Binaural processing model based on contralateral inhibition. III. Dependence on temporal 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 paper and two accompanying papers [Breebaart *et al.*, J. Acoust. Soc. Am. **110**, 1074–1088 (2001); **110**, 1089–1104 (2001)] describe a computational model for the signal processing of the binaural auditory system. The model consists of several stages of monaural and binaural preprocessing combined with an optimal detector. Simulations of binaural masking experiments were performed as a function of temporal stimulus parameters and compared to psychophysical data adapted from literature. For this purpose, the model was used as an artificial observer in a three-interval, forced-choice procedure. All model parameters were kept constant for all simulations. Model predictions were obtained as a function of the interaural correlation of a masking noise and as a function of both masker and signal duration. Furthermore, maskers with a time-varying interaural correlation were used. Predictions were also obtained for stimuli with time-varying interaural time or intensity differences. Finally, binaural forward-masking conditions were simulated. The results show that the combination of a temporal integrator followed by an optimal detector in the time domain can account for all conditions that were tested, except for those using periodically varying interaural time differences (ITDs) and those measuring interaural correlation just-noticeable differences (jnd's) as a function of bandwidth. © *2001 Acoustical Society of America.* [DOI: 10.1121/1.1383299]

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

I. INTRODUCTION

This is the third paper describing our binaural signal detection model and its ability to predict binaural detection thresholds in a great variety of experimental conditions. This model basically consists of three stages (Breebaart *et al.*, 2001a). The first stage simulates the effective signal processing of the basilar membrane and the inner hair cells 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, and the central processor matches the representations of the presented stimuli to templates (derived during previous presentations); on this basis the model indicates which interval contains the signal.

In the second paper of this series (Breebaart *et al.*, 2001b), model predictions for binaural detection were discussed as a function of the spectral parameters of the stimuli, keeping the temporal parameters constant. All stimuli had a duration which was long compared to the temporal resolution of both the monaural and binaural stages of the model (i.e., 200 ms or longer). We demonstrated that the model is very successful in describing the threshold dependence on spectral stimulus parameters and that this success can, to a large extent, be attributed to an optimal combination of information across auditory channels.

In the current paper, we focus on the temporal properties of the stimuli, keeping the spectral parameters constant. Two important temporal properties are studied intensively. The first concerns temporal *integration*. It has been shown that the binaural system is able to integrate binaural cues temporally if such a process enhances a detection task. For example, an increase of the signal duration in an NoSπ condition results in a decrease of the signal threshold for signal durations up to 300 ms (Zwicker and Zwicker, 1984; Wilson and Fowler, 1986; Wilson and Fugleberg, 1987; Bernstein and Trahiotis, 1999). The second property is related to the temporal *resolution* of the binaural auditory system. Several studies have revealed that the auditory system is sluggish in its processing of interaural differences. For example, the minimum audible angle of a sound source strongly depends on its velocity (Perrott and Musicant, 1977). Experiments using time-varying interaural intensity differences (IIDs) revealed that IID detection shows a low-pass behavior with a cutoff frequency of about 20 Hz (Grantham, 1984). The detection of dynamic ITDs seems to be even worse; Grantham and Wightman (1978) showed that ITD detection has a low-

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

pass response with a cutoff frequency of 2 to 5 Hz. Detection experiments performed with a masker which has a time-varying interaural correlation show that the binaural auditory system has a time constant between 44 and 243 ms (Grantham and Wightman, 1979; Kollmeier and Gilkey, 1990; Culling and Summerfield, 1998; Akeroyd and Summerfield, 1999), which is rather high compared to the 4 to 44 ms for monaural processing (Kollmeier and Gilkey, 1990; Plack and Moore, 1990). The aim of the current study is to demonstrate that the model presented in Breebaart *et al.* (2001a) can also account for these temporal phenomena.

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 paper (i.e., temporal 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 equivalent rectangular bandwidth (ERB) estimates published by Glasberg and Moore (1990). Due to the limited bandwidth of the filters in the gammatone filterbank, ringing occurs which influences forward-masking thresholds for very short signal delays.

(ii) A chain of five adaptation loops is included in the peripheral preprocessor. These adaptation loops limit the detectability of short low-level signals presented shortly after the offset of a high-level masker. Due to this limitation, the monaural detection model by Dau *et al.* (1996a) has been successful in predicting detection performance in monaural nonsimultaneous masking conditions (Dau *et al.*, 1996b, 1997). Because the current model includes the same stages as the model by Dau *et al.* (1996a), this predictive scope is inherited by our model. Furthermore, because the binaural interaction follows the peripheral adaptation (cf. Kohlrausch and Fassel, 1997), binaural forward masking will also be limited in its steepness through the presence of the adaptation loops.

(iii) Central temporal window. In the binaural processor, EI-type elements calculate the squared-difference signal between the outputs of the peripheral processor for each auditory filter. These difference signals are convolved with a double-sided exponential window with an equivalent rectangular duration (ERD) of 60 ms to account for a limited binaural temporal resolution. Because this window operates on the difference signal, the same window is used to analyze IIDs and ITDs as well as binaural detection data.

(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) 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 influence of the internal noise, and to accumulate information about the signal by adapting its observation interval (matched temporal integrator).

B. Procedure and stimuli

The procedure, the method of generation of stimuli, and the model calibration were the same as those described in the spectral paper (Breebaart *et al.*, 2001b). In particular, all model parameters were kept fixed for all simulations described in this paper and were the same as in the previous paper. For details regarding the procedure and model calibration, refer to Breebaart *et al.* (2001b). In all simulations, the duration, level, on- and offset ramps, bandwidth, and onset delay of both the maskers and signals equaled the values used in the experiments with human subjects. If more data sets from various authors with different experimental settings were used, the experimental 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, we either calculated binaural masking level differences (BMLDs) or normalized the thresholds with the spectral level of the masking noise.

III. SIMULATIONS

A. NρSπ and NρSm correlation dependence for wideband noise

This section deals with the detection of a signal in the presence of a masker with various (fixed) values of the interaural correlation (ρ). Similar stimuli were discussed in the spectral paper (Breebaart *et al.*, 2001b, Sec. III E) focusing mainly on the bandwidth dependence of the masker. In this section, the nonstationary behavior of interaural differences in such conditions will be discussed. By ''nonstationary'' we mean that the expected values of the statistical properties, such as the interaural correlation, are constant, but that these properties evaluated on a short-time basis change as a function of time within each interval. It is therefore valuable to discuss NρSπ thresholds in the current paper, since these experiments reflect the detection of a change in the *distribution* of interaural differences rather than the detection of the *presence* of interaural cues.

Robinson and Jeffress (1963) measured thresholds for a wideband NρSπ condition. They used a 150-ms, 500-Hz tone as signal presented in a 150-ms noise masker. The masker had a spectral level of 50 dB/Hz. Their data are shown in the left panel of Fig. 1 (open symbols), together

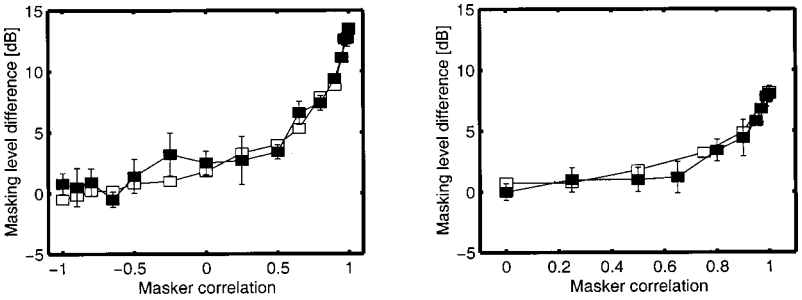

FIG. 1. NρSπ (left panel) and NρSm (right panel) BMLDs as a function of the interaural masker correlation. The white symbols are experimental data adapted from Robinson and Jeffress (1963) for NρSπ and from Wilbanks and Whitmore (1968) for NρSm. The black symbols are model predictions.

with the model predictions (filled symbols). The binaural masking level difference (BMLD) is shown as a function of the interaural masker correlation. For a correlation of $+1$, the condition corresponds to NoSπ and a large BMLD is observed, which decreases with decreasing correlation. If the masker is interaurally uncorrelated ($\rho=0$), a BMLD of only 2 to 3 dB is present, which completely disappears if the correlation is decreased further towards -1.

