# The Role of the Precedence Effect in Sound Source Lateralization

by

Daniel E. Shub

BSE Bioengineering
University of Pennsylvania, 1997

SUBMITTED TO THE DEPARTMENT OF ELECTRICAL ENGINEERING AND COMPUTER SCIENCE IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE IN ELECTRICAL ENGINEERING AT THE MASSACHUSETTS INSTITUTE OF TECHNOLOGY

JUNE 2001



Signature of Author: _ _____
Department of Electrical Engineering and Computer Science
April 24, 2001

Certified by: _____ _____
H. Steven Colburn
Professor of Biomedical Engineering
Boston University
Thesis Supervisor

Accepted by: _____ _____
Arthur C. Smith
Department of Electrical Engineering and Computer Science
Chairman, Committee for Graduate Students

The Role of the Precedence Effect in Sound Source Lateralization

by

Daniel E. Shub

Submitted to the Department of Electrical Engineering and Computer Science on April 24, 2001
in Partial Fulfillment of the requirements for the Degree of Masters of Science in Electrical Engineering

## Abstract

Sound source lateralization can be accomplished in many real world environments with a simple cross-correlation model. The precedence effect predicts that in certain situations, such as sounds with a single loud reflection, sound source lateralization performance of normal hearing listeners could be better than performance predicted by a simple cross-correlation model. This thesis examines sound source lateralization of ongoing broadband noise targets, alone and in the presence of a single ongoing broadband noise jammer. Different jammer conditions were tested when the jammer was either a simple reflection of the target or a sound uncorrelated with the target. Two measurements were made for each subject: The identification threshold of ±300 μsec ITD targets was determined through a 1-interval 2-alternative forced-choice identification experiment; the lateralization threshold was determined through a 1-interval lateralization experiment. In both cases the minimum target-to-jammer level for criterion performance was estimated as the threshold. The identification threshold was 3.2 dB lower (P value of 0.027) for simple reflection jammers compared to jammers that were uncorrelated with the target. The lateralization threshold was 2.85 dB lower (P value of 0.040) for the reflection jammers relative to the uncorrelated jammers. A comparison between the lateralization thresholds of normal-hearing listeners and two cross-correlation based models was also made. Both models obtained lateralization thresholds as low as -13 dB, up to 10 dB better than normal-hearing performance. There was a slight trend, in agreement with current understanding of the precedence effect, for both the identification thresholds and the lateralization thresholds of the simple reflection jammers to be dependent on the jammer ITD. Simple reflection jammers with an ITD of 0 μsec had thresholds that were 1.9 dB lower (P value of 0.104) than the thresholds of simple reflection jammers with an ITD of 643 μsec.

Thesis Supervisor: H. Steven Colburn
Title: Professor of Biomedical Engineering, Boston University

## Table of Contents


**I. INTRODUCTION** ..... **5**

**II. METHODS** ..... **9**

A. SUBJECTS AND APPARATUS ..... 9
B. IMPULSE RESPONSES ..... 9
*i. Target* ..... *9*
*ii. Jammer* ..... *10*
C. STIMULI ..... 11
D. EXPERIMENTAL PROCEDURE ..... 13
*i. Experiment 1: Lateralization-in-quiet* ..... *14*
*ii. Experiment 2: Identification* ..... *14*
*iii. Experiment 3: Lateralization* ..... *15*
*iv. Procedure* ..... *16*
E. MODEL ..... 17
F. DATA ANALYSIS ..... 17

**III. RESULTS** ..... **20**

A. EXPERIMENT 1: LATERALIZATION-IN-QUIET ..... 20
B. EXPERIMENT 2: IDENTIFICATION ..... 21
C. EXPERIMENT 3: LATERALIZATION ..... 22
D. MODEL ..... 23

**IV. DISCUSSION** ..... **24**

**V. CONCLUSIONS** ..... **27**

**ACKNOWLEDGEMENTS** ..... **28**

**FIGURES** ..... **29**

**APPENDIX** ..... **42**

**REFERENCES** ..... **79**


## Figure list


Fig. 1 Common magnitude spectrum of the target and jammer stimuli. (Spectrum from Gardner and Martin 1994) 29

Fig. 2 The first 10 msec of the left and right ear impulse responses (IRs) corresponding to the different jammer conditions. 30

Fig. 3 Example cross-correlation functions of the sound stimuli IRs for the different stimulus conditions over a range of target-to-jammer ratios. (TJRs) 31

Fig. 4 Confusion matrices for all subjects in the target-in-quiet condition. 32

Fig. 5 Statistics summarizing the performance of subject S3 during the target-in-quiet condition as a function of 100-trial block number. 33

Fig. 6 Identification and lateralization thresholds and *smart* and *dumb* model predictions for the different jammer conditions. 34

Fig. 7 Confusion matrices of subject S4, jammer condition B'. 35

Fig. 8 Normalized $R^2$ for subject S2 and all jammer conditions. The continuous curves are cumulative Gaussian fits. 36

Fig. 9 Cumulative Gaussian fits to the normalized $R^2$ data for all subjects in condition B. 37

Fig. 10 Confusion matrices of Model 1, the *dumb* cross-correlation model, for jammer condition B'. 38
Fig. 11 Cumulative Gaussian fits to the normalized $R^2$ data of Model 1 for all the jammer conditions. 39
Fig. 12 Confusion matrices of Model 2, the *smart* cross-correlation model, for jammer condition B'. 40
Fig. 13 Cumulative Gaussian fits to the normalized $R^2$ data of Model 2 for all the jammer conditions. 41

## Table List

Table 1 Properties of the jammer IRs for the 6 different jammer conditions. 11
Table 2 Average values and the standard errors of the different statistics for the target-in-quiet condition. 21
Table 3 Identification thresholds for each subject and condition. 22
Table 4 Lateralization thresholds for each subject and condition. 23

# I. Introduction

Many factors affect sound source lateralization. These factors include reverberation and competing sound sources. The precedence effect predicts that the inter-relationship of the target (defined in this thesis as the sound source the listener is instructed to lateralize) and the jammer (defined in this thesis as a competing sound source) can affect sound source lateralization. Better lateralization performance is predicted under some circumstances when the target and jammer are correlated then when they are uncorrelated. This thesis is primarily interested in the role that the precedence effect plays in sound source lateralization. In this section, the background inspiring this thesis is presented.

Mills (1958) showed that in anechoic environments without competing jammers normal-hearing listeners can perform sound source localization tasks quite well: just noticeable differences of less than two degrees are not uncommon in anechoic conditions. Good and Gilkey (1996) and Lorenzi et al. (1999) studied the effect of noise on sound source localization. Their results showed that at signal-to-noise ratios (SNRs) less than -10 dB, sound source localization performance is seriously degraded. In studies of sound localization in reverberation, Hartmann (1983), Giguere and Abel (1993) and Rakerd and Hartmann (1985) obtained results similar to those in studies of sound localization in noise, namely degraded performance. Although much previous work uses the standard notation of SNR, this thesis is primarily concerned with target-jammer interactions and will therefore refer to the more descriptive target-to-jammer ratio (TJR).

The two major cues used for lateralization are interaural time differences (ITDs) and interaural level differences (ILDs). Wightman and Kistler (1992) studied the effect of conflicting ILD and ITD cues in terms of sound source localization performance in headphone studies. The stimuli were processed with head related transfer functions (HRTFs) to create virtual images and

the ILDs and the ITDs were then manipulated. For sounds containing low frequencies, horizontal plane localization was dominated by the ITD cue. In fact horizontal plane localization, for wide-band sound stimuli with only an ITD cue was quite good, almost the same as anechoic performance. It seems that wide-band stimuli can be accurately lateralized with only an ITD cue.

Kuhn (1977) measured ITD values, at the ears of a mannequin, for a given a source angle and presented expressions for calculating ITDs. The relationship between ITD and angle is a function of many variables. The major contributing variables are head size and the frequency of the source. Lower frequency sound waves have larger ITDs compared to sound waves of higher frequencies. The ITD of a high frequency (above approximately 3 kHz) source can be calculated with Eq. 1 for a given source angle, θ, and head size.

$$\text{ITD} = \frac{\text{Head Radius}}{\text{Speed of Sound}} * (\sin(\theta) + \theta) \tag{1}$$

It should be noted that Eq. 1 is a nonlinear function. For a 1-degree change in angle, from 0 degrees, there is a change in ITD of 8.7 μsec, whereas for the 1-degree change in angle from 90 degrees, there is a change in ITD of only 4.4 μsec. Since ITD is the major cue introduced to the stimuli used in this thesis, the sound sources that were presented during the psychophysical experiments have equal ITD spacing, as opposed to the more typical equal angle spacing used by Hartmann (1983) and others.

Vedula (2000) and Martin (1993) explored computational methods for locating sound sources in anechoic conditions without jammers. Although numerous methods of ITD parameter extraction were examined, the cross-correlation method adequately calculates the sound source location in the horizontal plane. As the level of noise and/or reverberation is increased, the cross-correlation method is expected to begin to fail. This thesis explores how well normal hearing

listeners perform sound source lateralization tasks for conditions in which these simple cross-correlation methods are expected to fail.

The precedence effect is of particular interest in this study because it could enhance performance of sound source lateralization beyond a simple cross-correlation method. The precedence effect is a well known perceptual phenomena (Litovsky et al. 1999) involving stimuli that contain lead and lag sound components. The precedence effect is most clearly observed when the lag sound is simply a time-delayed version of the lead sound, but where the lead and lag sounds come from different locations. For clicks and brief noise bursts, if the lead-lag delay is on the order of 5 to 10 msec, one sound source is heard, and the perceived direction is close to that of the lead (i.e., the lead component has a larger influence on the direction.) If the lead-lag separation is long, say 30 msec, two sound sources are heard coming from directions corresponding to the lead and lag directions. If the lead-lag separation is very short, say 1 msec, only one sound source is heard, coming from a combination of the lead and the lag directions. It is generally hypothesized that the precedence effect could lead to increased performance of a subject's ability to lateralize sound sources containing only ITD cues under circumstances when a simple cross-correlation method fails to detect the correct source location.

This thesis helps understand the role of the precedence effect in sound source lateralization. For long duration stimuli made up of target and jammer components, the relationship between lateralization judgments and the ITD of the target component is explored for various jammer conditions. In some conditions the target and jammer have a lead-lag relationship while in other conditions the target and jammer are uncorrelated. This thesis compares lateralization performance of jammer conditions with common cross-correlation functions and different auto-correlation functions.

One model (Zurek, 1987) of the precedence effect postulates periods of attenuated lateralization influence following stimulus onsets. For ongoing stimuli, this period of attenuated is achieved by multiplying the stimulus with an attenuating "precedence" window whenever a "peak" in the stimulus occurs. This windowing of the stimulus decreases the importance of future "peaks" that occur within the window's duration. For stimuli containing correlated target and jammer components, as reflected by peaks in the auto-correlation function, the application of a "precedence" window theoretically allows for improvements in lateralization performance. When the target and jammer components are not correlated there is no theoretical improvement in lateralization when a "precedence" window is applied to the sound signal.

The target-jammer time delay can play an important role in the precedence effect. When the target and jammer are uncorrelated the precedent effect does not predict improvements in lateralization performance regardless of the jammer ITD. When the target and jammer are correlated, the Zurek (1987) model of the precedence effect can predict varying levels of improvements as a function of jammer ITD. This is due to the temporal pattern of the precedence window. Longer target-jammer delays result in smaller attenuation of the importance of the future "peaks." Since target-jammer delay is a function of jammer ITD, differences in lateralization performance are predicted for precedent-like stimuli as a function of jammer ITD.

Both psychoacoustic listening experiments and modeling were conducted to assess the role of the precedence effect in lateralization. In Sec. II, the methods and procedures used for the listening experiments and modeling are presented. Section III contains the results, and in Sec. IV there is a discussion of the results. Finally, conclusions about the results are presented along with some suggestions for future work.

# II. Methods

## A. Subjects and Apparatus

Four male subjects, S1, S2, S3, and S5 and one female subject, S4, participate in the experiments. Subjects S1, S3, S4 and S5 were paid to participate; subject S2 was the author. Listeners ranged from 20 to 26 years in age. All subjects reported no hearing impairments and audiometric testing showed normal hearing. (Pure tone thresholds were no poorer that 15 dB HL between 250 Hz and 8 kHz.) Subjects had varying levels of experience with auditory listening experiments and lateralization tasks. Testing was conducted in a sound proof chamber in the Hearing Research Center of Boston University. Stimuli were presented through high quality headphones (AKG MODEL K240) to the listeners. During experiments subjects sat in front of a computer monitor and a mouse pointer method was used to record lateralization responses.

## B. Impulse Responses

In an attempt to understand the role of the precedence effect in sound source lateralization, stimuli with two major components (target and jammer) were explored. Listeners were instructed to lateralize the target component. The properties of the jammer component were controlled so that in some conditions the target-jammer relationship was of a precedence nature while in other conditions the target-jammer relationship was uncorrelated. A description of the properties of the target-alone and jammer-alone impulse responses (IRs) follows this section. The properties of the combined stimulus containing both the target and jammer components are presented in Sec. II.C.

### i. Target

The target IRs were created from an anechoic HRTF. Specifically, the HRTF was from a measurement on the left ear of the Knowles Electronics Mannequin for Acoustics Research

(KEMAR) for a source located directly in front of the head, 0 degrees [obtained from Gardner and Martin (1994)]. The target IRs varied only in ITD corresponding to changes in target direction. As the target angle was varied, the left and right target IRs were simply delayed versions of each other. The same target IRs were used throughout all conditions and experiments. Figure 1 shows the frequency magnitude spectrum of the target IRs.

### ii. Jammer

There were five different jammer conditions, identified as A, A', B, B' and C. In condition A, the jammer IRs were simply a 5-msec delayed version of the IRs for the 0 μsec ITD target. (This 0 μsec ITD corresponds with a 0-degree azimuth jammer source; a source directly in front of the listener.) Similarly in condition A', the jammer IRs were a 5-msec delayed version of the target IRs with an ITD of 643 μsec. (This 643 μsec ITD corresponds with a 90-degree azimuth source; a source located at the left ear of the listener.) In conditions B, B' and C, the jammer IRs were 1-second tokens of noise, with the onset of the noise delayed by 5 msec from the start of the jammer IRs. In condition B, the jammer had an ITD of 0 μsec, whereas in condition B' the jammer had an ITD of 643 μsec. In condition C, the jammer was statistically independent between the left and right IRs (independently chosen phase spectra.) For all the jammer conditions, the jammer IRs had the same magnitude frequency spectrum as the target IR (shown in Fig. 1) but the phase spectra was different for each condition. Table 1 presents the general properties of the jammer IRs.

Table 1 Properties of the jammer IRs for the 6 different jammer conditions.

| Condition | ITD | Duration | Near Ear Latency | General Jammer Properties |
|---|---|---|---|---|
| A | 0 μsec | 3 msec | 5 msec | Reflection from 0 degrees |
| A' | 643 μsec | 3 msec | 5 msec | Reflection from 90 degrees |
| B | 0 μsec | 1 sec | N/A | Independent noise from 0 degrees |
| B' | 643 μsec | 1 sec | N/A | Independent noise from 90 degrees |
| C | N/A | 1 sec | N/A | Interaurally uncorrelated noise |

Note. Figure 2 shows the actual IRs. No ITD is imposed on the jammer for stimulus conditions C. The latencies stated correspond to the latency of the near ear only: There is an additional latency of 643 μsec for the far ear in conditions A'.

## C. Stimuli

In general the sound stimulus was a 500-msec noise burst with a linear rise/decay time of 25 msec. Specifically, the convolution of a 5-second token of white noise with the target IRs for an ITD generated the target waveform for that ITD. Similarly, the convolution of the same 5-second token of white noise with the jammer IRs for the specific condition generated the jammer waveform, which was then scaled to obtain the desired TJR. On every trial a 500-msec window, with a rise/decay time of 25 msec, was multiplied with the target and jammer waveforms to create the target and jammer stimuli. The location of this window was randomly chosen from trial to trial, and was chosen to avoid both onset and offset effects by avoiding window locations that included the first or last second of the target and jammer waveforms. The target and jammer signals were then summed to create the sound stimuli. A personal computer generated the sound stimuli, which were presented through Tucker Davis Technologies digital-to-analog converters with a sampling rate of 44.1 kHz. The target signal was presented at a sound pressure level of 65 dB SPL for all conditions and the jammer signal level varied according to the TJR.

