detection of binaural tones as a function of masker bandwidth,'' J. Acoust. Soc. Am. **50**, 623–636.

Zerbs, C. (**2000**). ''Modeling the effective binaural signal processing in the auditory system,'' Ph.D. thesis, Oldenburg University, Germany.

Zurek, P. M. (**1991**). ''Probability distributions of interaural phase and level differences in binaural detection stimuli,'' J. Acoust. Soc. Am. **90**, 1927–1932.
Zurek, P. M., and Durlach, N. I. (**1987**). ''Masker-bandwidth dependence in homophasic and antiphasic tone detection,'' J. Acoust. Soc. Am. **81**, 459–464.
Zwislocki, J., and Feldman, R. S. (**1956**). ''Just noticeable differences in dichotic phase,'' J. Acoust. Soc. Am. **28**, 860–864.