The right panel shows data for a monaural signal (i.e., NρSm), adapted from Wilbanks and Whitmore (1968). In this experiment, a 200-ms signal was used, also with a frequency of 500 Hz. The spectral masker level was 33 dB/Hz. The data show a similar curve as the NρSπ condition, with two important differences. First, the BMLD at $\rho=+1$ is 6 dB smaller for the NρSm condition. Second, almost no BMLD is observed in the NρSm condition with $\rho=0$, while the NρSπ condition still shows a BMLD of a few dB at this masker correlation.

To understand why the BMLDs decrease with a decrease in the masker correlation, it is useful to first have a closer look at the way these partially correlated maskers are generated. Usually, the Nρ masker is obtained by combining the waveforms of two or three independent noise sources. We will focus on the method using two noise sources, keeping in mind that the method using three noise sources is in principle similar (cf. van der Heijden and Trahiotis, 1997). If two independent noise sources having time-domain waveforms given by $N_1(t)$ and $N_2(t)$ are used to generate a noise with an interaural correlation of ρ, the left and right channels $\mathrm{L}(t)$ and $\mathrm{R}(t)$ consist of the following linear combination of these noises:

$$\begin{aligned}\mathrm{L}(t)&=\tfrac{1}{2}\sqrt{2}\sqrt{1+\rho}N_1(t)+\tfrac{1}{2}\sqrt{2}\sqrt{1-\rho}N_2(t),\\ \mathrm{R}(t)&=\tfrac{1}{2}\sqrt{2}\sqrt{1+\rho}N_1(t)-\tfrac{1}{2}\sqrt{2}\sqrt{1-\rho}N_2(t).\end{aligned}\tag{1}$$

If such a stimulus is presented to the model, and if we neglect the processing of the peripheral preprocessor, the waveforms $\mathrm{L}(t)$ and $\mathrm{R}(t)$ enter an EI-type element that optimally cancels the masker (no internal delay or level adjustment, i.e., $\tau=\alpha=0$). In our implementation (see Breebaart *et al.*, 2001a), the output (E) of the EI-type element is then given by

$$E(t)=(\mathrm{L}(t)-\mathrm{R}(t))^2.\tag{2}$$

Substitution of Eq. (1) into Eq. (2) results in

$$E(t)=(2-2\rho)N_2^2(t).\tag{3}$$

In a very similar way, it can be shown that the addition of an interaurally out-of-phase signal $[S_\pi(t)]$ to the same masker results in

$$E(t)=(2-2\rho)N_2^2(t)+4S_\pi^2(t).\tag{4}$$

The processes in the model that follow the EI-type element basically consist of time averaging resulting in a running average E of the energy of $(2-2\rho)N_2(t)$, followed by the logarithmic input–output relation of the EI-type elements. As can be observed from Eq. (3), a decreasing interaural correlation ρ results in an increasing amount of masker energy that cannot be canceled by the EI-type elements. As described in the first paper (Breebaart *et al.*, 2001a), this results in higher signal thresholds due to the logarithmic input–output behavior. The above explanation also holds for the NρSm condition, except for the fact that the amount of signal energy in Eq. (4) is decreased by a factor of 4. Thus, to achieve a similar change in E as for an Sπ signal, the signal level must be increased by 6 dB, an effect that is clearly found in the data for $\rho>0.7$. For lower correlation values, the signal level in the NρSm condition approaches the monaural threshold and hence thresholds remain constant if the correlation is reduced further.

B. NρSπ thresholds for narrow-band noise

Breebaart and Kohlrausch (2001) measured NρSπ thresholds as a function of both the correlation and the masker bandwidth. The masker duration was 300 ms. A 500-Hz sinusoid with a duration of 200 ms was used as the signal. Breebaart and Kohlrausch (2001) found that for a narrow-band masker with a bandwidth of 10 Hz, thresholds varied more with the interaural correlation than for a wideband masker. The narrow-band masker resulted in a much steeper curve for correlations between 0.8 and $+1$ compared to the wideband case. This is depicted in Fig. 2. Both the 10-Hz-wide (squares) and the 1000-Hz-wide (triangles) data are shown as a function of the masker correlation. The black symbols denote the model predictions; the white symbols are experimental data. The separation between the narrow-band and the broadband data is due to the choice of a constant overall masker level of 65 dB SPL, which gives a higher spectral level for the 10-Hz-wide masker.

Breebaart and Kohlrausch (2001) argued that the differences between the 10-Hz curve and the 1000-Hz curve are due to the fact that two different factors limit the detection process: internal errors and external variability. For the wideband masker, the thresholds are determined by the internal

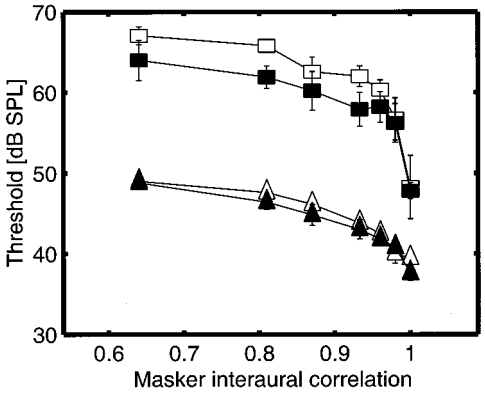

FIG. 2. Running-noise NρSπ thresholds for a masker bandwidth of 10 Hz (squares) and 1000 Hz (triangles). The overall masker level was 65 dB SPL for both bandwidths. The black symbols are model predictions; the white symbols are data adapted from Breebaart and Kohlrausch (2001).

errors. In our model, the specific relation between threshold and interaural correlation results from the logarithmic input–output curve present in the EI-type elements which was explained in Sec. II A.

For the narrow-band condition, other factors are important. The temporal sluggishness filter effectively calculates a running average of the output of Eq. (3) if a masker alone is present. This output increases if the signal is present. Because this output serves as a decision variable, the model must look for fluctuations in this variable that are attributable to the addition of the signal. Since the squared waveform of $N_2(t)$ is present in Eq. (3), a running average of $N_2(t)$ is obtained, multiplied with a scalar which depends on the interaural correlation. If the bandwidth of N_2 is very small, the energy estimate E shows large fluctuations due to the limited number of degrees of freedom in the noise. If the standard deviation of these fluctuations is larger than the *change* in E due to the addition of the signal with a certain level, it is very unlikely that the model is able to detect the signal because of the large amount of *stimulus variability*. This is exactly what happens in the narrow-band NρSπ condition. Instead of being limited by internal errors, the thresholds are limited by external stimulus fluctuations if $\rho<1$. Since the amount of E energy increases with decreasing correlation [as can be observed from Eq. (3)], the amount of fluctuations in E also increases with decreasing correlation, hence resulting in increasing thresholds. This property is not altered by the logarithmic input–output function of the EI-type elements, since both the fluctuations and the change in the output due to the signal are transformed by the same process. Thus, for the narrow-band condition with ρ at or below 0.98, thresholds are relatively high due to the "external" variability in E. Only when $\rho=1$ does external variability play no role, and thresholds are only determined by the internal error and are therefore relatively low.

C. Interaural cross-correlation discrimination

Gabriel and Colburn (1981) measured just-noticeable differences (jnd's) in interaural cross correlation from two reference correlations (0 and +1) at several bandwidths. The total noise level was kept constant at 75 dB SPL and the stimuli were spectrally centered at 500 Hz. At a reference correlation of +1, their results indicated that for bandwidths less than 115 Hz the correlation jnd was equal to about 0.004, while for larger bandwidths, the jnd increased monotonically with the noise bandwidth. On the other hand, at a reference correlation of 0, the jnd *decreased* with increasing bandwidth, having a value of 0.7 for narrow-band stimuli (3 Hz) and 0.3 for the broadband case (4500 Hz). Their results are summarized in Fig. 3. The left panel shows the correlation jnd's at a reference correlation of +1, the right panel at a reference correlation of 0. The white symbols are the experimental data for different subjects; the black symbols denote model predictions.