For a given sound stimulus (fixed target ITD, TJR and jammer condition), one can construct transfer functions and overall stimulus IRs, using the fact that the convolution,

windowing, scaling and summation operations are all linear. Figure 2 shows sample IRs of this type for a -100 μsec ITD target and a TJR of 0 dB for conditions A and A' and -10 dB for conditions B, B' and C. In each IR of conditions A and A', two prominent transients appear separated by about 5 msec. The first of these transients corresponds with the target and the second to the jammer. Jammer conditions B, B' and C show only one major transient followed by noise. Only the first 10 msec of the IRs are shown, although the noise in the IRs of jammer conditions B, B' and C has a duration of 1 second. Figure 3 shows the cross-correlation functions of IRs similar to those shown in Fig. 2 for various TJRs. In all cases the target has an ITD of –100 μsec. Each column of Fig. 3 presents the cross-correlation functions of all the different jammer conditions for a fixed TJR.

Examining the cross-correlations of condition A, in the top row of Fig. 3, there are three clusters of activity. Two of the clusters are side lobes, with delays near ±5 msec. Since these clusters are outside the physiologically relevant range of ±1 msec, they are ignored in the modeling presented in this thesis. For a TJR of 0 dB, there are two major peaks in the central cluster within the range of ±1 msec. The peak in the central cluster at –100 μsec corresponds to a lag in which the left and right *targets* are highly correlated. Similarly, the peak in the central cluster at 0 μsec corresponds to a lag in which the left and right *jammers* are highly correlated. The magnitude of the peak associated with the target is not a function of TJR, whereas the magnitude of the peak associated with the jammer is a function of TJR. The reason the peak associated with the target is not seen in the –10 dB TJR is due to the scale of the figure. There is no "noise" outside the three clusters of activity. Results of the cross-correlation for condition A' are shown in the second row of Fig. 3. The cross-correlation functions for condition A' are similar to those of conditions A. The major difference is a shift in the peak locations associated

with the jammer, due to the change in the jammer ITD from 0 μsec to 643 μsec. Again there is no noise outside the three major clusters of activity.

Conditions B and B', which correspond to jammer IRs that are spread out in time, are similar to each other but are substantially different from A and A' in that the side clusters are missing, additionally there is "noise" outside the central cluster. For a TJR of 0 dB, conditions B and B' have the same two large peaks corresponding to the target and jammer ITDs. Again the magnitude of the peak associated with the target is not a function of TJR, whereas the magnitude of the peak associated with the jammer is a function of TJR. The level of the "noise" outside the central cluster of activity is inversely proportional to the TJR.

For Condition C, there is only a single cluster of activity, which in turn contains only one peak. This peak is a result of the correlation between the left and right target IRs. The magnitude of this peak is not a function of TJR (The peak is not noticeable for the –10 dB TJR due to the scale.) Similar to conditions B and B', there is "noise" in the cross-correlation function outside the central cluster of activity, again the level of this "noise" is a function of TJR.

## D. Experimental Procedure

Subjects were first trained with a lateralization procedure for the target-in-quiet condition, to familiarize them with the stimulus and response method. After the subjects became adequately familiar with the target-in-quiet conditions they began running the different jammer conditions. For each jammer condition the subjects would first undergo a 1-interval 2-alternative identification procedure. After completing the identification procedure the subjects would begin lateralization procedures for the different TJRs and jammer conditions. The specifics of how the subjects were run and the procedures for the different experiments follows below.

### i. Experiment 1: Lateralization-in-quiet

Experiment 1 used the target-only stimuli exclusively. On each trial, the 500-msec target stimulus ITD was randomly chosen from 13 equally spaced ITDs between –600 and +600 μsec, corresponding to locations spanning the frontal hemifield. The subjects were asked to listen to target stimuli and locate the target stimulus by clicking a mouse at the appropriate left-right position along the image of a bar (hash marks were used to divide the bar into quarters.) Before the responses were recorded subjects were instructed to familiarize themselves with the stimulus set, especially the perceptual extremes, in terms of left/right, of the target. After familiarization, subjects underwent 100-trial block test procedures. Feedback was not given during experiment 1.

### ii. Experiment 2: Identification

Experiment 2 was a 1-interval 2-alternative forced choice identification experiment. On each trial, the target stimulus ITD was randomly chosen to be either ±300 μsec. (approximately ±35 degrees) The experiment was carried out for stimulus conditions A, A', B, B' and C. The jammer stimuli were consistent with the stimulus condition being tested and the TJR was adaptively chosen based upon a 2-down-1-up adaptive paradigm.

Subjects were asked to listen to target stimuli and to identify the target stimulus as either left or right through a dialog box and a mouse "click". Before measurements were taken, subjects listened to targets, from the set of target ITDs being tested, and jammers, for the conditions being tested and with randomly chosen levels; they were instructed to familiarize themselves with the left/right position of the two target stimuli as well as the jammer stimuli. After familiarization responses were recorded for a total of 10 reversals. The step size for TJR was 2 dB initially (for the first 4 reversals), and then the step size was reduced to 1 dB for the remainder of the

procedure. The mean of the last six reversals was taken as the subject's identification threshold for the conditions being tested.

For each trial a random 500-msec sample of the jammer stimuli was played. The jammer was followed by a 250-msec interval of quiet, which in turn was followed by the jammer stimuli, again, summed with the corresponding target stimuli with a randomly chosen ±300 μsec ITD. The subjects were told in the instructions for the experiment that the target stimulus was played during the second burst of noise. Implicit feedback was not given during experiment 2.

### iii. Experiment 3: Lateralization

Experiment 3 was also performed for stimulus conditions A, A', B, B' and C. On each trial, the target stimulus ITD was randomly chosen from 13 equally spaced ITDs between –600 and +600 μsec, corresponding to locations spanning the frontal hemifield. The jammer stimuli were consistent with the stimulus condition being tested. The jammer level was determined from the TJR being tested. The subjects were asked to listen to target stimuli and locate the target stimulus by clicking a mouse at the appropriate left-right position along the image of a bar (hash marks were used to divide the bar into quarters.) Before responses were recorded, subjects were instructed to familiarize themselves with the stimulus set, especially the perceptual extremes, in terms of left/right, of the target, by playing token samples of the target alone stimulus. Additionally they were asked to familiarize themselves with jammer stimuli in the same manner. After familiarization, subjects underwent 100-trial block test procedures. For each trial a random 500-msec sample of the jammer stimuli was played. The jammer was followed by a 250-msec interval of quiet, which in turn was followed by the jammer stimuli, again, summed with the corresponding target stimuli with a randomly chosen ITD. The subjects were told in the

instructions for the experiment that the target stimulus was played during the second burst of noise. Feedback was not given during experiment 3.

### iv. Procedure

A jammer condition sequence was created for each subject, randomly determining the order that the jammer conditions A, A', B, B' and C would be run in. Eight TJRs were examined, (+12 dB, +8 dB, +4 dB, +0 dB, -4 dB, -8 dB, -12 dB and -16 dB.) For each subject and jammer condition a TJR sequence was created containing the eight TJR in a pseudo random order. The first TJR of all the TJR sequences was always +12 dB. This was done to allow the subject additional time to familiarize themselves with the experiment.

Each subject first underwent a series of lateralization-in-quiet runs. These measurements were concluded when the subjects had completed a minimum of six 100-trial blocks and the difference in the standard deviations of the marked location (averaged over the actual ITDs) was less than 100 μsec between the last two successive 100-trial blocks.

The subject then began testing of the different stimulus conditions. First the subjects completed an identification run for the jammer condition (determined by the subject's stimulus condition sequence). The subject then completed seven, 100-trial block, lateralization runs. The jammer condition was held constant throughout the seven, 100-trial blocks. The TJR was varied between 100-trial blocks according to the TJR sequence for the subject and stimulus condition. Before the third 100-trial block, of the series of seven lateralization runs, the subject underwent another 100-trial block lateralization-in-quiet run. After completing all the stimulus conditions, if a subject's schedule permitted, additional 100-trial blocks were run of various jammer conditions and TJRs.

## E. Model

Two relatively simple cross-correlation models were designed. Model 1 was a *dumb* model that did not make use of the known location of the jammer, and model 2 was a *smart* model that did make use of the known location. Both models are wide-band, long-time cross-correlation models. The left and right ear sound stimulus was cross-correlated for lags between ±1 msec for each stimulus. Model 1 used the ITD of the peak in the cross-correlation function as the predicted ITD. Model 2 used information about the jammer condition being tested; it calculated a cross-correlation for a single jammer-only stimulus and used this as a template throughout a testing block. (Possibly similar to what subjects do during the familiarization stage of the psychoacoustic experiments.) For each stimulus the smart model subtracts this jammer only cross-correlation template from the target plus jammer cross-correlation function. The ITD of the peak in this resulting function was taken as the predicted ITD.

Both models were run for a 500-trial block for jammer conditions A, A', B, B' and C with TJRs between +10 dB and –20 dB, in 2 dB steps, for each jammer conditions. On each trial the target stimulus ITD was randomly chosen from 13 equally spaced ITDs between –600 and +600 μsec. The jammer stimulus was consistent with the jammer conditions and TJR being tested.

## F. Data Analysis

One of the first steps in analyzing the data involves transforming the subject's response on the bar, with the mouse pointer method, to an ITD. This is accomplished by arbitrarily calling the left extreme of the bar –1, the right extreme of the bar +1. The detected angle for a given lateralization response can then simply be calculated with the arcsin function. From this detected

angle combined with Eq. 1 the detected ITD for a given lateralization response can be determined.

There are three major divisions of the analysis of the data. One is based on the least-mean-squared best-fitting straight line (LMS-line) relating the actual ITD of the target and the lateralization responses, i.e. the detected ITD, of the subject. This first division includes the intercept of the LMS-line, the root mean squared (RMS) difference between the detected ITDs and the LMS-line, the slope of the LMS-line, and the square of the correlation coefficient ($R^2$) of the detected ITD and actual ITD results. As discussed below, the $R^2$ statistic has the best properties for exploring the differences for both inter-subject and inter-condition differences. The second division of the analysis of the data is based upon the TJR dependence of a summary statistic. This second division of the analysis is used to determine both lateralization and identification thresholds for each subject and condition. The final division of the data analysis involves determining the statistical significance of these thresholds.

The intercept of the LMS-line is an indicator of the perceived location of a 0 μsec ITD target stimulus. The intercept value relies heavily upon the "guessing" strategy employed by the subjects when they can no longer correctly determine the target ITD; therefore it is not a good summary statistic. The RMS difference between the detected ITDs and the LMS-line also relies heavily upon the "guessing" strategy and is not a good summary statistic. However in experiment 1, the RMS difference does help to show that the subjects are able to correctly and repeatable lateralize the target stimuli with the mouse pointing response method. The slope of the LMS-line is an indicator of how much of the response bar was used by the subject. Although subjects were instructed to use as much of the response bar as possible, different subjects used different amounts of the bar, therefore the slope is not a good summary statistic. The square of

the correlation coefficient is a measure of the differences in the linearity of the response mapping from actual ITD to detected ITD. $R^2$ seems to be a good summary statistic. An even better summary statistic in terms of highlighting the differences between the different jammer conditions seems to be a normalized version of $R^2$.

The normalized $R^2$ for a the detected ITD responses and the actual ITDs for a given subject, jammer condition and TJR is defined simply as the $R^2$ for that subject, jammer condition and TJR divided by the $R^2$ for that subject in quiet. Equation 2 defines the normalized $R^2$.

$$Normalized \quad R^2_{condition} = \min\left[1, \quad \frac{R^2_{condition}}{R^2_{quiet}}\right] \tag{2}$$

The second division of the data analysis is based upon the results of the cumulative Gaussian function that best fits the normalized $R^2$ versus TJR data for each subject and condition. The cumulative Gaussian function is taken as the psychometric function of the subject ability to lateralize the target in the jammer conditions. The normalized $R^2$ is limited to a maximum value of 1 since a cumulative Gaussian function is also limited to a maximum value of 1. The lateralization threshold TJR is defined as the TJR for which the psychometric function has a normalized $R^2$ value of 0.5.

The final division of the data analysis was a single factor ANOVA to determine the significance of the differences between the results for the different conditions. Numerous pairings were made. The differences between conditions A and A' were examined, as well as the differences between A and A' and B and B' to name just a few of the comparisons examined.

# III. Results

## A. Experiment 1: Lateralization-in-quiet

The purpose of Experiment 1 was to determine whether the subjects were able to correctly lateralize the target stimuli in the quiet condition. Figure 4 presents the confusion matrices (detected ITDs versus actual ITDs) of each subject, for the target-in-quiet condition. The bin size for the confusion matrices is 100 μsec, and the area of the symbol is proportional to the number of points in each bin. In general the subjects place targets with different ITDs at different locations along the bar. Additionally for each 100-trial block, the slope, intercept, RMS difference and $R^2$ statistics were calculated. Figure 5 shows, for subject S3, the values of these statistics as a function of 100-trial block number. The statistics appear to be time-invariant after the first few blocks; no steadily increasing or decreasing trends can be seen as a function of block number. This implies that the subjects were given adequate time, a minimum of 6, 100-trial blocks to become familiar with both the target stimulus and response method used during the experiment.

Table 2 presents the mean and the standard error, over 100-trial blocks, of the statistics for each subject. It should be noted that although all the statistics for the subjects are time invariant (Fig. 5) there are differences between subjects. For example, subject S1 has a much higher RMS difference than S5, and subject S5 has a much shallower slope than S2. As mentioned in Sec. II.F, the differences in these statistics arise from the mouse pointing method. For instance the intercept, for each subject, is a function of the perceived center of the bar. The slope is a measure of how much of the bar the subjects used in their responses. In summary, all the subjects were able to correctly locate the target for the target-in-quiet condition. The values of $R^2$ presented in Table 2 are used in the calculations of the normalized $R^2$ in Experiment 3.

Table 2 Average values and the standard errors of the different statistics calculated over all the 100-trial blocks of Experiment 1 for each subject.

| | S1 | S2 | S3 | S4 | S5 |
|---|---|---|---|---|---|
| Slope | 0.72±0.02 | 0.78±0.02 | 0.73±0.03 | 0.57±0.02 | 0.48±0.03 |
| Intercept (μsec) | 15.0±8.1 | 6.7±7.6 | 4.9±4.2 | 17.6±6.8 | 2.9±5.7 |
| $R^2$ | 0.65±0.03 | 0.88±0.02 | 0.89±0.01 | 0.89±0.01 | 0.80±0.02 |
| RMS Difference (μsec) | 196.7±9.8 | 106.9±7.1 | 91.1±2.6 | 76.9±4.6 | 85.5±2.9 |

## B. Experiment 2: Identification

Experiment 2 determined the identification threshold for the different jammer conditions summarized in Table 1. The identification threshold is defined as the lowest target-to-jammer ratio (TJR) for which the subjects were able to correctly identify, at least 70% of the time, which target ITD (either ±300 μsec) was presented in the presence of the jammer. Table 3 presents the identification thresholds of the individual subjects, the threshold averaged across subjects and the standard deviation across subjects, for the five different jammer conditions. Figure 6 shows the data in Table 3 graphically along with model predictions that are discussed later in Sec. III.D.

The standard deviation of the thresholds for a given condition between subjects is relatively modest, on the same order of magnitude as the variance in a subject's threshold from day to day (data not shown). There is, however, a significant 3.2 dB difference (P value of 0.027) between the mean identification thresholds for the precedent-like jammer conditions, jammer conditions A and A', and the corresponding non-precedent-like jammer conditions, jammer conditions B and B'. Due to the nature of the jammer conditions this difference in identification threshold is interpreted as a result of the precedence effect.

Table 3. Identification thresholds for each subject and condition as well as the across subject average threshold and standard deviation.

| Condition | S1 Threshold (dB) | S2 Threshold (dB) | S3 Threshold (dB) | S4 Threshold (dB) | S5 Threshold (dB) | Average Threshold (dB) | Standard Deviation (dB) |
|---|---|---|---|---|---|---|---|
| A | -10.0 | -9.3 | -12.1 | -14.8 | -7.5 | -10.75 | 2.80 |
| A' | -10.7 | -9.7 | -7.2 | -7.8 | -6.8 | -8.4 | 1.68 |
| B | N/A | -6.0 | -9.0 | -6.7 | -6.3 | -7.0 | 1.36 |
| B' | -3.7 | -8.0 | -4.8 | -11.3 | -1.3 | -5.8 | 3.90 |
| C | -5.0 | -3.0 | -6.3 | -3.3 | -4.0 | -4.3 | 1.35 |

## C. Experiment 3: Lateralization

Experiment 3 determined the lateralization threshold for the different jammer conditions. Example confusion matrices of subject S4, for jammer condition B', for the different TJRs, are presented in Fig. 7. For high TJRs the subject has performance similar to the performance in quiet (cf, Fig. 4). As the TJR decreases, performance is less dependent on the target ITD. This change with TJR is not seen in all statistics. (See the appendix for complete data) For example, the RMS difference of the subject is not a function of TJR. Some subjects however exhibit a change in the RMS difference statistic as a function of TJR. For jammer condition B', the intercept of S4 (in Fig. 7) is a function of TJR. Again, this result is consistent neither across subjects nor across conditions. Changes in the intercept and RMS difference statistics depends on how the subject responds when they cannot detect the correct ITD of the target. Some subjects have biases towards the ITD of the jammer, others to the center of the bar. Both the slope and the $R^2$ statistics vary consistently with the TJR, and across conditions. The $R^2$ statistic seems to be the most consistent measure for interpreting the results across conditions. The normalization of the $R^2$ statistic leads to the best statistic for making comparisons across both subjects and conditions.