The model predictions for a reference correlation of +1 (left panel) show a completely different behavior from the experimental data: the experimental data increase with increasing bandwidth, while the model predictions show a monotonic decrease with increasing bandwidth. Only for the data at 40 and 115 Hz is there a close resemblance between model predictions and experimental data. The decrease in correlation jnd with increasing bandwidth for the model can be explained as follows. For all bandwidths, the reference condition has a correlation of +1. Consequently, the stimulus can be eliminated completely by the model. Thus, the reference intervals result in no internal activity for EI-type elements that are optimally tuned to this detection task. If the interaural correlation of the stimulus is reduced, the masker cannot be canceled completely, which results in some activity in the model. If the stimulus bandwidth is small (i.e., 3 to 10 Hz), this cue for detection is available in the on-frequency filter and in some adjacent filters due to spread of excitation. If the bandwidth is increased, the number of auditory filters that contains information about the change in the correlation increases since the change in the interaural correlation occurs

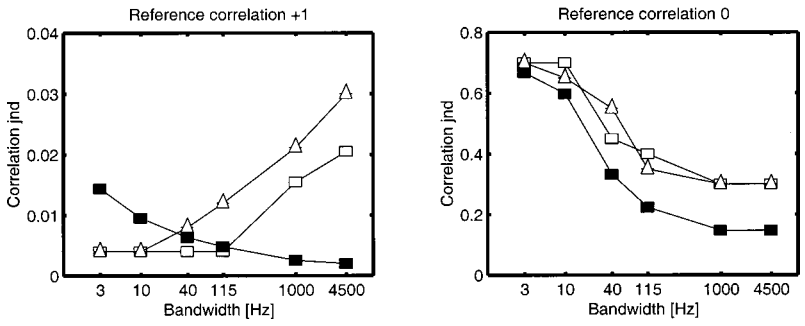

FIG. 3. Just-noticeable differences in the interaural correlation at a reference correlation of +1 (left panel) and 0 (right panel) as a function of the bandwidth of the stimulus. The white symbols are experimental data adapted from Gabriel and Colburn (1981) for two different subjects. The filled symbols are model predictions.

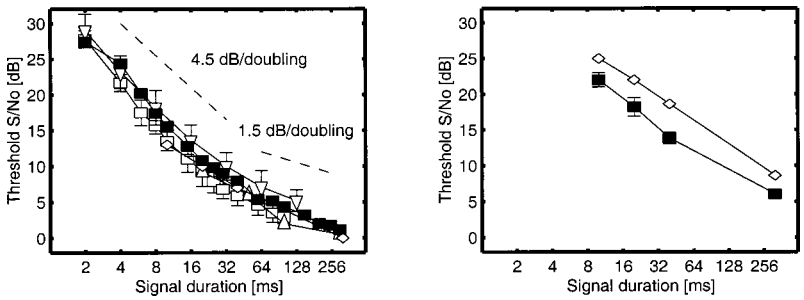

FIG. 4. NoSπ thresholds as a function of the signal duration. Data in the left panel are for a signal frequency of 500 Hz, in the right panel for a signal of 4 kHz. The white squares are data adapted from Wilson and Fugleberg (1987), the upward triangles from Yost (1985), the downward triangles from Wilson and Fowler (1986), and the diamonds from Bernstein and Trahiotis (1999). The black symbols are model predictions. Thresholds are given as signal level *re* masker spectrum level.

at the complete spectrum of the stimulus. This enables the model to integrate information across auditory filters, resulting in an effective reduction of the internal noise. This in turn results in lower thresholds, as observed in the model predictions. This effect is not observed in the experimental data, however.

The right panel of Fig. 3 shows data for the reference correlation of 0. Both model predictions and experimental data show a decrease in the correlation jnd with increasing bandwidth, although the model is more sensitive to changes in the correlation at bandwidths beyond 40 Hz. This decrease in the correlation jnd can be explained by considering stimulus uncertainty. If the bandwidth is very small (3 Hz), the fluctuations in the output of the EI-type elements are very large (see Sec. III B.). Increases in the bandwidth result in more degrees of freedom in the masker and hence less uncertainty in the output of the EI elements. This decreased uncertainty is reflected by the decrease in the correlation jnd. For bandwidths beyond 115 Hz, increases in the bandwidth have almost no effect on the stimulus uncertainty because these parts of the stimulus fall outside the auditory filter and hence thresholds remain constant.

D. NoSπ signal duration

An important property that has a very strong effect on thresholds in an NoSπ condition is the duration of the signal. The threshold behavior in this experimental paradigm basically reflects the ability of the binaural auditory system to *integrate* information over time. Several studies showed an increase in detection performance of up to 25 dB if the signal duration is increased from 2 to 250 ms (cf. Yost, 1985; Wilson and Fowler, 1986; Wilson and Fugleberg, 1987; Bernstein and Trahiotis, 1999). The results of these studies are shown in Fig. 4 (white symbols) for a center frequency of 500 Hz (left panel) and for 4 kHz (right panel). Since in these studies different noise levels were used ranging from 26.3 to 47 dB/Hz, we expressed thresholds as the ratio between signal level and spectral masker density. The thresholds decrease with a slope of 4.5 dB/oct for durations up to about 60 ms, while for longer durations, this slope is shallower (1.5 dB/oct). The model predictions are shown by the black symbols in Fig. 4. They were derived for a 500-ms wideband noise masker with a spectral energy density of 36.2 dB/Hz (similar to Wilson and Fugleberg, 1987).

The model predictions show a very similar signal-duration dependence to the experimental data. These results can be explained as follows. First, consider a signal of very short duration (2 ms). In this case, the signal interval contains interaural differences which are present within a short period of time compared to the duration of the binaural window (about 60 ms). As described in the first model paper (Breebaart *et al.*, 2001a), the output of each EI-type element is averaged over time by a temporal integrator. Consequently, the cue for detection at the *averaged* output of such a temporal window is strongly reduced for a very short signal. Therefore, the signal must have a relatively high level to elicit a change in the averaged output that can be detected by the model. If the duration of the signal is increased but does not exceed the duration of the temporal window, the average output of the temporal window increases and hence the signal level decreases at threshold. Since the output of the temporal averager is proportional to the signal energy, this process accounts for a decrease of 3 dB for each doubling of the of signal duration, as long as the signal duration does not exceed the time constant of the temporal averager. For durations exceeding 60 ms, this process does not influence the detection process.

An additional effect of 1.5 dB per duration doubling results from the reduction of the internal error with increasing signal duration by the optimal detector. A longer signal duration means that the average output is available for a longer time. This enables the model to reduce the internal error in the decision variable in a similar way as was described for spectral integration of information (see Breebaart *et al.*, 2001b, for details). After the temporal averager, an additive noise is combined with the output of the EI-type elements, followed by an optimal detector. If, after the temporal integrator, the cue for detection is available for a long time, the optimal detector can decrease the amount of noise in its decision variable by integrating the EI-element output over the total signal duration. In this way, a doubling in the signal duration results in a doubling in the overall difference in output between masker and masker plus signal, while the amount of noise increases with the square root of 2. Hence, the detectability of the signal is increased, which results in a lower threshold. Thresholds are expected to decrease with 1.5 dB per doubling of signal duration. Thus, the combined effect of the processes described above accounts for the 4.5 dB per doubling for signal durations below 60 ms and 1.5 dB beyond 60 ms.

The slope of 4.5 dB/doubling for signal durations below 60 ms and 1.5 dB/doubling beyond 60 ms is also present in the model simulations at 4 kHz (right panel of Fig. 4). Except for an overall difference of about 4 dB, these model predictions are very close to the experimental data, although

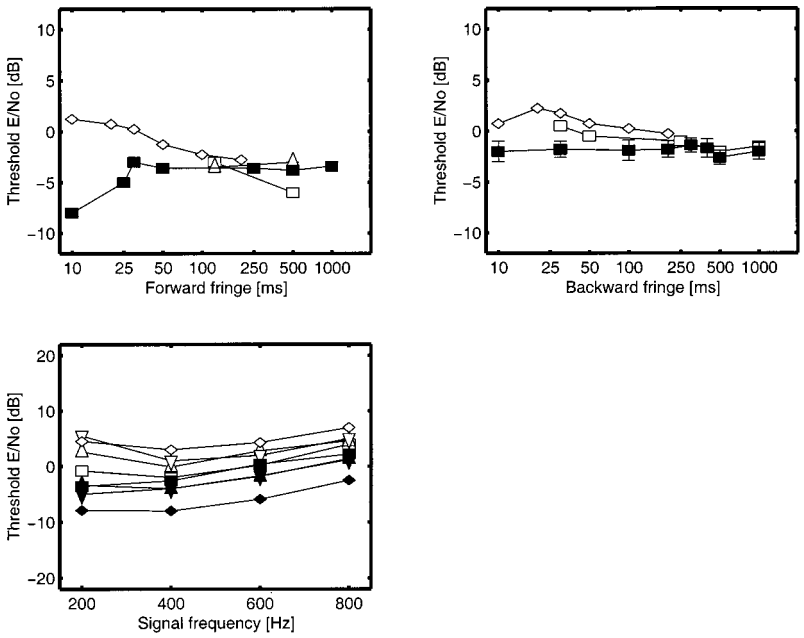

FIG. 5. NoSπ thresholds as a function of the masker duration. The upper-left panel denotes experimental data with a forward fringe, the upper-right panel with a backward fringe. The lower panel shows thresholds for a signal temporally centered in a masker for four different masker durations (squares denote 500 ms, upward triangles 100 ms, downward triangles 50 ms, and diamonds 25 ms) as a function of the center frequency. The white squares in the upper panels are experimental data for a 500-Hz, 32-ms signal adapted from Robinson and Trahiotis (1972), the upward triangles are from the same study with a 256-ms signal, the diamonds are adapted from Zwicker and Zwicker (1984) for a 400-Hz signal. The black symbols are model predictions for a 500-Hz, 20-ms signal.

the difference in slope below and beyond 60 ms cannot be deduced from the experimental data.