Figure 8 shows the normalized $R^2$, for subject S2, plotted as a function of TJR for all the jammer conditions. Also shown in Fig. 8 is the best fitting cumulative Gaussian function for each jammer condition. The transition region of the psychometric function seems to be captured well by the shape of the cumulative Gaussian function. Figure 9 shows the best fitting cumulative Gaussian functions for all the subjects and jammer condition B.

The lateralization threshold is defined as the TJR at which the normalized $R^2$ statistic is equal to 0.5. Table 4 contains a tabulation of the lateralization thresholds, which are presented graphically in Fig. 6. Similar to Experiment 2, there is a 2.85 dB difference (Significant with a P value of 0.040) in the thresholds between the precedent-like jammer conditions, conditions A and A' and the corresponding non-precedent-like jammer conditions, conditions B and B'. As explained in Section I, due to the nature of the stimuli the differences are interpreted as a result of the precedence effect.

Table 4 Lateralization thresholds for each subject and jammer condition as well as the across subject average and standard deviation.

| Condition | S1 Threshold (dB) | S2 Threshold (dB) | S3 Threshold (dB) | S4 Threshold (dB) | S5 Threshold (dB) | Average Threshold (dB) | Standard Deviation (dB) |
|---|---|---|---|---|---|---|---|
| A | -9.5 | -9.5 | -4.4 | -13.6 | -8.8 | -9.2 | 3.3 |
| A' | -9.7 | -9.6 | -5.7 | -6.4 | -7.3 | -7.8 | 1.8 |
| B | -7.6 | -6.6 | -5.7 | -5.5 | -0.4 | -5.2 | 2.8 |
| B' | -2.6 | -11.3 | -5.3 | -8.1 | -2.9 | -6.0 | 3.7 |
| C | -2.4 | -4.2 | -5.2 | -2.6 | -4.1 | -3.7 | 1.2 |

## D. Model

Figure 10 shows the confusion matrices for Model 1, the *dumb* model, for condition B' and a variety of TJRs. From these confusion matrices it is apparent that when the model fails to predict the target ITD it predicts the jammer ITD. Figure 11 presents the normalized $R^2$ statistic as a function of TJR for the different jammer conditions along with the best fitting cumulative

Gaussian functions for Model 1. The lateralization thresholds of Model 1 are approximately 0 dB for conditions A, A', B and B' and –13 dB for jammer condition C.

Figure 12 shows the confusion matrices for Model 2, the *smart* model, for condition B' and a variety of TJRs. Model 2 either predicts the target ITD or a random ITD. Figure 13 presents the normalized $R^2$ statistic as a function of TJR for the different jammer conditions and the best fitting cumulative Gaussian function for Model 2. The lateralization thresholds of Model 2 are approximately -13 dB for all jammer conditions.

Figure 6 graphically shows the lateralization threshold predicted by both models as well as the lateralization thresholds determined in experiment 2. This allows for a direct comparison between the human subject and model performances. It is interesting to note that the *smart* model outperforms the human subjects for all conditions.

## IV. Discussion

This thesis shows three major points. The first is the differences between conditions A and A' and B and B'. This difference, as explained earlier, is predicted to be a direct result of the precedence effect. The second major point is found in the differences between conditions A and A' and the similarities between conditions B, B' and C. The similarities in conditions B, B' and C help to show that the differences between conditions A and A' are a results of how the precedence effect is implemented by the auditory system. Finally an overall comparison between the psychophysically measured thresholds of normal-hearing individuals can be compared to the performance of simple cross-correlation models.

In an attempt to illustrate the three major points of this thesis some similarities between the identification and lateralization thresholds need to be explained. Some of these similarities seem to be incidental, while others are a direct result of the stimulus properties. For example the

average identification threshold (4.3 dB) and the average lateralization thresholds (3.7 dB) for jammer condition C are similar only because of the choice of ITD values in experiment 2. For any given conditions the identification threshold should be dependent upon the size of the difference between the left and right target ITDs. (In the case of this work the difference was 600 μsec.) Therefore the similarities of the identification threshold levels, for any jammer condition, with the lateralization threshold levels of either corresponding or non-corresponding jammer conditions simply reflect the choice of the difference in ITD of the left and right targets. Other similarities between the identification thresholds and the lateralization thresholds may be more fundamental. For example, the differences between the mean identification thresholds of conditions A and B (-3.8 dB) can be attributed to the precedence effect.

The lead-lag relationship between the target and jammer in condition A is thought to allow for improved lateralization performance from the precedence effect. The relationship between the target and the jammer in condition B is such that lateralization performance is not improved by the precedence effect. The similarities between the interaural cross-correlations of conditions A and B lead cross-correlation based models to achieve similar lateralization performance in both conditions. A similar explanation can be for the differences in lateralization performance between conditions A' and B'.

Further evidence for the precedence effect playing a role in lateralization performance comes from examining the differences between conditions A and A' and the differences between conditions B and B'. It is argued here that the difference in thresholds between conditions B and B' will be smaller than the difference in the thresholds between conditions A and A'. The stimuli do not rule out the possibility of differences in lateralization or identification performance based upon the jammer ITD. The precedence effect, as described in Sec. I, predicts a difference in the

threshold level as a function of jammer ITD. In conditions A' the far ear stimulus has an additional 643 msec of lag. The decrease in importance of "peaks" in the stimuli associated with the precedence window is time dependent; the decrease in importance decreases with time. With this in mind, the differences in threshold between conditions A and A' should be slightly larger than the differences in threshold between conditions B and B'. The results however do not show this finer point of the precedence effect as clearly as expected.

The difference in thresholds between conditions A and A' was 1.9 dB. The results of the ANOVA analysis however yielded a P value of 0.104; meaning the results are not significant. This does not however tell the entire story. In the view of the author, the results are significant due to the limited number of subjects and the relatively small effect. The author believes that there is in fact a difference in the thresholds for conditions A and A'. This belief is based upon the fact that in both the lateralization and the identification experiments the average thresholds for condition A was higher than the average thresholds for condition A'. (1.6 dB difference for the lateralization experiment and 2.4 dB difference for the identification experiment.) However, the difference between conditions B and B' were not consistent between the lateralization experiment and the identification experiment. It seems the consistencies between conditions A and A' and the inconsistencies between conditions B and B' yield further evidence that there is in fact a slight effect of jammer location for the precedent like stimuli of conditions A and A'.

In terms of the goal of this thesis and the determination of the role the precedence effect in sound source lateralization, the differences between conditions A and A' and B and B' for both the lateralization and identification thresholds suggests approximately a 3 dB role for the precedence effect. Since this work only examined a single 5 msec lead-lag separation, the

importance of the precedence effect for different lead-lag separations is not known, however, it seems that the precedence effects role is significant in the cases examined in this work.

Although it seems the precedence effect has been found to be significant an interesting result is found when a comparison between the performance of normal-hearing listeners and the smart model is made. The smart model outperformed normal hearing listeners by up to 10 dB in some situations. It seems that normal-hearing subjects are not efficient at performing peak extraction from cross-correlation functions. The results of this thesis suggest a more detailed comparison between the performance of simple cross-correlation models and human performance needs to be done. Neither model explored in this thesis takes into account the precedence effect. It would be interesting to examine the performance of a cross-correlation based precedence model in terms of the performance of non-precedence cross-correlation models and normal-hearing listeners. In addition to further modeling efforts, an examination of the role of the precedence effect for other stimuli (clicks, narrow-band noises, ongoing noise burst and multiple jammers) should be carried out.

## V. Conclusions

The role of the precedence effect in both sound source identification and sound source lateralization, for wide-band ongoing targets in the presence of a perfectly correlated, 5 msec delayed, wide-band ongoing jammer of a known location is approximately 3 dB. For all the jammer conditions tested a *smart* cross-correlation model was able to outperform the human subjects by a minimum of 5 dB and up to nearly 10 dB. More jammer conditions (the number of jammer sources, target-jammer delays and spatial separations) need to be investigated to begin to fully understand the role of the precedence effect in sound source lateralization.

## Acknowledgements

Work supported by NIH. (Grant Numbers R01 DC00100 and T32 DC00038)

## Figures

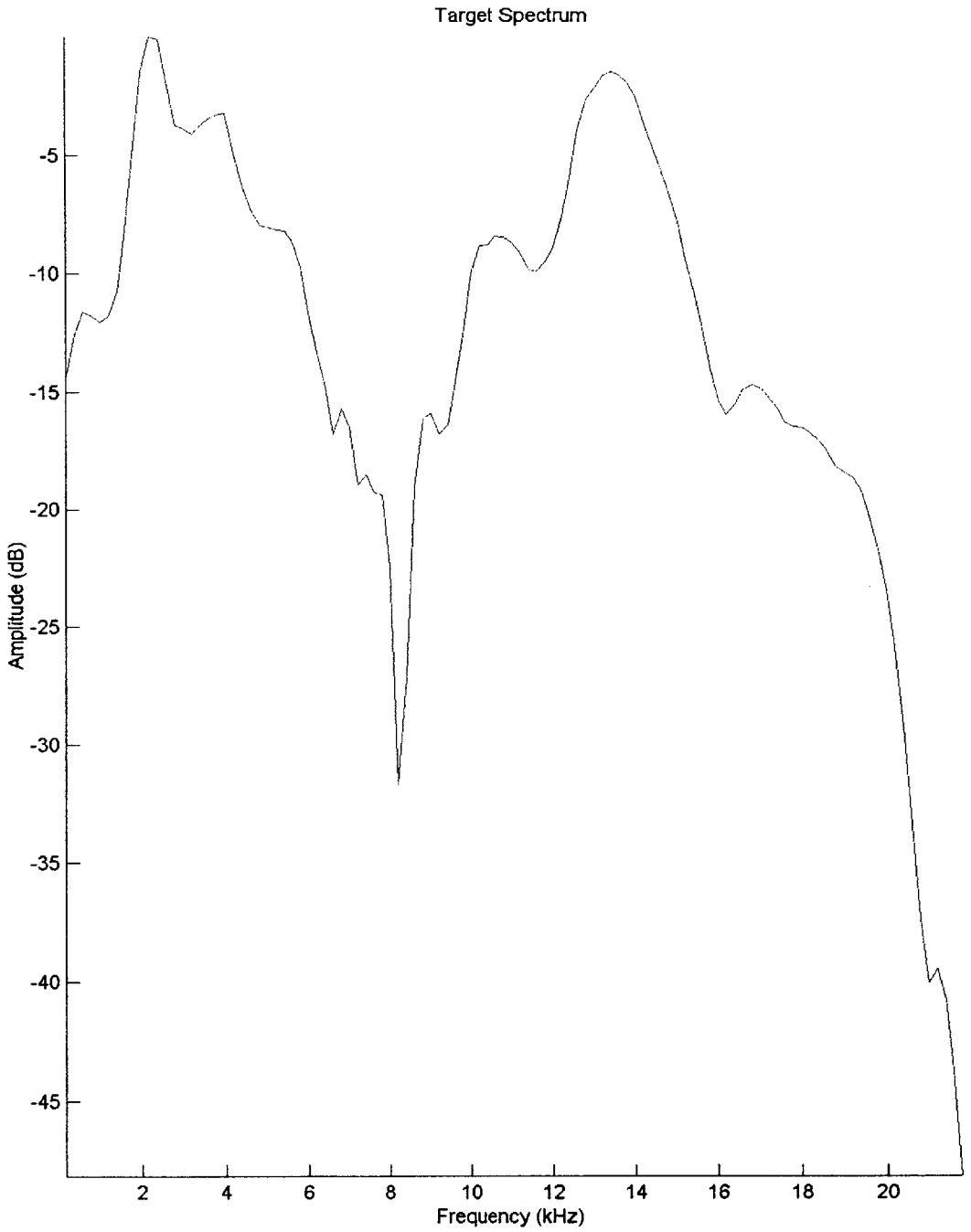


Fig. 1 Common magnitude spectrum of the target and jammer stimuli taken from a 0-degree HRTF measurement made on KEMAR obtained from Gardner and Martin (1994). The amplitude is in dB re the maximum.

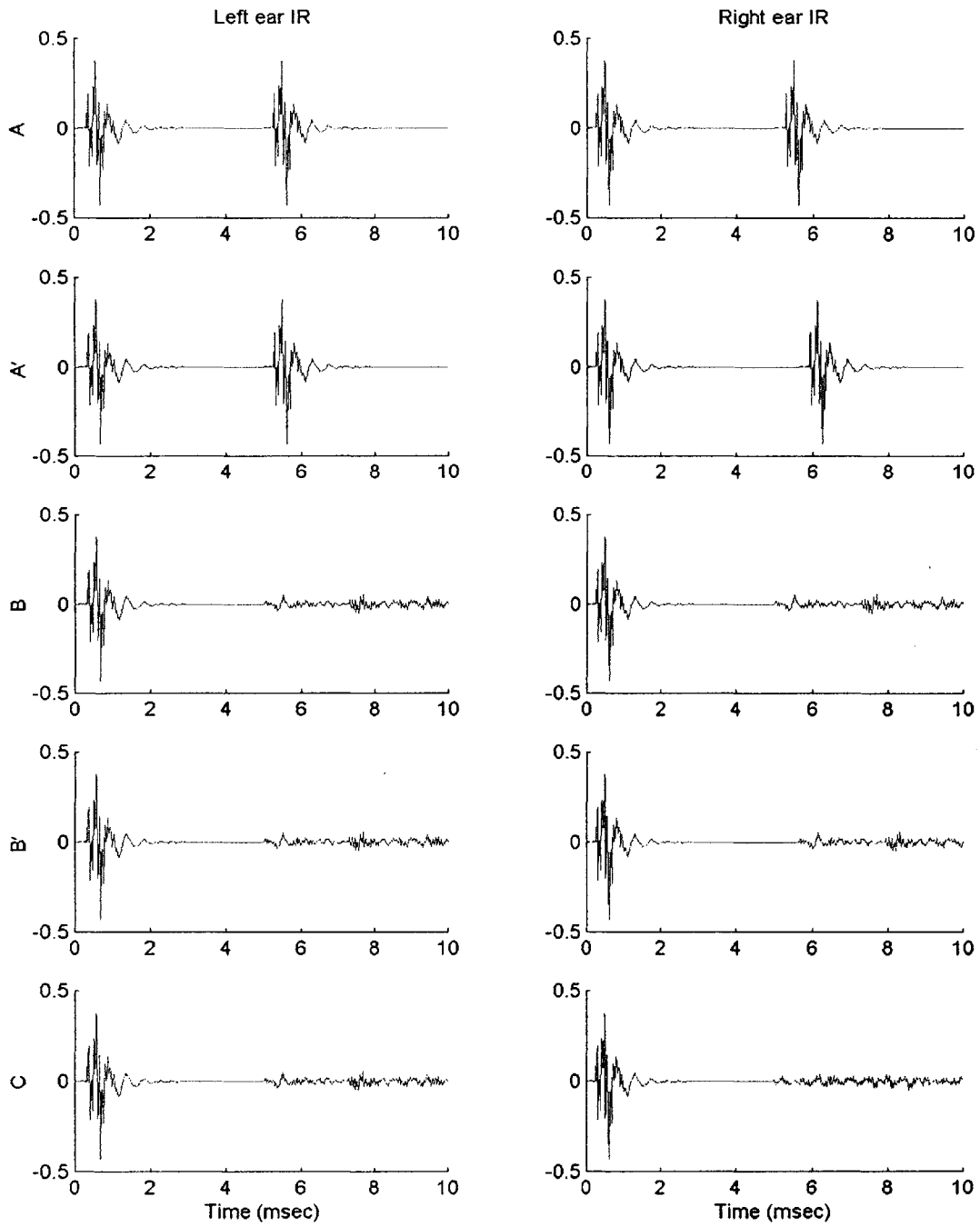


Fig. 2 The first 10 msec of the left and right ear impulse responses (IRs) for example sound stimuli of types corresponding to the different jammer conditions. In all cases the target has an ITD of –100 μsec. The target-to-jammer ratio (TJR) is 0 dB for conditions A, A' and –10 dB for conditions B, B' and C. The amplitude scale is consistent across conditions and ears.