E. NoSπ masker duration

Besides reports on the duration of the signal, several studies have been published on the effect of the duration of the masker. Basically three configurations have been tested. The first uses a forward noise fringe of variable length (Robinson and Trahiotis, 1972; Zwicker and Zwicker, 1984; Yost, 1985), while in the second configuration, a backward noise fringe is used (Trahiotis *et al.*, 1972; Zwicker and Zwicker, 1984; Yost, 1985). A third condition includes a signal which is temporally centered in the masker (Kohlrausch, 1986). The duration of the fringe in these studies varied between 10 and 256 ms; the center frequency was always 500 Hz (except for Zwicker and Zwicker, 1984, using 400 Hz). Most experiments were performed with relatively short signals (10 to 32 ms). The experimental results are shown in Fig. 5. The upper-left panel shows data for a forward fringe, the upper-right panel for a backward fringe. The lower panel shows four different curves for four different masker durations (25, 50, 100, and 500 ms) as a function of center frequency. The white symbols are data from different data sets, the black symbols are the model predictions for a 20-ms, 500-Hz signal added to a broadband masker with a level of 50 dB/Hz (similar to Trahiotis *et al.*, 1972). Thresholds are expressed as the ratio between signal energy and spectral masker density to enable comparison between data sets.[1]

The predicted thresholds hardly change if the masker duration is varied. This is in contrast to what is observed if the *signal* duration is changed, as discussed in Sec. D. A maximum decrease of 5 dB is observed experimentally if the forward masker fringe is increased (top-left panel of Fig. 5). The model predictions show no effect of the masker duration at all, except for very short forward fringes. The absence of a distinct effect of masker duration is expected, since the output of an EI-type element optimally tuned for this condition has no activity for a masker alone and has some activity during the presence of the signal. Since this activity is used as a cue for the detection, and since it does not depend on the duration of the masker, the model's predictions do not, in principle, depend on the masker duration. The decrease in thresholds for very short forward fringes is a consequence of the monaural adaptation loops (Dau *et al.*, 1996a). These loops are not completely charged during the first 25 ms of the stimulus and a substantial overshoot in the output of the adaptation loops exists just after the onset. If interaural differences are presented within this short period of time, they will result in a stronger change at the output of EI-type elements than if they were presented with a longer forward masker fringe. Therefore, the thresholds are up to 5 dB lower during the first 25 ms. This effect is clearly visible in the upper-left panel of Fig. 5 for short forward fringes and in the lower panel (with the centered signal). In the latter case, simulated thresholds are more than 10 dB lower than the experimental data for a 25-ms masker (2.5-ms forward fringe). For masker durations of 50 ms and beyond, the model predicts nearly constant thresholds, while the experimental data show a slight decrease with increasing masker duration. On the other hand, the backward fringe has no influence on the state of the adaptation loops; hence, the thresholds do not depend on the duration of the backward fringe, as demonstrated in the right panel of Fig. 5.

F. Maskers with phase transitions in the time domain

The previous sections dealt with the ability of the binaural auditory system to *integrate* information over time. Besides integration, another very important temporal property that can be measured is the temporal *resolution* of the system. In analogy to the frequency-domain phase transition which we discussed in the accompanying paper (Breebaart *et al.*, 2001b), an interaural phase transition can be applied in the time domain. As the spectral phase transition enabled the estimation of the spectral resolution, the time-domain

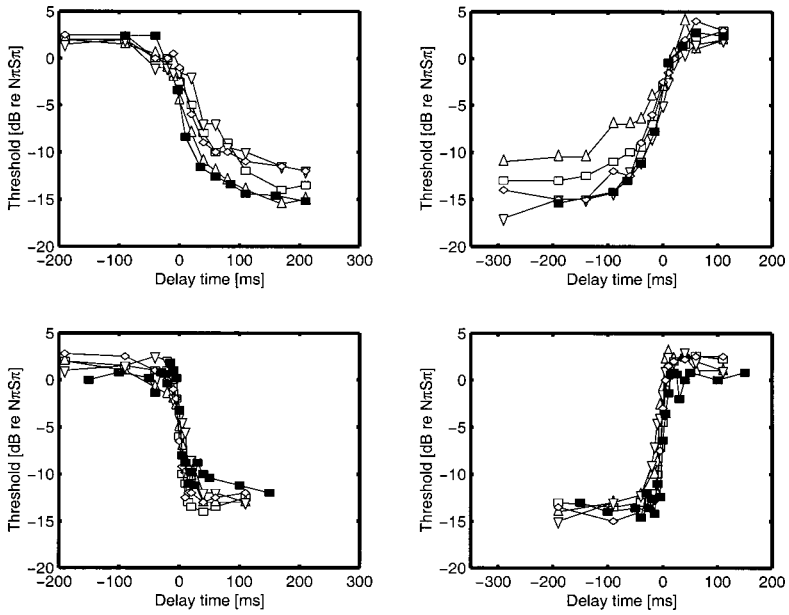

FIG. 6. $N\pi oS\pi$ (upper-left panel) and $No\pi S\pi$ (upper-right panel) thresholds as a function of the temporal position of the signal center relative to the masker–phase transition. The lower-left and lower-right panels correspond to the monaural $N\pi(-15dB)N\pi S\pi$ and $(-15dB)N\pi N\pi S\pi$ conditions, respectively (see the text). The 0-dB point on the ordinate denotes the nontransient $N\pi S\pi$ thresholds. White symbols are experimental data for different subjects adapted from Kollmeier and Gilkey (1990); the black symbols are model predictions.

equivalent enables estimation of the temporal resolution. One possible realization of such a phase transition is a masker which is first interaurally in phase and then interaurally out of phase. This condition is referred to as $No\pi S\pi$ if an interaurally out-of-phase signal is used. In a similar way, $N\pi oS\pi$ refers to a masker that is interaurally out of phase first, followed by an in-phase noise. If the signal is centered within the in-phase masker portion, the effective condition is $NoS\pi$ and a large BMLD is observed. On the other hand, if the signal is presented during the out-of-phase masker portion, no BMLD is expected. Experimental data have shown that for signal positions close to the masker phase transition, a gradual change of the threshold is observed (Kollmeier and Gilkey, 1990; Holube *et al.*, 1998). The experimental data of individual subjects from Kollmeier and Gilkey (1990) are shown in Fig. 6. These data were measured with a 500-Hz, 20-ms signal added to a broadband noise with a spectral energy density of 40 dB/Hz. The thresholds are plotted as a function of the signal center relative to the occurrence of the phase transition of the masker. The upper-left panel corresponds to an $N\pi oS\pi$ condition, the upper-right panel to $No\pi S\pi$. In both conditions, a gradual change in threshold is observed near the phase transition (0 ms). The model predictions (black symbols) have a similar gradual change in threshold as the experimental data and follow the lower bound of the four subjects.

The gradual change as observed in Fig. 6 is thought to reflect the *temporal resolution* of the binaural auditory system. In the model, this resolution is limited by the temporal averager at the output of all EI-type elements; the stepwise masker correlation change is heavily smoothed by the averager, and hence thresholds show a gradual change instead of a stepwise one.