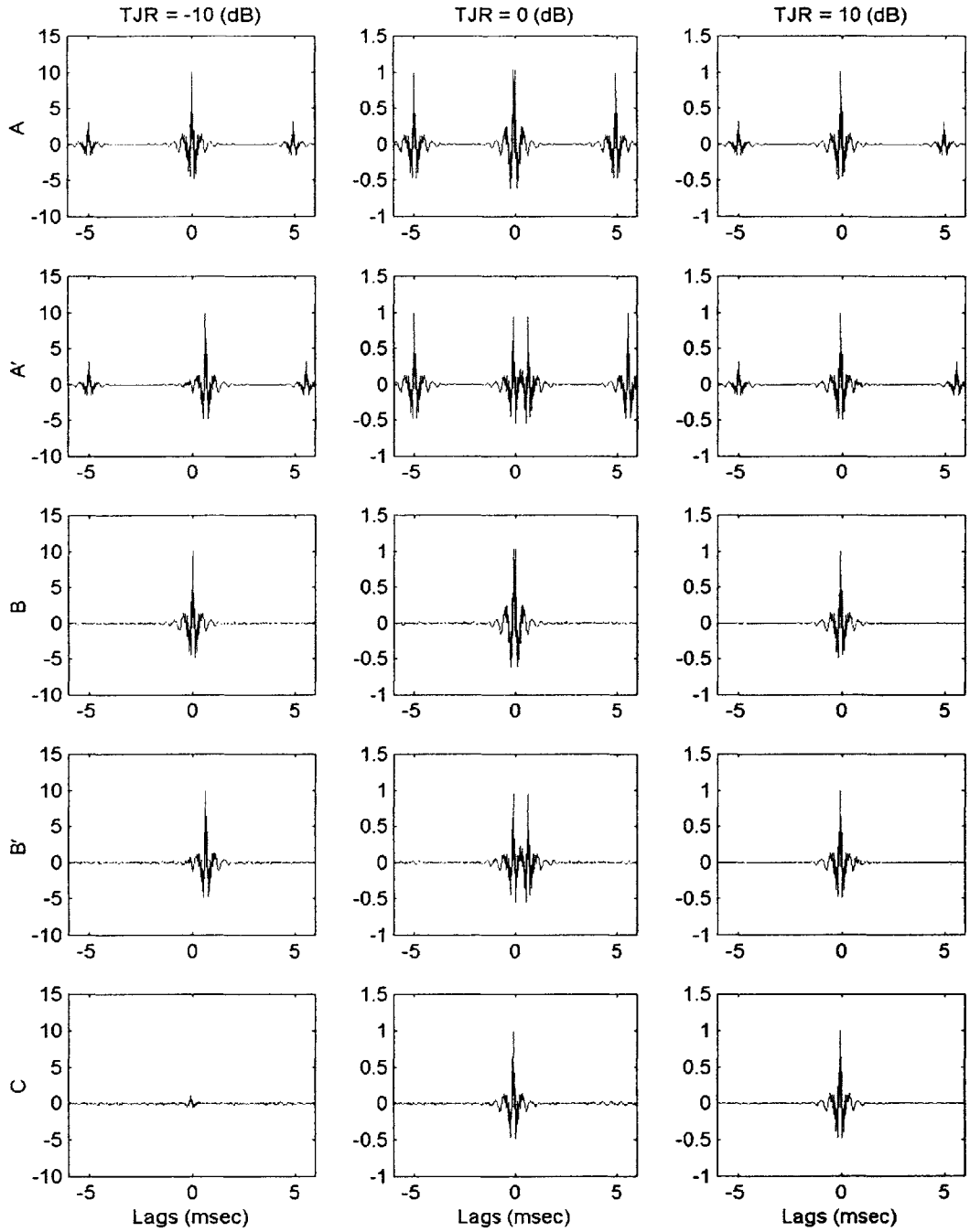


Fig. 3 Example cross-correlation functions of the sound stimuli IRs, similar to those presented in Fig. 2, for the different stimulus conditions over a range of target-to-jammer ratios (TJRs). The target has an ITD of −100 μsec in all cases. Note the change in the scale of the ordinate for the different TJRs. The amplitude scale values are consistent across conditions and TJRs.

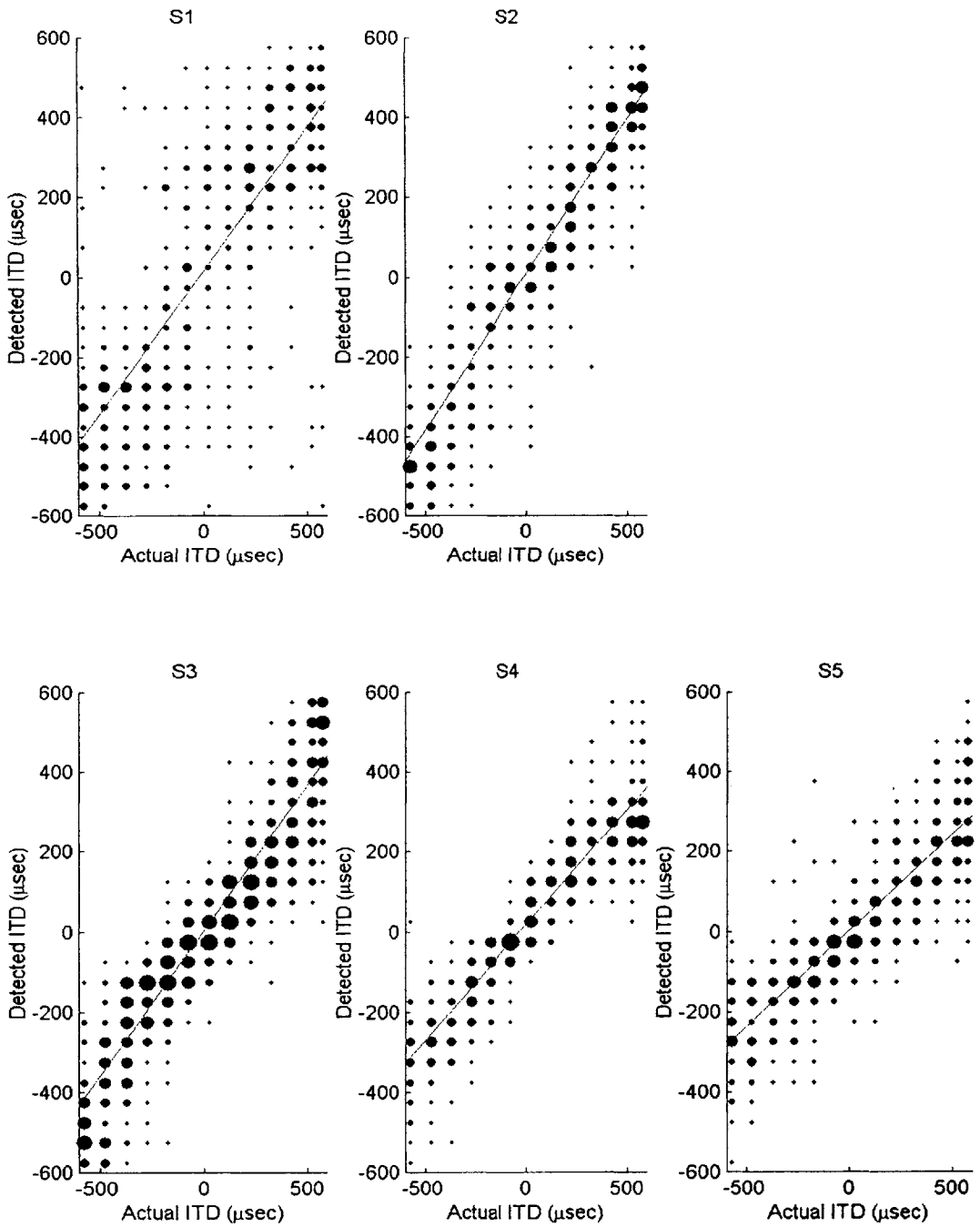


Fig. 4 Confusion matrices for all the subjects in the target-in-quiet condition

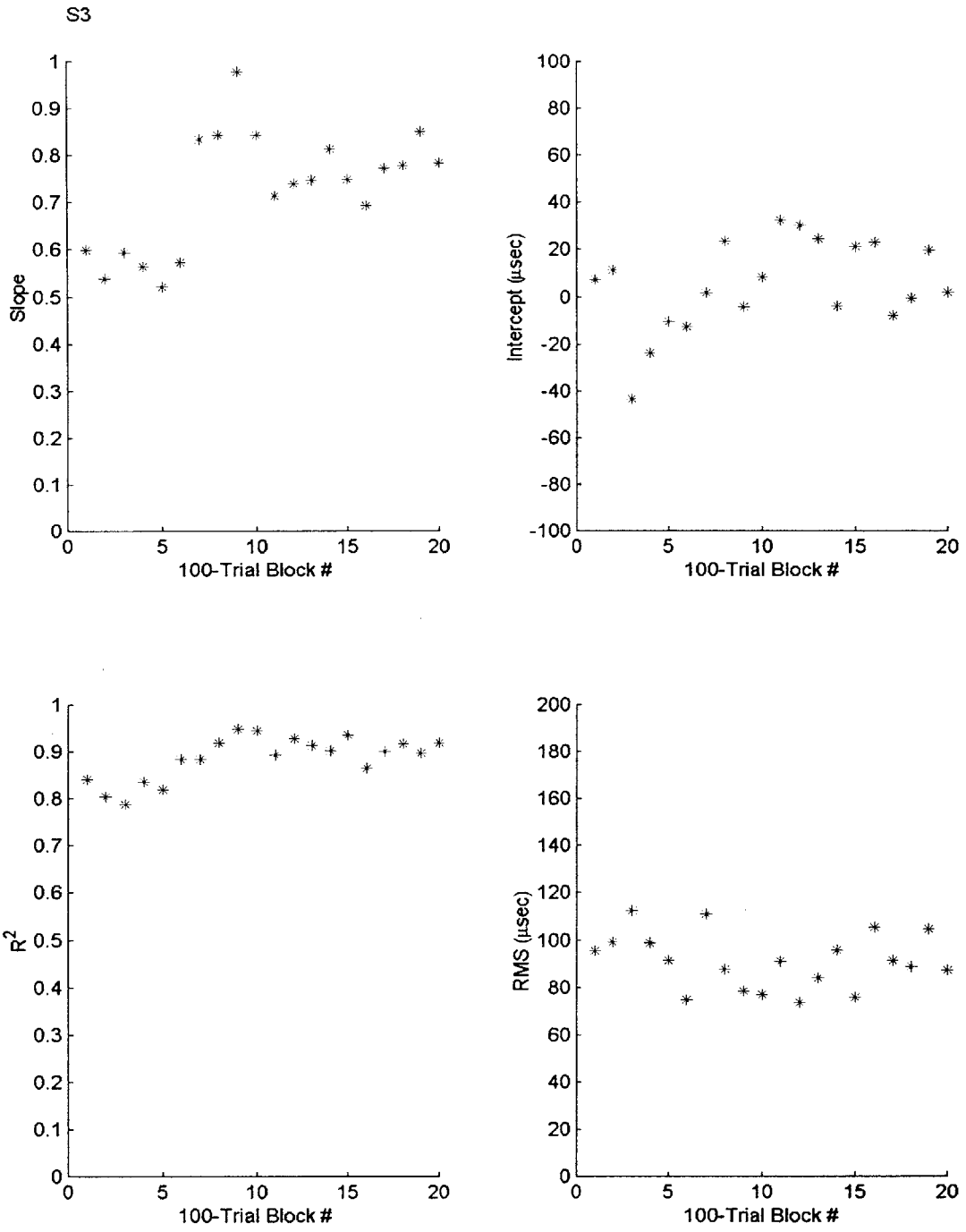


Fig. 5 Statistics summarizing the performance of subject S3 during the target-in-quiet condition as a function of 100-trial block number. The top left panel shows the slope of the LMS-line and the top right panel shows the intercept of the LMS-line. The bottom left panel shows $R^2$ between the detected ITD and the actual ITD. The bottom right panel shows the RMS difference between the detected ITD and the LMS-line.

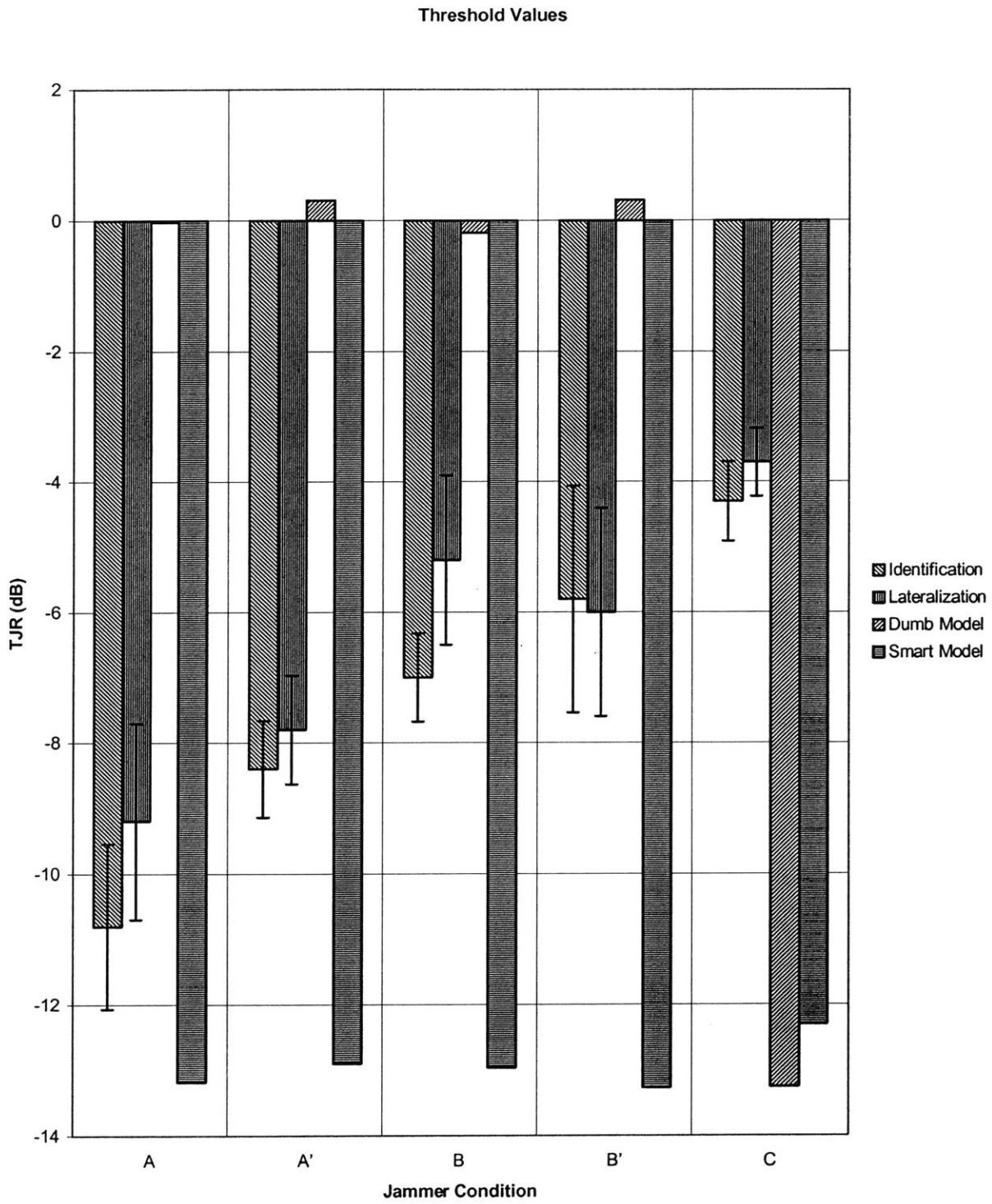


Fig. 6 Threshold values for the different jammer conditions. Across subject average values for the identification and lateralization thresholds, error bars represent the standard error. The *smart* and *dumb* model thresholds are the results of 500-trial lateralization runs.