The lower panels in Fig. 6 show data obtained for a corresponding ''monaural'' condition. Both the signal and the two masker portions were presented interaurally out of phase ($N\pi N\pi S\pi$), and one of the masker portions was decreased in level by 15 dB. The data in the lower-left panel were obtained for a masker that drops by 15 dB at the second half of the total stimulus [this condition is referred to as $N\pi(-15dB)N\pi S\pi$], while the data in the lower-right panel were obtained for a 15-dB increase at the stimulus center. The open symbols denote data from different subjects adapted from Kollmeier and Gilkey (1990); the filled symbols are model predictions. Clearly, the time constant for processing monaural cues is much shorter than for the processing of binaural cues. The ability of the present model to also predict monaural forward- and backward masking relies on a completely different process from that involved in predicting the binaural temporal resolution effects. The monaural data are predicted due to the presence of the adaptation loops prior to any binaural interaction. Since the model incorporates all the stages of the monaural model described by Dau *et al.* (1996a, b), it inherits the predictive power of that model for all cases where no binaural interaction is needed.

An extension of the experiment with one masker phase transition in the time domain is obtained by using two phase transitions. Culling and Summerfield (1998) measured detection thresholds of a 500-Hz 20-ms $S\pi$ signal which was added to a broadband in-phase noise masker (No) of variable duration, preceded and followed by 400 ms of interaurally uncorrelated noise. The spectral energy density of the noise was 40 dB/Hz. We refer to this condition as $NuouS\pi$. Culling and Summerfield (1998) found that thresholds decrease by up to 12 dB with increasing No duration from 20 to 960 ms. The data are shown in Fig. 7. The white symbols are data for three different subjects; the black symbols are model predictions.

For No durations between 20 and 400 ms, the modeled thresholds agree well with the subjects' data. For further increases in No duration there is a discrepancy, because the experimental data tend to decrease further, while the modeled thresholds remain constant. This indicates that the temporal *resolution* of the human binaural auditory system is very well represented by the model, but that some really long-term processes with a temporal extension of 500 ms and more are not covered by our present model structure.

Besides using stepwise correlation changes in the

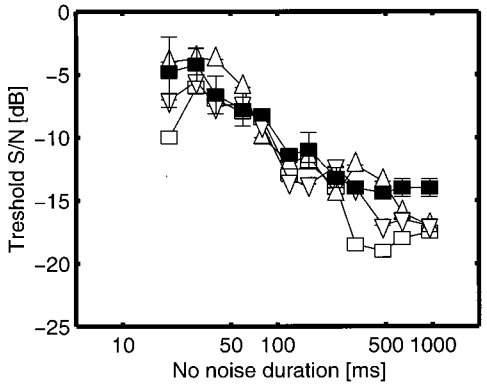

FIG. 7. NuouSπ detection thresholds as a function of the duration of the No noise. White symbols are experimental data for different subjects adapted from Culling and Summerfield (1998); black symbols are model predictions.

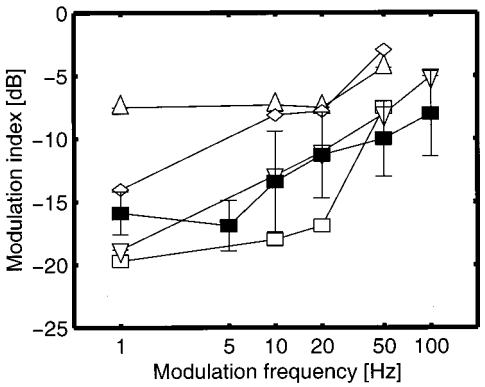

FIG. 9. Modulation depth (in dB) required for discrimination of interaural in-phase modulation from out-of-phase modulation as a function of the modulation frequency. The white symbols denote results for different subjects adapted from Grantham (1984); the black symbols are model predictions.

masker, experiments have also been performed with a sinusoidally changing interaural masker correlation. Holube *et al.* (1998) measured the detectability of a 500-Hz, 20-ms signal as a function of the correlation modulation period. The signal was always centered at a temporal position where either an No or Nπ noise was present. The masker duration was 2500 ms for the modulation periods of 2- and 1 s and 750 ms for shorter periods. A bandlimited masker (0.1 to 1 kHz) was used with an overall level of 75 dB SPL. The results (white symbols) combined with model predictions (black symbols) are shown in Fig. 8. The left panel corresponds to a signal presentation at a position where the noise was interaurally in phase, the right panel where it was out of phase. Thresholds are expressed relative to the monaural (NoSo) thresholds.

The thresholds for the signal presentation centered on No decrease with increasing modulation period. This is the result of the decreasing amount of Nπ noise at the input of the temporal window of the EI-type element during the presentation of the signal. A similar argument holds for the presentation during the Nπ masker phase. With decreasing period, an increasing amount of Nπ noise is present in the EI-type element during signal presentation, resulting in higher thresholds.

In contrast to earlier modeling approaches (Kollmeier and Gilkey, 1990; Holube *et al.*, 1998), our implementation does explain data for stepwise and sinusoidal correlation modulation with the *same* temporal window. We will come back to this observation in the General Discussion (Sec. IV).

G. Discrimination of dynamic interaural intensity differences

Grantham (1984) measured observers' ability to detect time-varying interaural intensity differences. The stimuli consisted of interaurally uncorrelated noises of which the envelopes were sinusoidally modulated. The task was to discriminate between a modulation which was interaurally in phase and a modulation which was interaurally out of phase, the latter resulting in interaural intensity differences. The noise used in this experiment had a bandwidth of 0.4 octaves centered at 500 Hz and had a level of 75 dB SPL. The stimulus duration was 1000 ms. The results of Grantham (1984) (open symbols) combined with the model predictions (filled symbols) are shown in Fig. 9.

As can be observed in Fig. 9, the thresholds increase with increasing modulation rate, indicating a low-pass modulation function of the binaural auditory system. Although the data show large variations across subjects, the model predictions are a good representation of the subject denoted by the downward triangles. The general trend is that of a low-pass filter; the modulation index required at threshold increases with increasing modulation frequency. In our model, this can be understood as follows. The unmodulated noise that was used had an interaural correlation of 0. If a sinusoidal modulation is superimposed on the noise waveform, the interaural correlation of the noise remains unchanged. Now, consider the output of an EI-type element centered at the stimulus center frequency and $\alpha=\tau=0$. For the interaurally in-phase

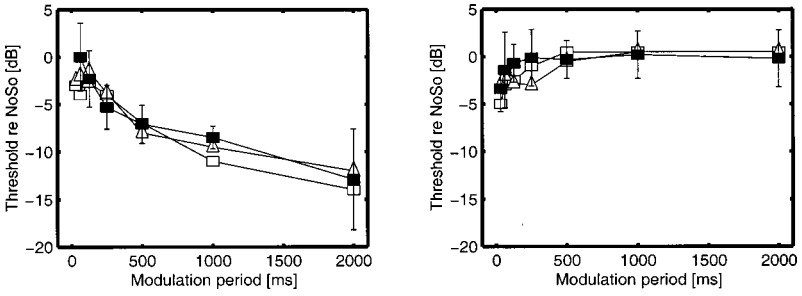

FIG. 8. Detection thresholds for an Sπ signal presented in a noise masker with a sinusoidally modulated interaural correlation. Thresholds are plotted relative to the monaural (NoSo) threshold. The left panel corresponds to a signal presentation at a point where the interaural masker correlation was +1, the right panel to a masker correlation of −1. The white symbols are data adapted from Holube *et al.* (1998); the black symbols are model predictions.

modulator (i.e., the reference stimuli), the output of the EI-type element has a similar pattern as the modulator: during positive modulator phases the masker energy is increased and hence the uncorrelated masker results in an increased amount of activity. On the other hand, a negative modulator phase results in a decrease in the EI activity. These modulations are, however, only present for low modulation frequencies (<10 Hz); for higher modulation frequencies the EI-output modulation depth decreases due to the temporal averaging. Thus, as long as the modulation period is beyond the time constant of the temporal averager of the EI-type elements, the externally presented monaural modulation is reflected by a modulation of the EI-type element output. The out-of-phase modulator results in hardly any modulation in the EI-type element output: every attenuation (i.e., negative modulator phase) of the left-ear signal is accompanied by an amplification of the right-ear signal (i.e., a positive modulator phase) and vice versa, resulting in only a very small effect on the EI activity. Thus, as long as the modulations in the EI-type output due to the in-phase modulator are clearly visible (i.e., slow modulations), the model shows a low modulation threshold which increases with increasing modulation frequency.

H. Discrimination of dynamic interaural time differences

Grantham and Wightman (1978) measured the detectability of sinusoidally time-varying interaural time differences present in a low-pass noise. The spectra of the noise stimuli were approximately flat between 10 and 3000 Hz. The presentation level was 70 dB SPL, the duration 440 ms. Grantham and Wightman (1978) found that the peak ITD required for detection strongly depends on the modulation frequency, having a value of about 30 μs at a modulation rate of 0 Hz, which increases to 90 μs at 20 Hz. Interestingly, the thresholds decrease again for higher modulation rates, reaching a value of about 30 μs at a modulation rate of 500 Hz. The results are shown in Fig. 10. The white symbols denote the experimental data; the black symbols are model predictions.

The bell-shaped curve which is seen in the experimental data is not observed in the model predictions. Considering the properties of the model, this is expected, since the average or peak interaural difference that occurs in this stimulus does not depend on the modulation rate. Therefore, the model predictions do not show a bell-shaped curve but decrease monotonically by a factor of about 2.5. This decrease

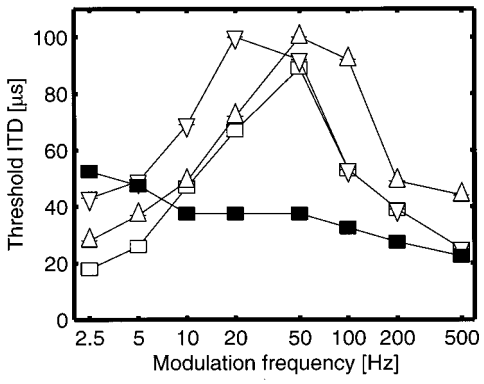

FIG. 10. Peak interaural time difference in microseconds at threshold as a function of the modulation frequency for the detection of sinusoidally varying interaural time differences. The open symbols are experimental data adapted from Grantham and Wightman (1978) representing different subjects; the black symbols are model predictions.

is related to the fact that the ITD at the onset of the signal is set to 0 and changes sinusoidally during the stimulus. If a modulation rate of 10 Hz is used, the first maximum in the ITD occurs 25 ms after the stimulus onset. At a modulation rate of 20 Hz, the maximum occurs at 12.5 ms, etc. As discussed in Sec. III E, interaural differences closer to the onset of the stimulus result in lower thresholds due to the overshoot in the peripheral adaptation loops. Therefore, the ITD thresholds shown in Fig. 10 decrease with increasing modulation rate.

I. Binaural forward masking

In the previous experiments, the signal was always presented simultaneously with the masker. If a short signal is presented *after* the masker, a phenomenon referred to as forward masking is observed. For signals that are presented at increasing delays with respect to the masker offset, the thresholds decrease gradually towards the absolute threshold instead of showing a stepwise change (cf. Punch and Carhart, 1973; Yama, 1992; Kohlrausch and Fassel, 1997). This elevation is observed for signal delays of up to 200 ms. Moreover, a binaural release of masking can be observed if the signal is presented interaurally out of phase compared to an in-phase signal. For example, Yama (1992) measured forward-masking thresholds for a 10-ms, 250-Hz signal combined with a low-pass (0–2.5 kHz, overall level of 70 dB SPL), 500-ms running noise. Linear ramps of 5-ms duration were used to gate both signal and masker. The results show a BMLD of about 14 dB for simultaneous masking, which

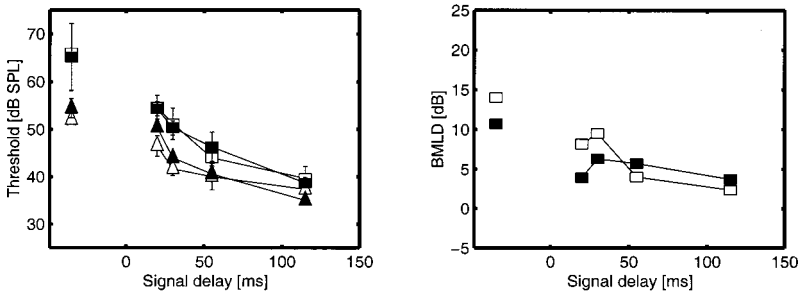

FIG. 11. Signal thresholds for NoSπ (triangles) and NoSo conditions (squares) as a function of the time difference between masker and signal offset (left panel) and corresponding BMLDs (right panel). White symbols are experimental data adapted from Yama (1992); black symbols are model predictions. The signal duration was 10 ms. The symbols at the left in both panels denote thresholds for simultaneous masking.

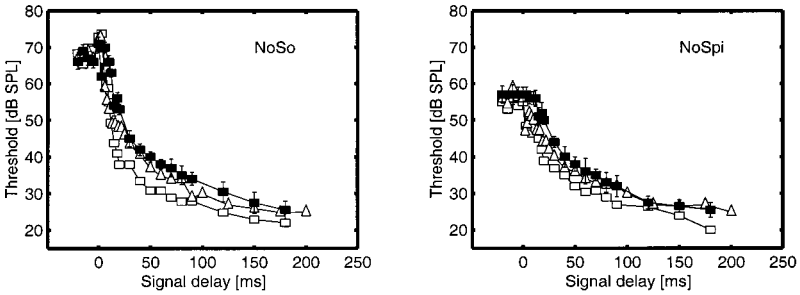

FIG. 12. Forward-masking thresholds for a 20-ms So signal (left panel) and an Sπ signal (right panel) as a function of the time difference between masker and signal offset. A 300-ms frozen-noise (No) served as the masker. The white symbols are experimental data adapted from Kohlrausch and Fassel (1997); the black symbols are model predictions.

decreases to a few dB for signal delays of 100 ms, as can be observed in the left panel of Fig. 11. Thresholds are shown as a function of the time difference between signal and masker offset. The squares denote monaural (i.e., NoSo) thresholds; the triangles denote binaural (NoSπ) thresholds. The right panel shows the corresponding BMLDs, for both the model and the experimental data.

As can be observed from Fig. 11, the model (black symbols) shows a similar threshold behavior as the experimental data. For simultaneous masking, a BMLD of 14 dB is observed in the experimental data and a few dB less for the model predictions. Both the binaural and monaural forward-masking thresholds show a decrease towards the absolute threshold, which is about 35 dB for both the So and Sπ signal. In the region of 0 to 100 ms, a substantial BMLD can be observed which is, however, smaller than the BMLD for simultaneous masking.

Results that seem to be in contradiction with those found by Yama (1992) were obtained by Kohlrausch and Fassel (1997). Their forward-masking experiments only showed BMLDs for signal delays up to 20 ms instead of the 100 ms found by Yama (1992). The data were obtained with a 300-ms frozen-noise masker which was spectrally flat between 20 and 1000 Hz. The overall masker level was 70 dB SPL. A 20-ms, 500-Hz signal was used. The same values were used to obtain the model predictions. However, the frozen-noise sample was different from the one used in the experiments. The results and model predictions are shown in Fig. 12. The left panel corresponds to monaural conditions (NoSo), the right panel to binaural conditions (NoSπ).

In the NoSo condition, thresholds start to decrease as soon as the offset of the signal occurs after the masker offset. In contrast, NoSπ thresholds remain constant for signal delays up to about 10 ms. For larger signal delays, the thresholds gradually decrease with signal delay towards the absolute threshold. This decrease is stronger for the monaural (NoSo) condition than the binaural (NoSπ) condition. Hence, a substantial BMLD is only found for signal delays up to 20 ms. Kohlrausch and Fassel (1997) also measured forward masking thresholds for a 20-ms masker. The waveform of this short masker was identical to the last 20 ms of the long masker. The results are shown in Fig. 13 in the same format as Fig. 12.

The main difference between the threshold behavior of 300-ms and 20-ms maskers is the slope of the forward-masking curve, which is steeper for the short masker. In the model, this steeper curve is the result of the fact that the adaptation loops which are part of the peripheral preprocessor are not completely ''charged'' if a masker of only 20 ms is used (Dau *et al.*, 1996b).