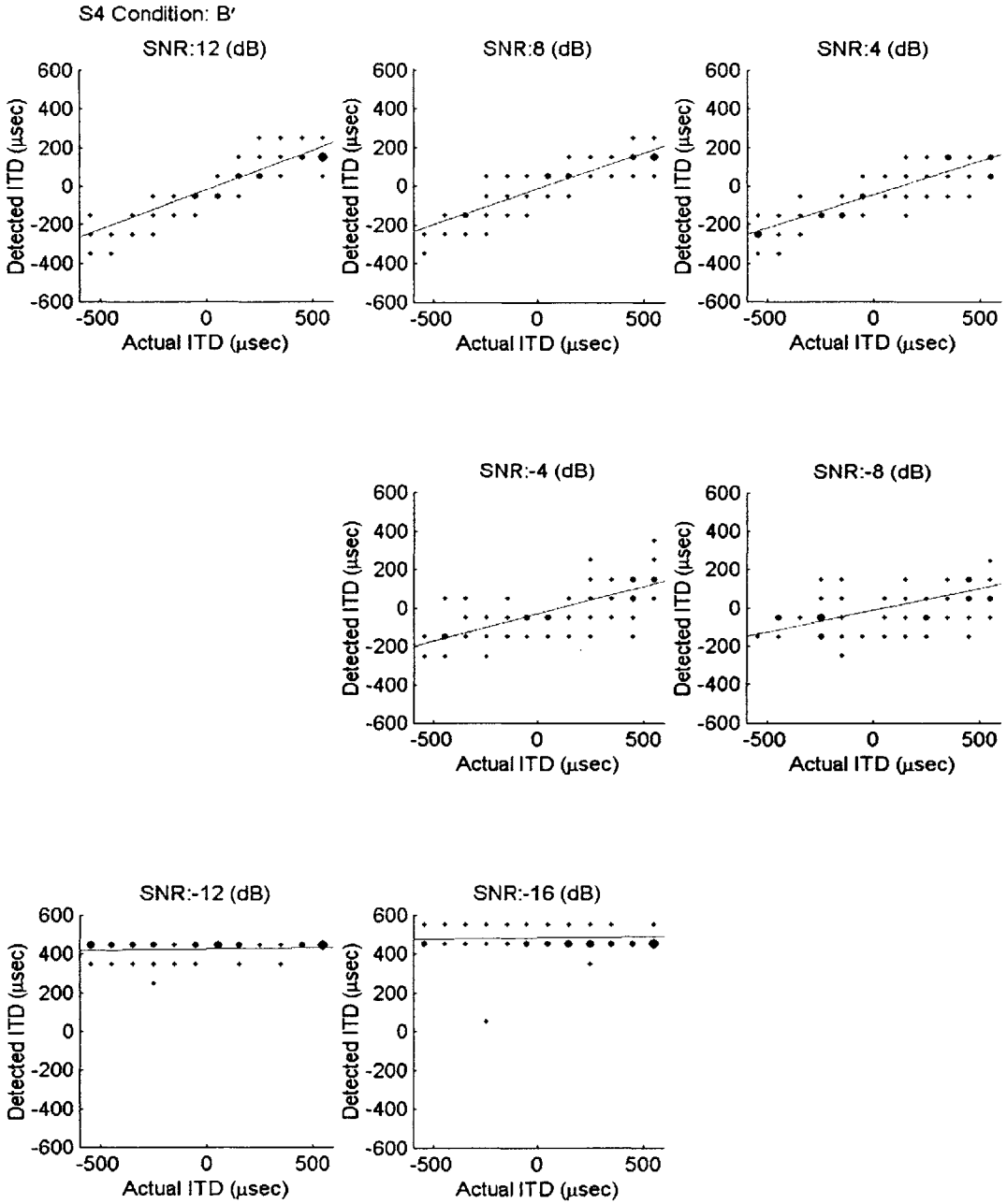


Fig. 7 Confusion matrices of subject S4 for jammer condition B'.

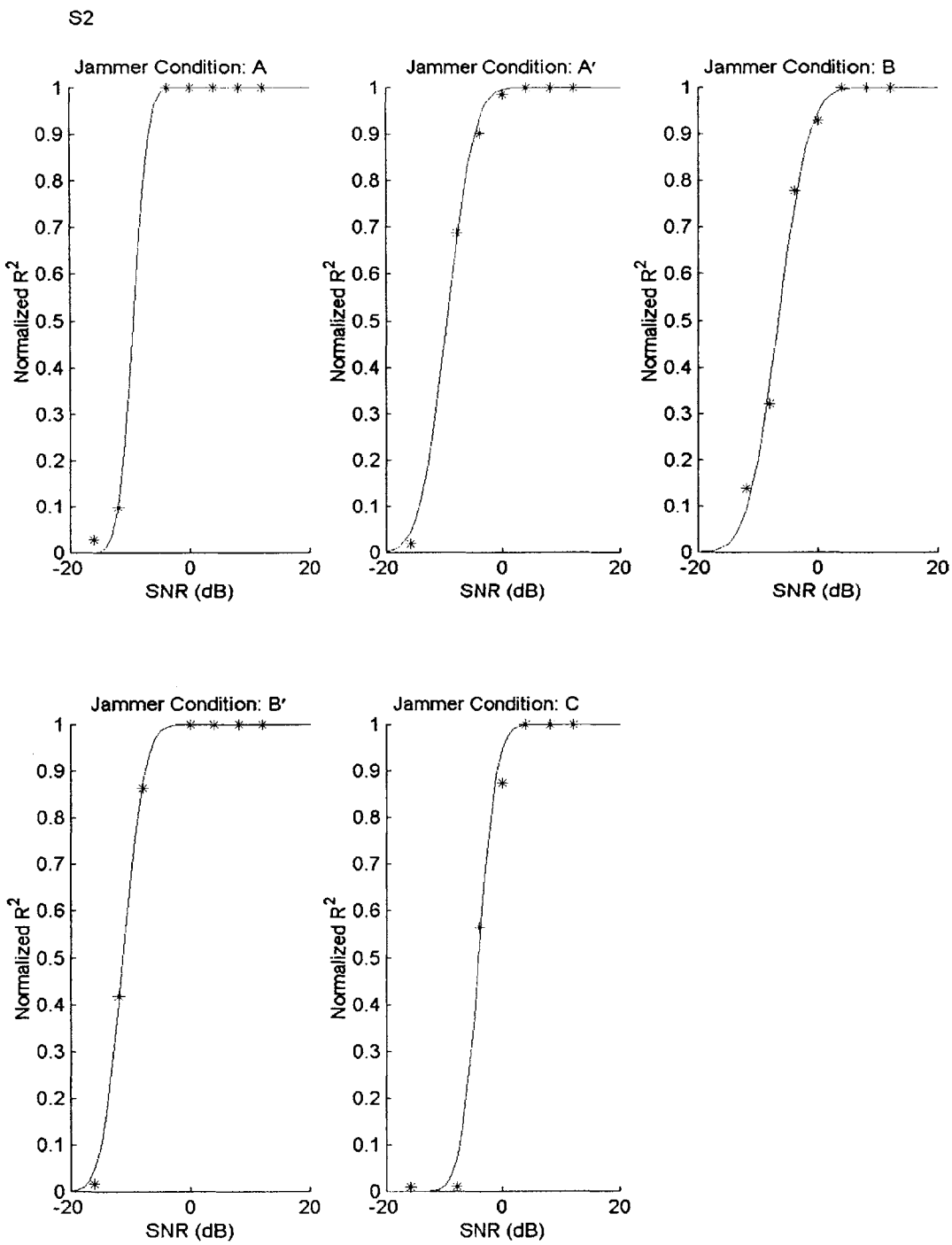


Fig. 8 Normalized $R^2$ of subject S2 for the different TJRs and jammer conditions. The continuous curves are cumulative Gaussian functions fit to the measured $R^2$ values shown.

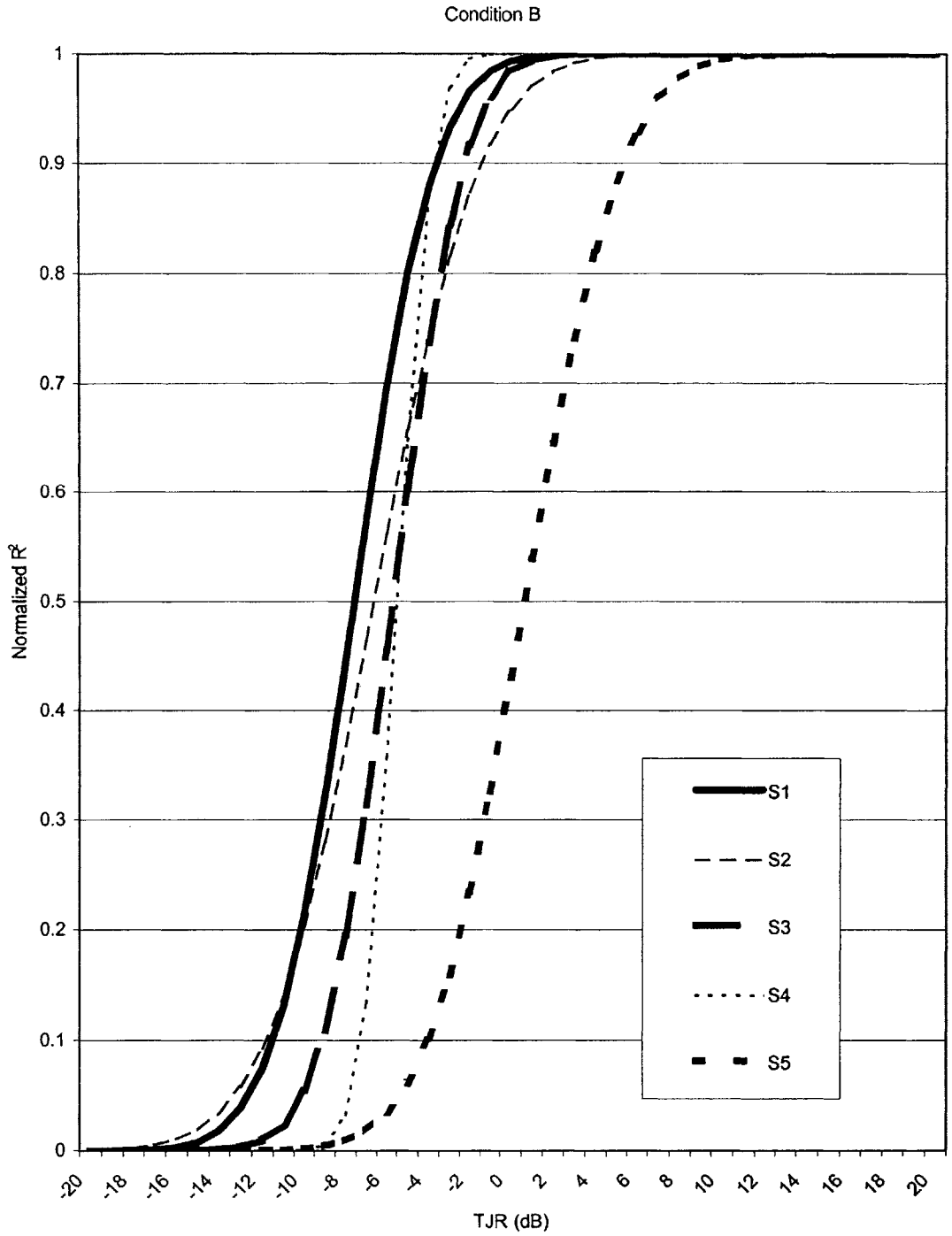


Fig. 9 Cumulative Gaussian functions, for each subject, that best fit normalized $R^2$ data for condition B.

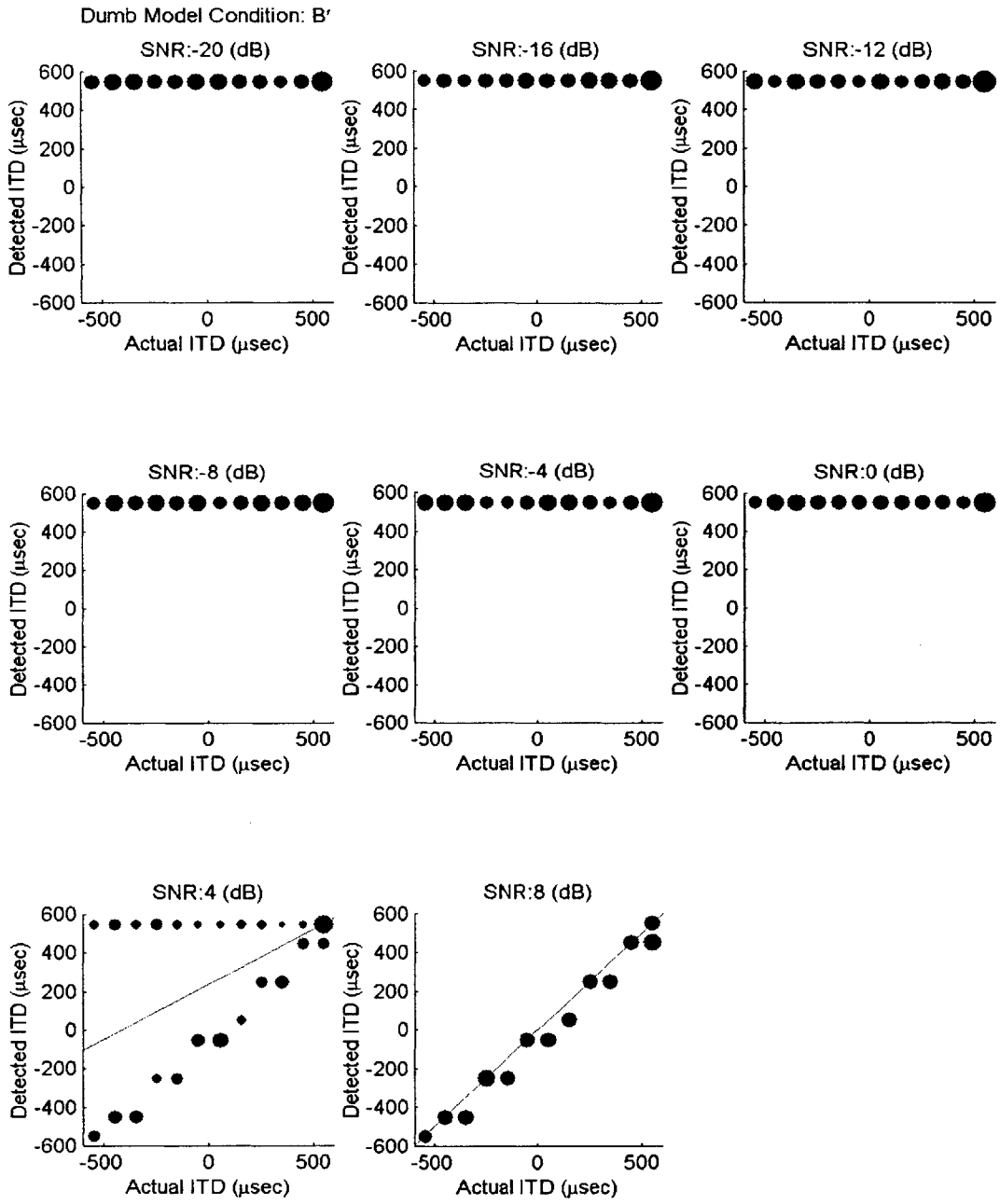


Fig. 10 Confusion matrices of Model 1, the *dumb* cross-correlation model, for jammer condition B' and different TJRs.

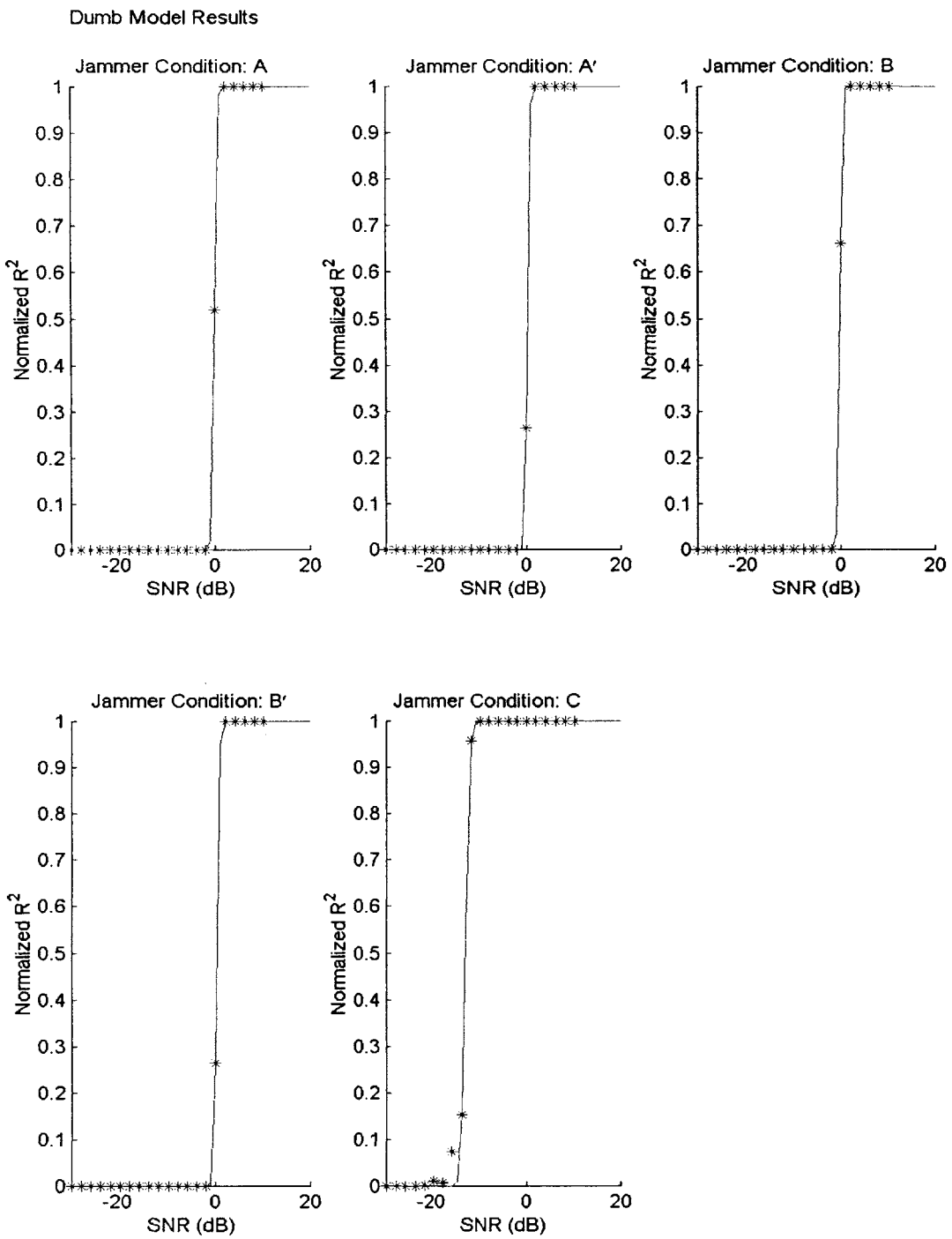


Fig. 11 Cumulative Gaussian functions that best fit the normalized $R^2$ data of Model 1 for all the jammer conditions.

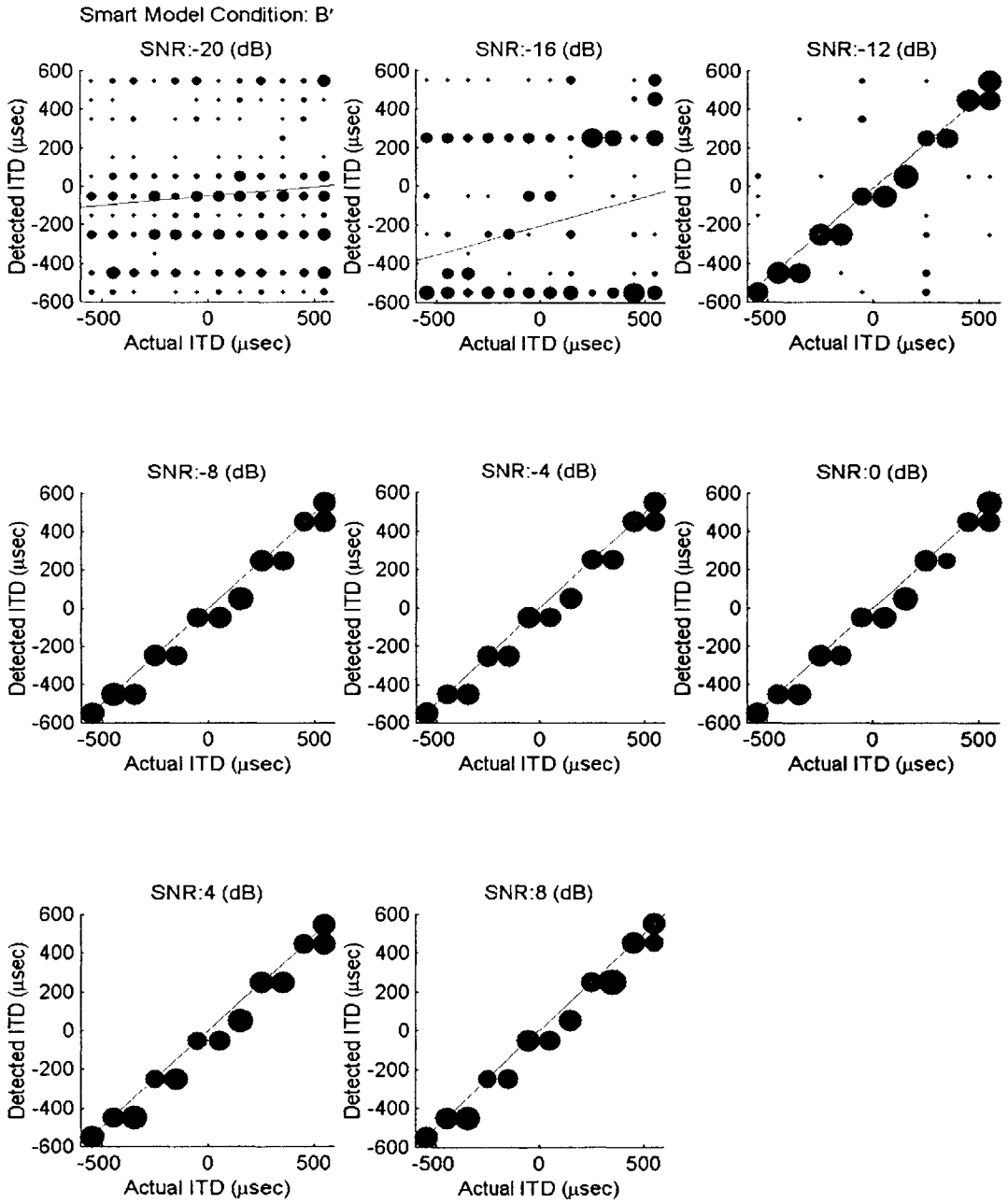


Fig. 12 Confusion matrices of Model 2, the *smart* cross-correlation model, for jammer condition B' and different TJRs

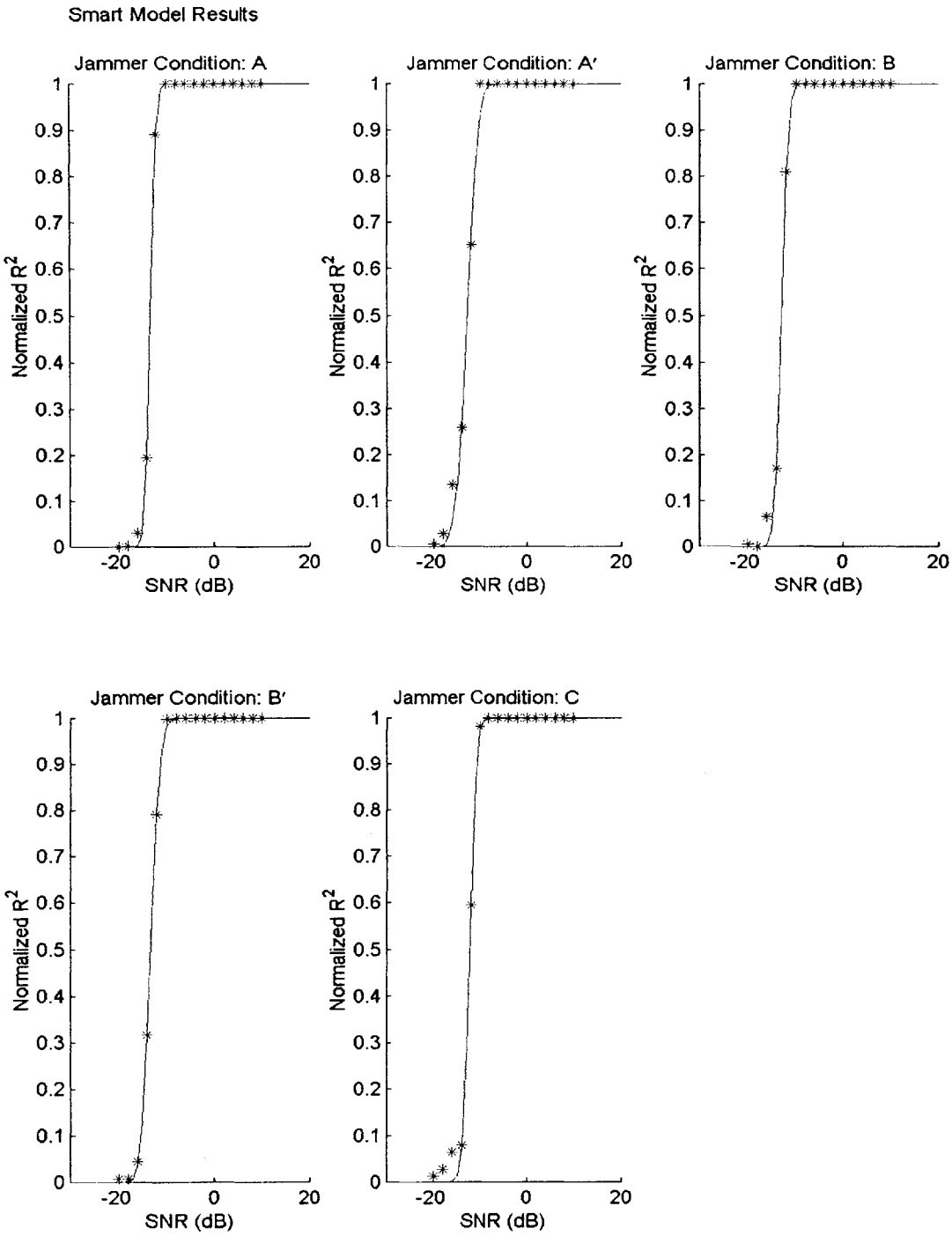


Fig. 13 the cumulative Gaussian functions that best fit the normalized $R^2$ data of Model 2 for all the jammer conditions.

## Appendix

The appendix consists of a series of figures representing all of the data collected and used in this thesis. The figures are identical in format to figures presented in the body of the thesis, varying only in the data of the subject and jammer condition being presented. Figures A-1 through A-4 are identical in format to Fig. 5, showing summary statistics for performance in quiet, but for subjects S1 through S5 respectively. Figures A-5 through A-32 are identical in format to Fig. 7, confusion matrices for a given subject and jammer conditions, but for subjects S1 through S5 respectively. Subject S3 performed certain jammer conditions twice. The results obtained from the first and second runs of these conditions are denoted S3 and S3' respectively. Figures A-33 through A-36 are identical in format to Fig. 8, normalized $R^2$ for the different jammer conditions, but for subjects S1 through S5 respectively.

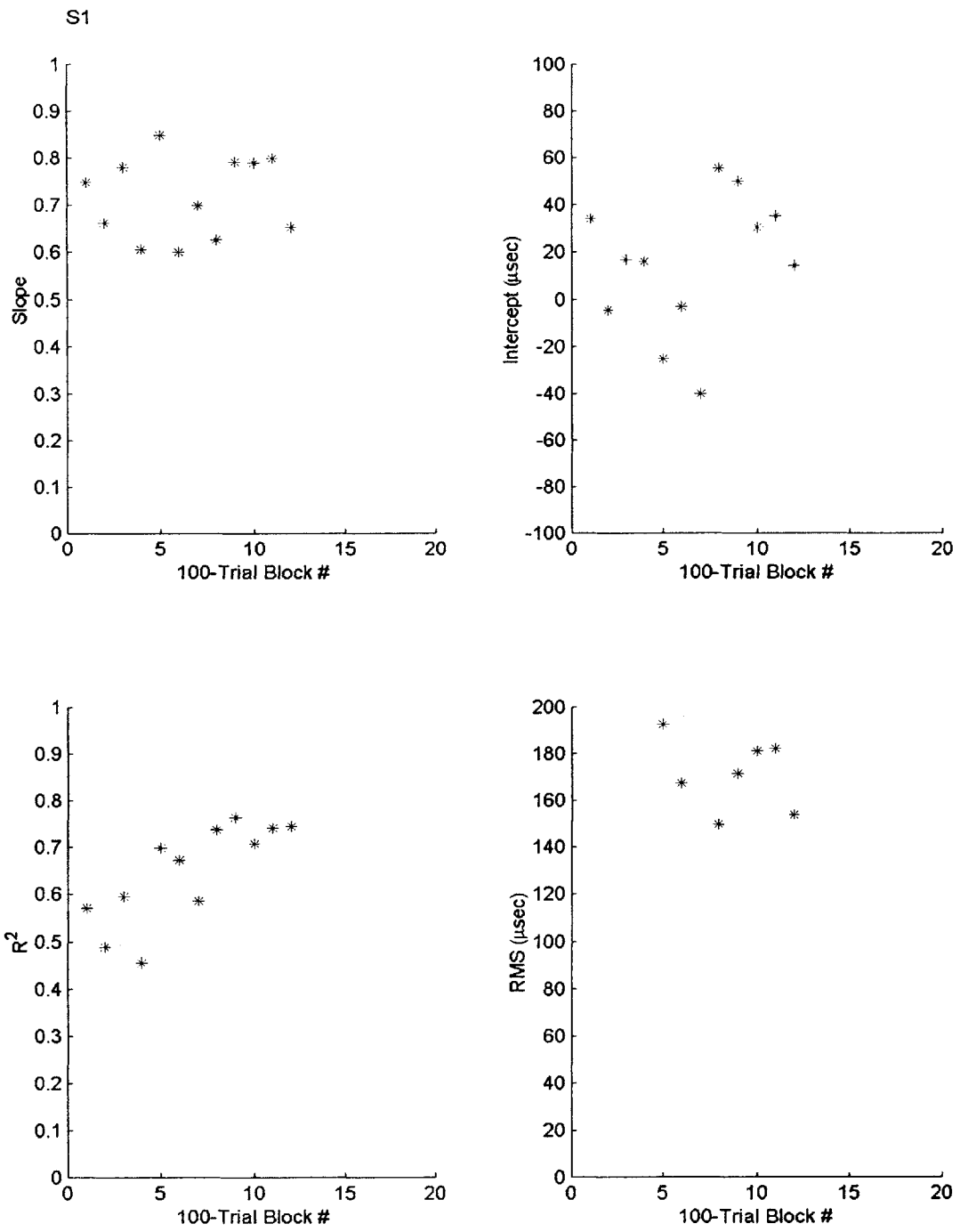
S1
Slope
100-Trial Block #
Intercept (µsec)
100-Trial Block #
R2
100-Trial Block #
RMS (µsec)
100-Trial Block #

Fig. A- 1

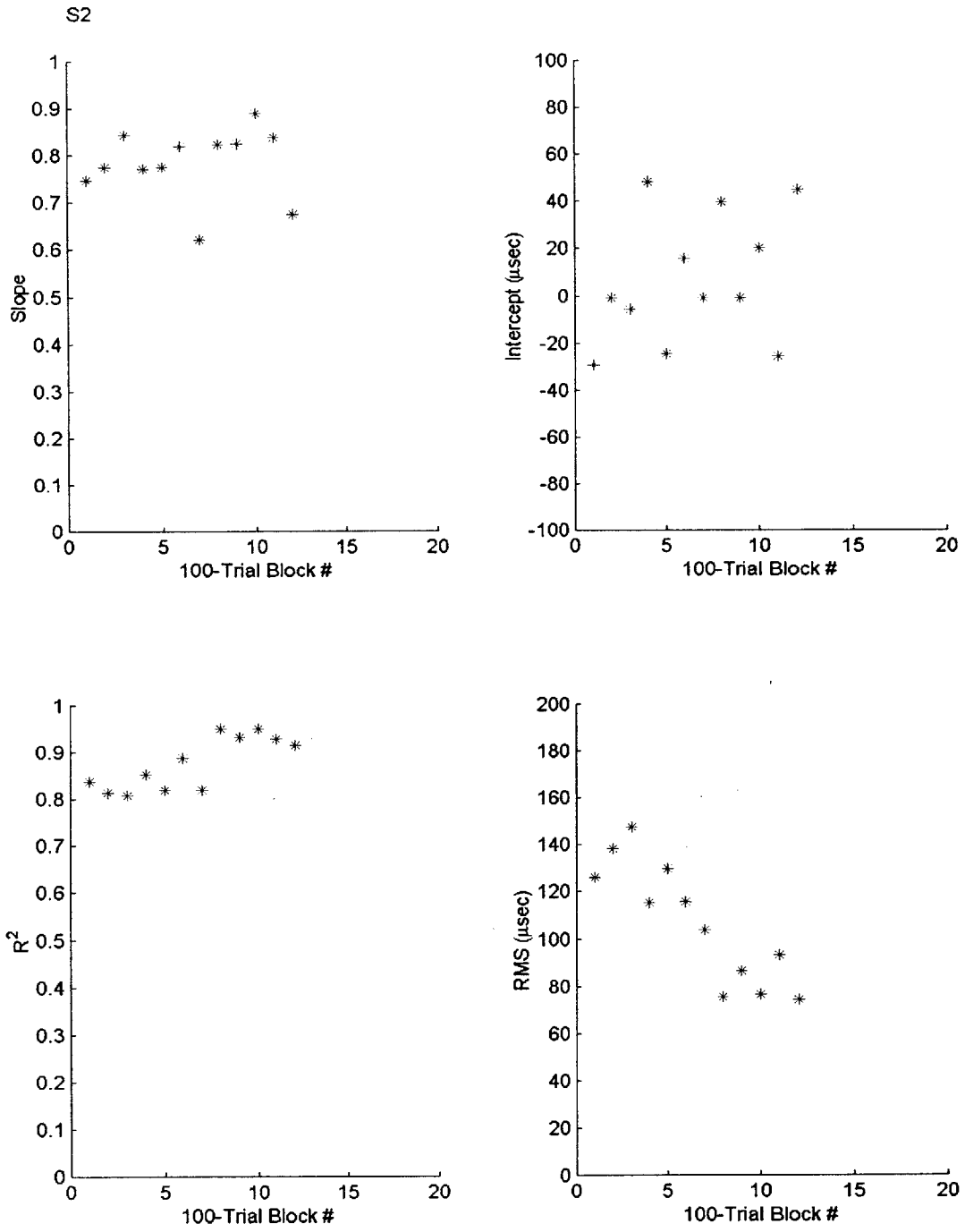
S2
Slope
100-Trial Block #
Intercept (µsec)
100-Trial Block #
$R^2$
100-Trial Block #
RMS (µsec)
100-Trial Block #