Also in this condition, the model shows BMLDs only for signal delays of up to 20 ms, perfectly in line with the experimental data. The difference in BMLD behavior between the conditions shown Fig. 11 on the one hand and in Figs. 12 and 13 on the other hand is related to the difference in signal duration that was used. Yama (1992) used a relatively short signal (10 ms), while Kohlrausch and Fassel (1997) used a 20-ms signal. If a short signal is presented to the model, the onset of the signal will result in an increase of the output of the peripheral adaptation loops compared to the output in the absence of the signal. On the other hand, if the signal is turned off, the adaptation loops are (at least partially) adapted to the (higher) input signal; hence, the signal offset results in a decrease of the output. Moreover, due to the adaptation of the system, the activity after the offset will be *less* than if *no signal* was present. An example of this property can be seen in the lower panel of Fig. 7 of Dau *et al.* (1996a): the presence of the signal results in both an overshoot at the signal onset and an undershoot at the signal offset. If the duration is sufficiently long compared to the temporal resolution of the monaural system, the model can use both the overshoot at the onset of the signal and the undershoot at the offset of the signal to detect the signal's presence. If a very short signal is used, however, the tempo-

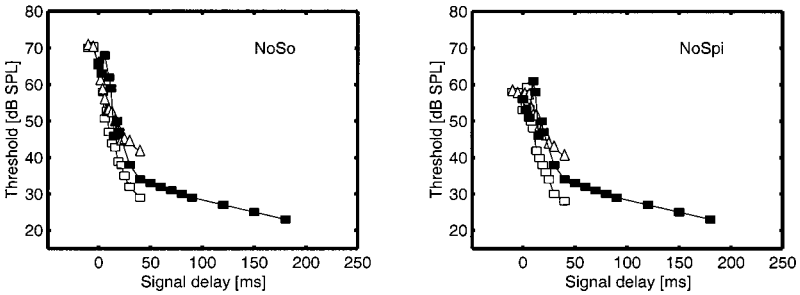

FIG. 13. Forward-masking thresholds for a 20-ms So signal (left panel) and an Sπ signal (right panel) as a function of the time difference between masker and signal offset. In this case, the masker had a duration of only 20 ms. The white symbols are experimental data adapted from Kohlrausch and Fassel (1997); the black symbols are model predictions.

ral averager at the output of the adaptation loops partially cancels the undershoot by the overshoot, resulting in a smaller overall effect at the output of the temporal averager. In the binaural case, the temporal window does not reduce the detectability of the signal because the window is applied *after* the computation of the *squared difference* between the left and right channels. Therefore, monaural thresholds are elevated more strongly if the signal duration is decreased from 20 to 10 ms than binaural thresholds. This is also observed in the model predictions. If a 20-ms signal is used (Figs. 12 and 13), both the monaural and binaural cues are about equally strong and no BMLD is observed for signal delays beyond 20 ms. When using a 10-ms signal, however, the monaural thresholds are elevated more than the binaural thresholds, resulting in a BMLD which is still present even for signal delays up to 100 ms.

IV. GENERAL DISCUSSION

We have shown that our binaural model is quite successful in describing the dependence of binaural detection thresholds on temporal stimulus properties. These properties include the effect of signal and masker duration, forward masking, and detection against stimuli with a time-varying interaural correlation. By means of a temporal integrator followed by an optimal detector in the time domain, the model accounts both for temporal resolution and for temporal integration properties within a single framework. It is interesting to note that a similar framework in the spectral domain is present in the model (i.e., a set of bandpass filters followed by an optimal detector in the frequency domain), also leading to very good predictions as a function of spectral stimulus parameters (see Breebaart *et al.*, 2001b).

Although many of the simulations shown in this paper and in the previous papers show a good fit between the data and the predictions, very similar results would probably be obtained if the basic EI interaction in the model was replaced by an EE (or cross-correlation) interaction. There are, however, some specific experimental conditions that may give rise to some modeling difficulties. In the first model paper (Breebaart *et al.*, 2001a), we expected that models based on the interaural cross correlation may not account for the effect of both signal and masker duration. Two methods of computing the interaural correlation were discussed. The first comprised computation over the complete stimulus (i.e., masker duration). We argued that this method would result in a strong increase of detection thresholds with an increase of masker duration, which is not found in experimental data. The second method comprised computation of the correlation only for the stimulus part that contains the signal. In this case, a maximum effect of 1.5 dB/oct of signal duration is expected, which is not in line with experimental results showing a stronger influence of signal duration for durations below about 60 ms. It is therefore difficult to explain the effects of both masker and signal duration with a model based on cross correlation. The simulation results in Sec. III D and III E show that a model based on EI-type interaction in combination with an optimal detector shows a performance which is more in line with the experimentally obtained results.

The simulations as a function of interaural correlation and bandwidth revealed that the detection performance of the model can in principle be limited by two sources of errors, namely stimulus uncertainty in the externally presented signals and errors in the internal accuracy (internal noise).

The data in Figs. 6 and 7 (stepwise correlation change in the time domain) revealed that a double-exponential window with time constants of 30 ms (equivalent rectangular duration, or ERD, of 60 ms) can account for the limited temporal resolution of the binaural system which is observed in different experiments, which often reveals different underlying temporal windows. The shape and ERD of the window were chosen to fit the experimental results of Kollmeier and Gilkey (1990). A somewhat larger value for the ERD of 100 ms was found by Culling and Summerfield (1998). Their estimate was based on a Gaussian window. They also fitted their data with a double-exponential window, resulting in ERDs between 48 and 117 ms, which is much closer to the temporal window we used throughout the whole paper. These results support the fact that the ERD itself is not a very valuable property to discuss without mentioning the window shape from which it is derived.

The data in Fig. 8 were obtained for a sinusoidally changing masker correlation. The ERD found by Holube *et al.* (1998) that fitted these data (using a double-exponential window) was 91 to 122 ms. A similar experiment by Grantham and Wightman (1979) revealed an ERD of 139 to 189 ms. Despite these rather large ERDs compared to the ERD of our model, the simulations give a good fit to the data. This suggests that not only the shape of the temporal window, but also the stimulus configuration has an influence on the ERD that is estimated from experimental data: experiments with stepwise correlation changes result in lower estimates of the time constants than sinusoidal correlation changes (Kollmeier and Gilkey, 1990; Holube *et al.*, 1998).

A possible explanation for these differences in the estimate of the ERD was given by Holube *et al.* (1998). They stated that ''the reason for this mismatch seems to be the different detection strategies employed for the various tasks that are affected by the consistency of binaural information across frequency and time.'' In their fitting procedure, Kollmeier and Gilkey (1990) and Holube *et al.* (1998) obtained the predicted thresholds by computing the weighted integration of the instantaneous interaural cross correlation *at the temporal center of the signal*. For the sinusoidal changes in the correlation, it is likely that this detection strategy results in the highest signal-to-masker ratio, given the fact that both the temporal window and the correlation modulation are symmetric around that moment. It is not obvious, however, that this strategy is also optimal for the stepwise correlation changes. In fact, an analysis of the optimal detector in our model revealed that within the framework of our model, the optimal position for detecting the signal is about 10 ms further away from the masker phase transition (i.e., off-time listening). Hence, by analyzing the correlation slightly away from the temporal center of the signal, lower thresholds are obtained. Consequently, the fitting procedure used by Kollmeier and Gilkey (1990) and Holube *et al.* (1998) for step-

wise correlation changes *underestimates* the time constants present in the system. This is perfectly in line with their experimental results: the time constants for stepwise correlation changes were about a factor of 2 lower than for the sinusoidal correlation changes.

One of the experimental findings that the present model could not account for is the bandwidth dependence of interaural correlation jnd's for a reference correlation of +1 (see Fig. 3). The experimental data show a substantial decrease in performance with bandwidth, while the model's performance increases. A possible explanation for this discrepancy is that the binaural auditory system can only integrate cues across frequency if these cues *are highly correlated* across auditory filters. Since the data show the strongest increase in the correlation jnd for bandwidths beyond the ERB of the auditory filters, it is likely that the increase in the thresholds results from across-frequency effects instead of within-filter stimulus properties. If a stimulus with a correlation of +1 is presented, the stimulus can be canceled completely. The reduction of the correlation can thus in the model be detected by an increase in the residual noise after optimal cancellation. If the noise is broadband, this residue is in principle independent across peripheral filters. Our model does not incorporate the correlation of cues across frequency. However, it could be possible that the binaural auditory system does.