Fig. A- 2

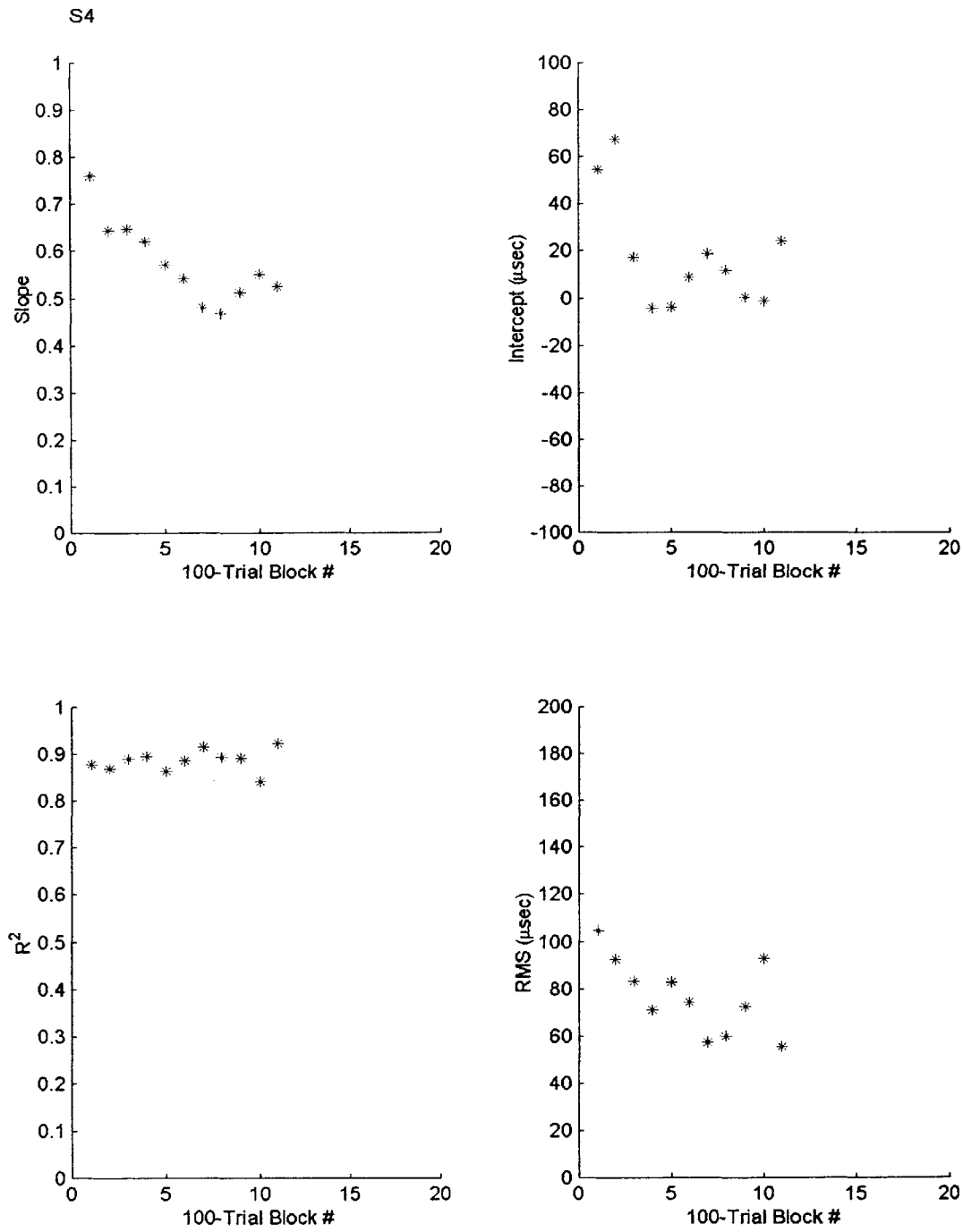
S4
Slope
100-Trial Block #
Intercept (µsec)
100-Trial Block #
$R^2$
100-Trial Block #
RMS (µsec)
100-Trial Block #

Fig. A- 3

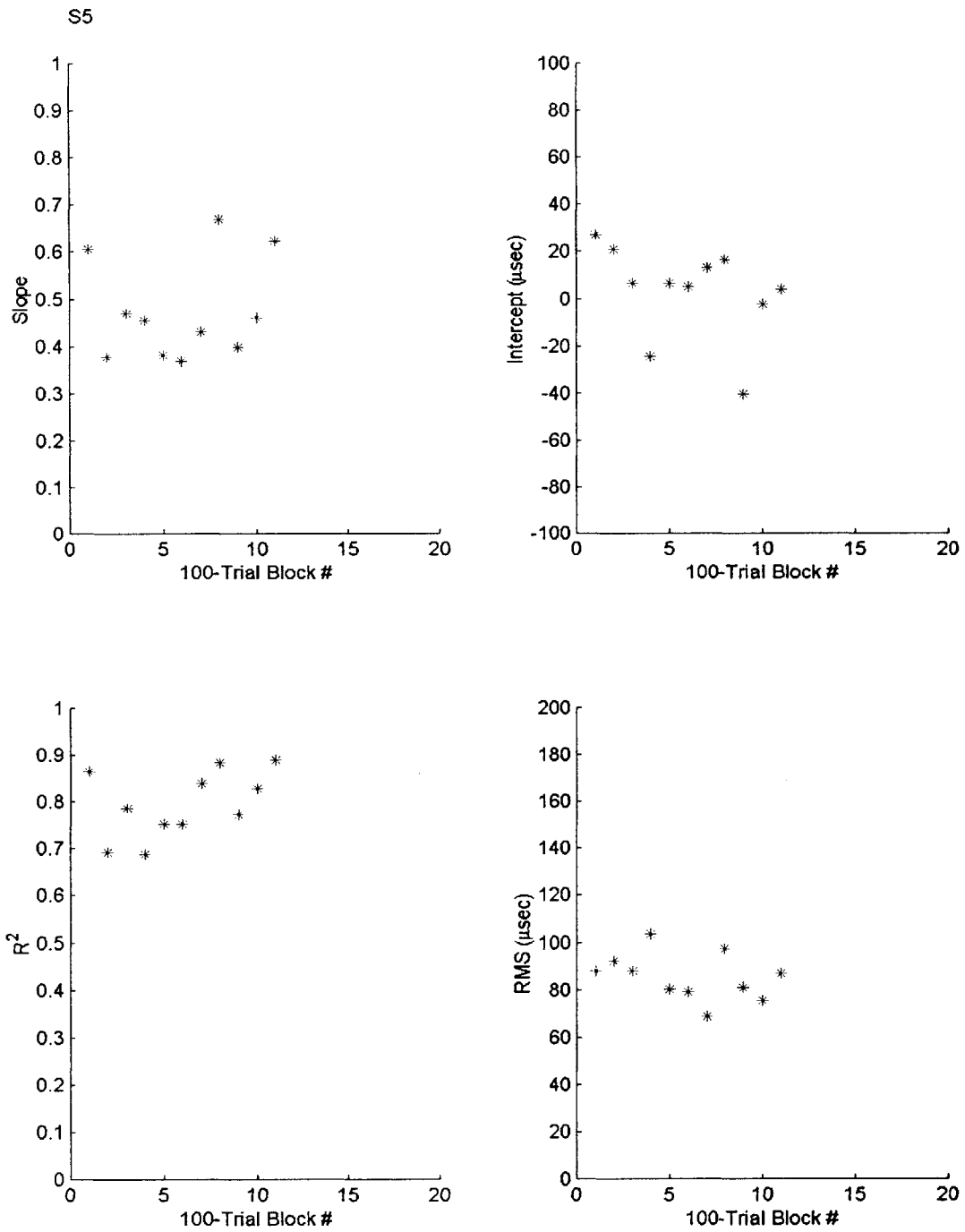
S5
Slope
100-Trial Block #
Intercept (µsec)
100-Trial Block #
$R^2$
100-Trial Block #
RMS (µsec)
100-Trial Block #

Fig. A- 4

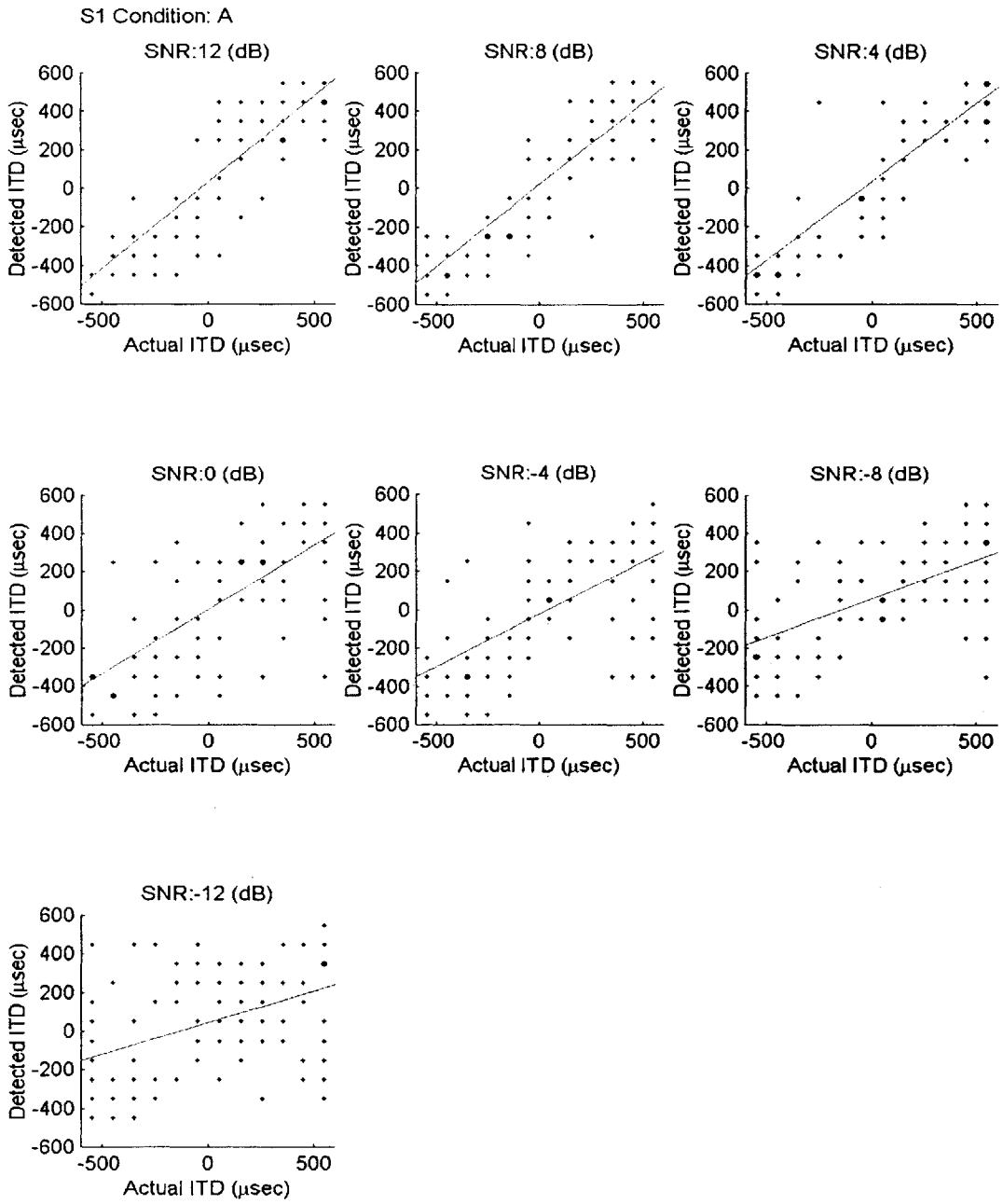
S1 Condition: A
SNR:12 (dB)
SNR:8 (dB)
SNR:4 (dB)
SNR:0 (dB)
SNR:-4 (dB)
SNR:-8 (dB)
SNR:-12 (dB)
Detected ITD (μsec)
Actual ITD (μsec)
600
400
200
0
-200
-400
-600
-500
0
500

Fig. A- 5

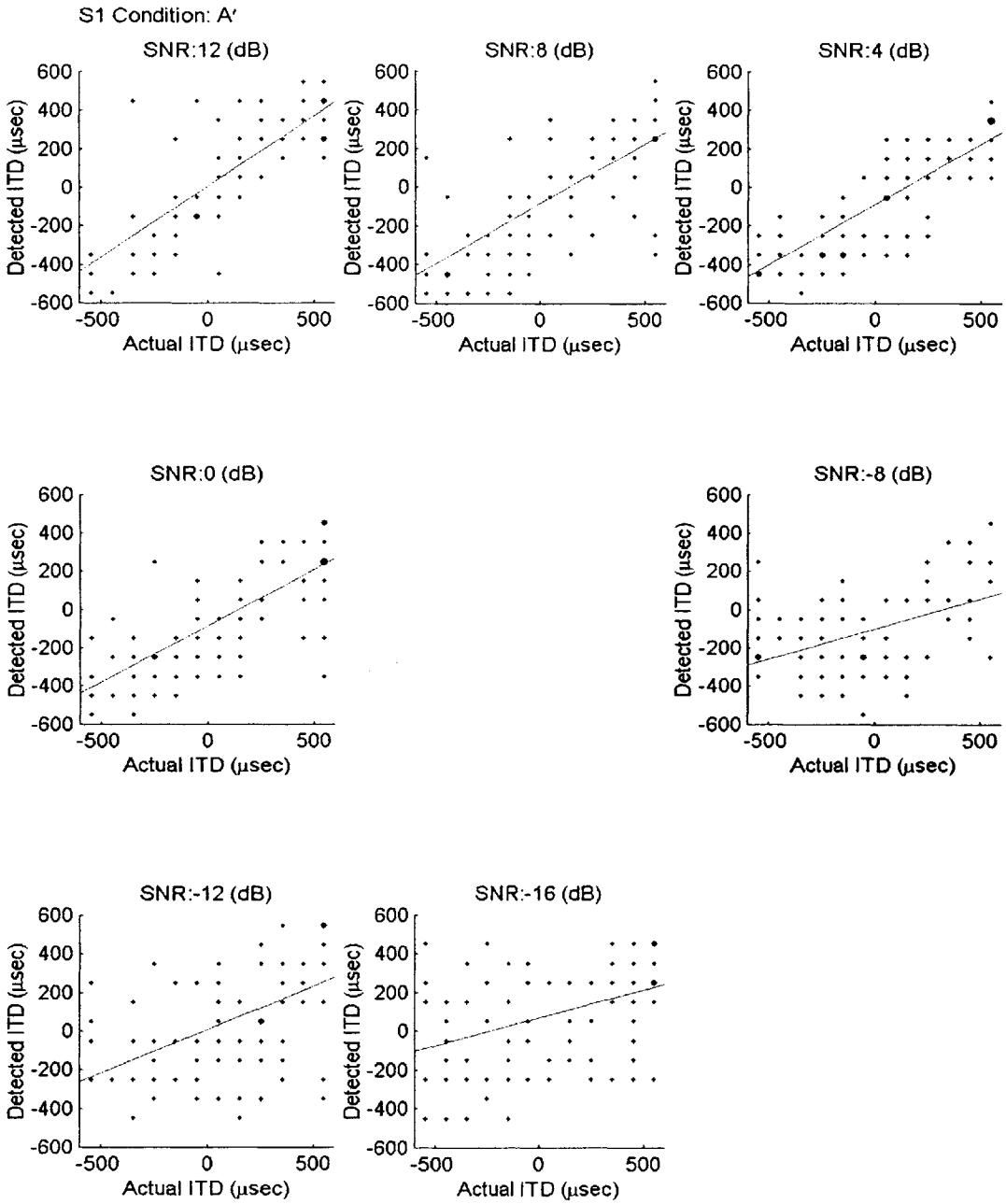
S1 Condition: A'
SNR:12 (dB)
SNR:8 (dB)
SNR:4 (dB)
SNR:0 (dB)
SNR:-8 (dB)
SNR:-12 (dB)
SNR:-16 (dB)
Detected ITD (μsec)
Actual ITD (μsec)
600
400
200
0
-200
-400
-600
-500
0
500