Another experimental result that cannot be accounted for by the model is the low-pass characteristic that is obtained with dynamically varying ITDs (Fig. 10). A model that can account for these data is the position-variable model of Stern and Bachorski (1983). In their model, the intracranial locus of the stimulus is estimated by computing a weighted centroid of the running cross-correlation function. The running cross-correlation function is computed using an exponentially decaying averaging window with a fixed time constant. If the ITD is modulated with a period that is longer than the duration of the temporal averager, the peak of the running cross-correlation function follows the externally presented ITD. Consequently, the weighted centroid modulates similarly. If the ITD modulation frequency is increased, the low-pass characteristic of the running cross correlation averages the ITD over time. This results in a lower but wider peak of the cross-correlation function. Moreover, the maximum displacement of the peak decreases. Consequently, the amplitude of the time-varying centroid of the cross correlation *decreases*. Since this centroid is used as a decision variable, thresholds increase with increasing modulation rate. In fact, the model of Stern and Bachorski (1983) was very successful in predicting the data shown in Fig. 10 for the left side of the bell-shaped curve.

It would in principle be possible to modify our model in such a way that it does not use increases in the activity in the EI-type pattern but rather an estimate of the position of the sound source as a decision variable. This can be facilitated by scanning the position of the minimum within the EI-type element activity pattern. The reason that we did not base the model's decision process on a position variable is that such an approach has a detrimental effect on the fits of other data. For example, in our second paper (Breebaart *et al.*, 2001b), NoSπ thresholds were discussed as a function of the bandwidth of the masking noise. The predicted thresholds as well as the experimental data show approximately constant thresholds for bandwidths up to twice the ERB of the peripheral filters. We expect that a position-variable model, independently of whether the binaural interaction is based on EE or EI processing, cannot account for this result. The addition of an Sπ signal to a diotic masker results in the presence of dynamically varying IIDs and ITDs in the stimulus. The rate of fluctuation of these differences depends on the bandwidth; a larger bandwidth corresponds to faster fluctuations. Consequently, the NoSπ thresholds for a position-variable model are expected to increase with increasing bandwidth, which is not in line with the experimental data. In summary, we do not see how the absence of an effect of the ITD modulation rate in bandlimited NoSπ conditions and the clear effect of ITD modulation shown by Grantham and Wightman (1978) can be explained with the same detection mechanism.

Finally, the strong overshoot in the peripheral adaptation loops which results in lower detection thresholds for a binaural signal presented during the first 25 ms of the masker is an unwanted effect. In a certain way, we can end this series of binaural modeling papers with a remark similar to one made at the end of the discussion by Dau *et al.* (1996b). The temporal dynamics and nonlinear compression effects of the adaptation loops are useful in understanding a number of binaural effects like binaural forward masking and the influence of overall interaural level differences on binaural unmasking and on lateralization, while for some specific conditions, their temporal dynamic is too strong. Obviously, we so far have not found the optimal realization of this stage. Therefore, we will, together with our colleagues in Oldenburg, continue in our efforts to improve the understanding of this part of our monaural and binaural models.

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, Richard Stern, and an anonymous reviewer for their valuable suggestions for improving the original manuscript.

[1]The data adapted from Zwicker and Zwicker (1984) were measured at 400-Hz center frequency instead of 500 Hz. In contrast to the other studies, their masking noise was not spectrally flat but had a spectral energy distribution that is referred to as *uniform masking noise*. The amplitude spectrum of this noise type is spectrally flat from 0 to 500 Hz and decreases with 10 dB/dec above 500 Hz. To be able to use their data, we calculated the spectral energy density of their masking noise at 400-Hz center frequency and used this value to plot the data.

Akeroyd, M. A., and Summerfield, A. Q. (**1999**). "A binaural analog of gap detection," J. Acoust. Soc. Am. **105**, 2807–2820.

Bernstein, L. R., and Trahiotis, C. (**1999**). "The effects of signal duration on NoSo and NoSπ thresholds at 500 Hz and 4 kHz," J. Acoust. Soc. Am. **105**, 1776–1783.

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. (**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. II. Dependence on spectral parameters,'' J. Acoust. Soc. Am. **110**, 1089–1104.
Culling, J. F., and Summerfield, Q. (**1998**). ''Measurements of the binaural temporal window using a detection task,'' J. Acoust. Soc. Am. **103**, 3540–3553.
Dau, T., Kollmeier, B., and Kohlrausch, A. (**1997**). ''Modeling auditory processing of amplitude modulation. II. Spectral and temporal integration,'' J. Acoust. Soc. Am. **102**, 2906–2919.
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.
Gabriel, K. J., and Colburn, H. S. (**1981**). ''Interaural correlation discrimination. I. Bandwidth and level dependence,'' J. Acoust. Soc. Am. **69**, 1394–1401.
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**). ''Discrimination of dynamic interaural intensity differences,'' J. Acoust. Soc. Am. **76**, 71–76.
Grantham, D. W., and Wightman, F. L. (**1978**). ''Detectability of varying interaural temporal differences,'' J. Acoust. Soc. Am. **63**, 511–523.
Grantham, D. W., and Wightman, F. L. (**1979**). ''Detectability of a pulsed tone in the presence of a masker with time-varying interaural correlation,'' J. Acoust. Soc. Am. **65**, 1509–1517.
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.
Kohlrausch, A. (**1986**). ''The influence of signal duration, signal frequency and masker duration on binaural masking level differences,'' Hear. Res. **23**, 267–273.
Kohlrausch, A., and Fassel, R. (**1997**). ''Binaural masking level differences in nonsimultaneous masking,'' in *Binaural and Spatial Hearing in Real and Virtual Environments*, edited by R. H. Gilkey and T. Anderson (Lawrence Erlbaum Assoc., Mahwah, NJ), Chap. 9, pp. 169–190.
Kollmeier, B., and Gilkey, R. H. (**1990**). ''Binaural forward and backward masking: evidence for sluggishness in binaural detection,'' J. Acoust. Soc. Am. **87**, 1709–1719.
Perrott, D. R., and Musicant, A. D. (**1977**). ''Minimum auditory movement angle: Binaural localization of moving sound sources,'' J. Acoust. Soc. Am. **62**, 1463–1466.
Plack, C. J., and Moore, B. C. J. (**1990**). ''Temporal window shape as a function of frequency and level,'' J. Acoust. Soc. Am. **87**, 2178–2187.
Punch, J., and Carhart, R. (**1973**). ''Influence of interaural phase on forward masking,'' J. Acoust. Soc. Am. **54**, 897–904.
Robinson, D., and Jeffress, L. (**1963**). ''Effect of varying the interaural noise correlation on the detectability of tonal signals,'' J. Acoust. Soc. Am. **35**, 1947–1952.
Robinson, D. E., and Trahiotis, C. (**1972**). ''Effects of signal duration and masker duration on detectability under diotic and dichotic listening conditions,'' Percept. Psychophys. **12**, 333–334.
Stern, R. M., and Bachorski, S. J. (**1983**). ''Dynamic cues in binaural perception,'' in *Hearing—Physiological Bases and Psychophysics*, edited by R. Klinke and R. Hartmann (Springer, Berlin), pp. 209–215.
Trahiotis, C., Dolan, T. R., and Miller, T. H. (**1972**). ''Effect of backward masker fringe on the detectability of pulsed diotic and dichotic tonal signals,'' Percept. Psychophys. **12**, 335–338.
van der Heijden, M., and Trahiotis, C. (**1997**). ''A new way to account for binaural detection as a function of interaural noise correlation,'' J. Acoust. Soc. Am. **101**, 1019–1022.
Wilbanks, W., and Whitmore, J. (**1968**). ''Detection of monaural signals as a function of interaural noise correlation and signal frequency,'' J. Acoust. Soc. Am. **43**, 785–797.
Wilson, R., and Fowler, C. (**1986**). ''Effects of signal duration on the 500-Hz masking-level difference,'' Scand. Audiol. **15**, 209–215.
Wilson, R., and Fugleberg, R. (**1987**). ''Influence of signal duration on the masking-level difference,'' J. Speech Hear. Res. **30**, 330–334.
Yama, M. (**1992**). ''Effects of temporal separation and masker level on binaural analysis in forward masking,'' J. Acoust. Soc. Am. **91**, 327–335.
Yost, W. A. (**1985**). ''Prior stimulation and the masking-level difference,'' J. Acoust. Soc. Am. **78**, 901–906.
Zwicker, U. T., and Zwicker, E. (**1984**). ''Binaural masking-level difference as a function of masker and test-signal duration,'' Hear. Res. **13**, 215–220.