Fig. A- 6

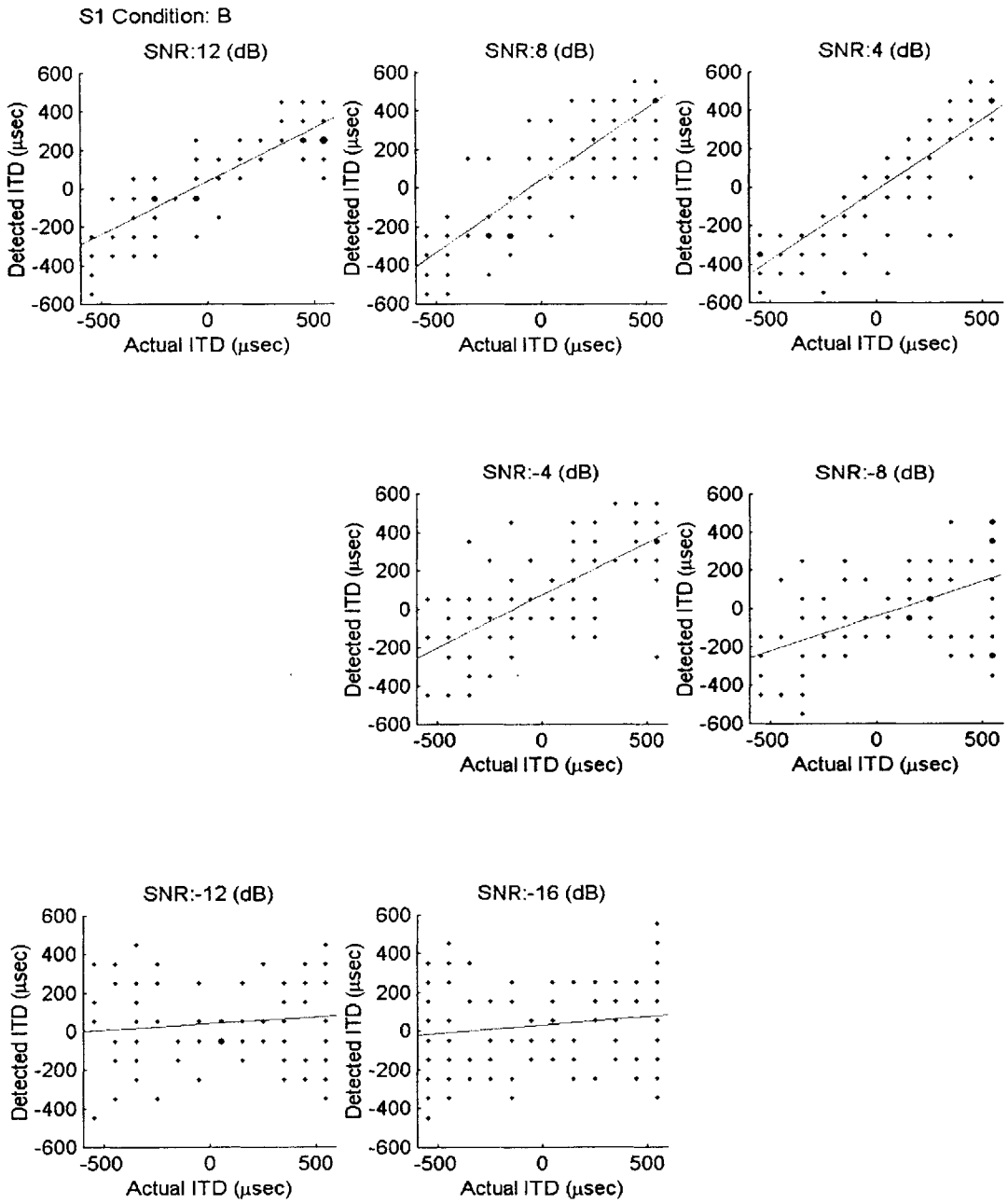
S1 Condition: B
SNR:12 (dB)
SNR:8 (dB)
SNR:4 (dB)
SNR:-4 (dB)
SNR:-8 (dB)
SNR:-12 (dB)
SNR:-16 (dB)
Detected ITD (μsec)
Actual ITD (μsec)
600
400
200
0
-200
-400
-600
-500
0
500

Fig. A- 7

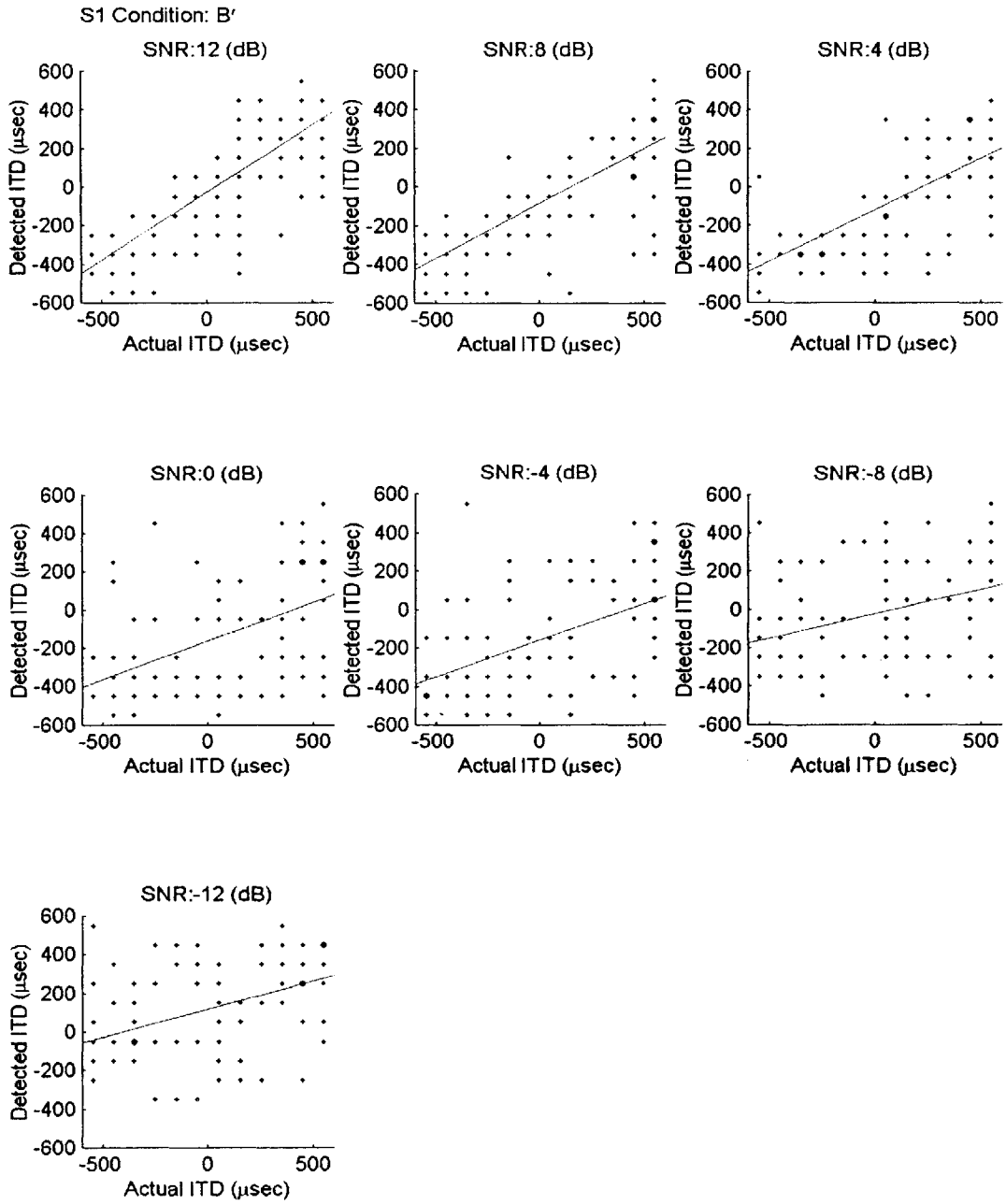
S1 Condition: B′
SNR:12 (dB)
SNR:8 (dB)
SNR:4 (dB)
SNR:0 (dB)
SNR:-4 (dB)
SNR:-8 (dB)
SNR:-12 (dB)
Detected ITD (μsec)
Actual ITD (μsec)
600
400
200
0
-200
-400
-600
-500
0
500

Fig. A- 8

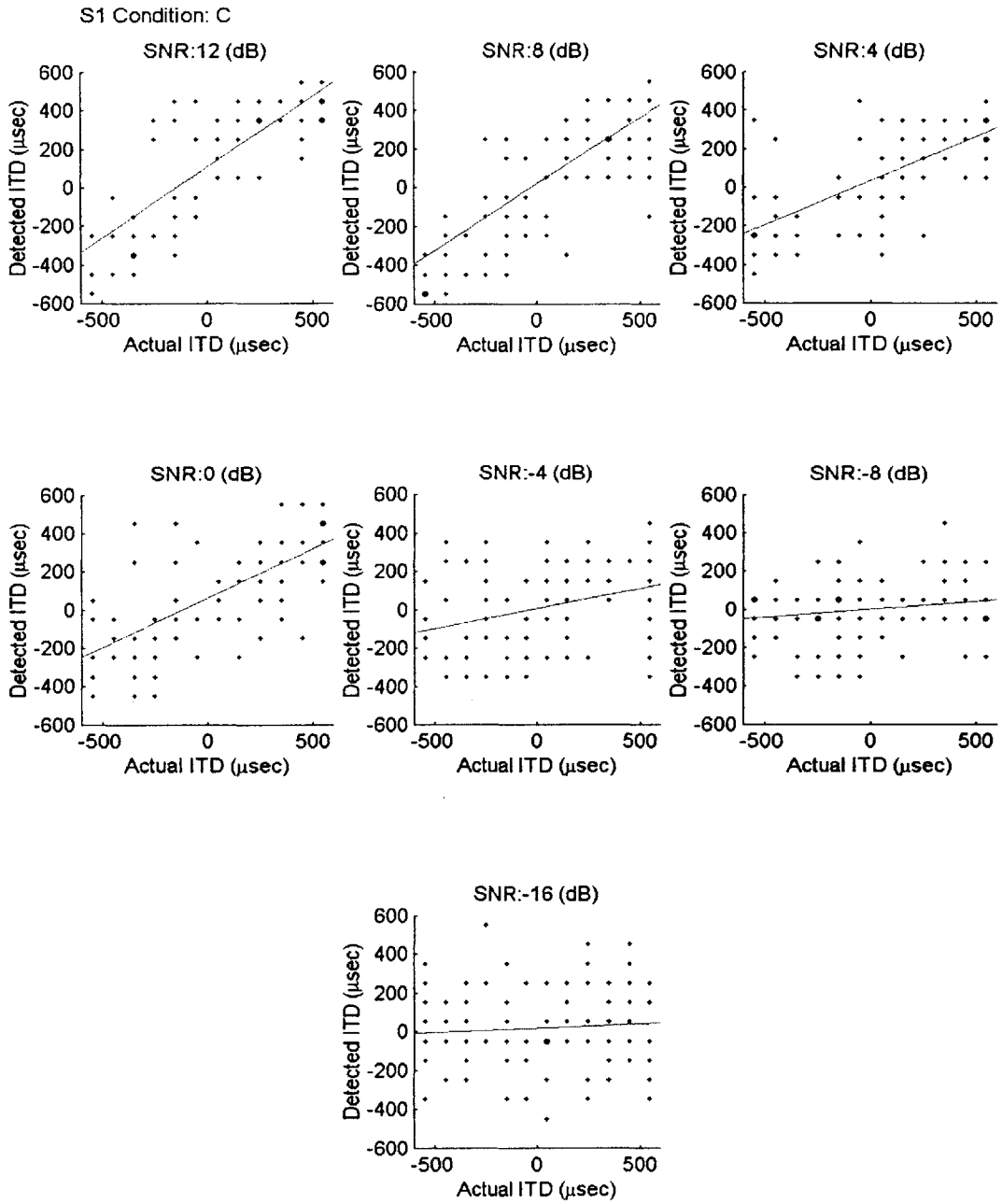
S1 Condition: C
SNR:12 (dB)
SNR:8 (dB)
SNR:4 (dB)
SNR:0 (dB)
SNR:-4 (dB)
SNR:-8 (dB)
SNR:-16 (dB)
Detected ITD (μsec)
Actual ITD (μsec)
600
400
200
0
-200
-400
-600
-500
0
500

Fig. A- 9

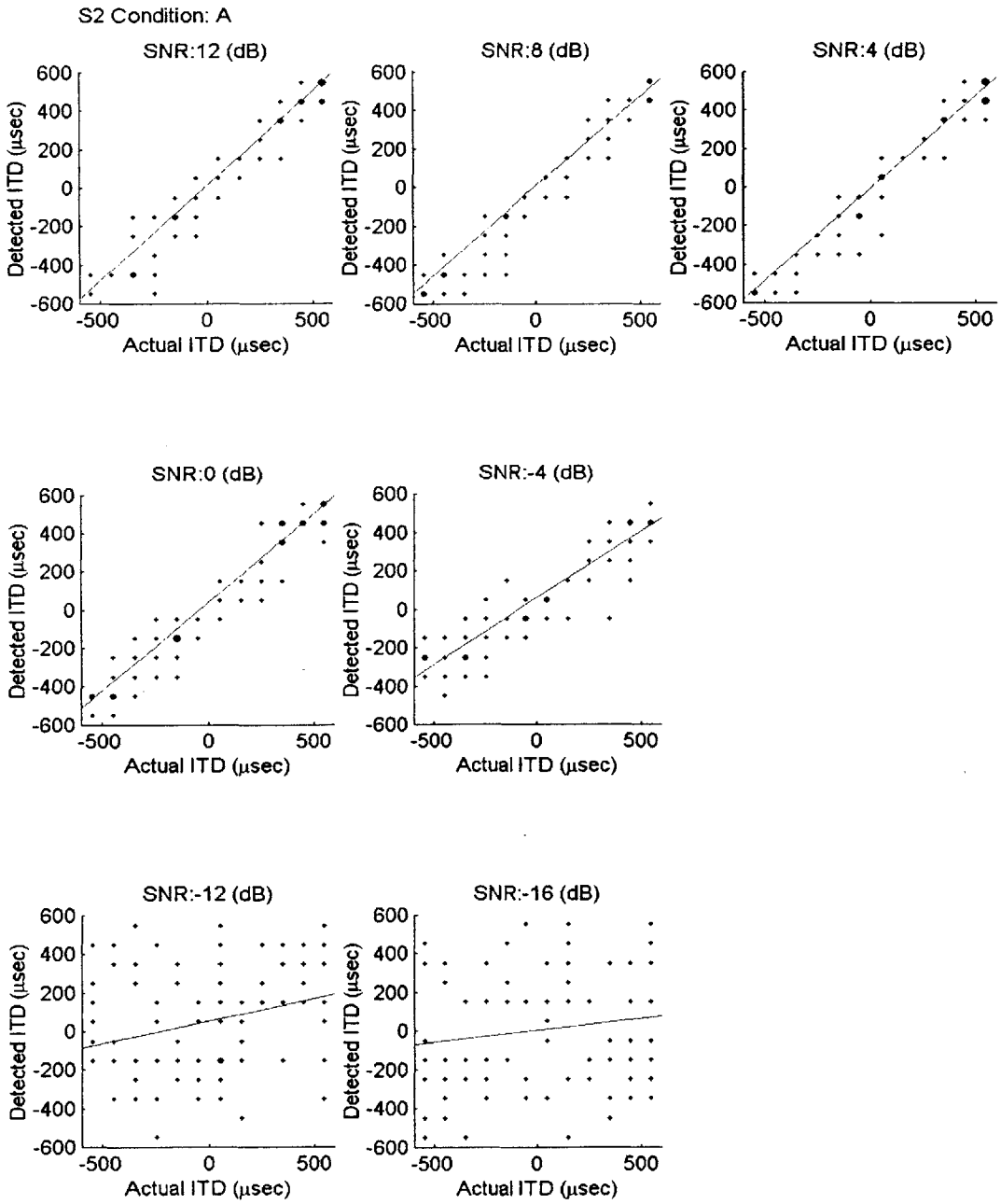
S2 Condition: A
SNR:12 (dB)
SNR:8 (dB)
SNR:4 (dB)
SNR:0 (dB)
SNR:-4 (dB)
SNR:-12 (dB)
SNR:-16 (dB)
Detected ITD (μsec)
Actual ITD (μsec)
600
400
200
0
-200
-400
-600
-500
0
500

Fig. A- 10

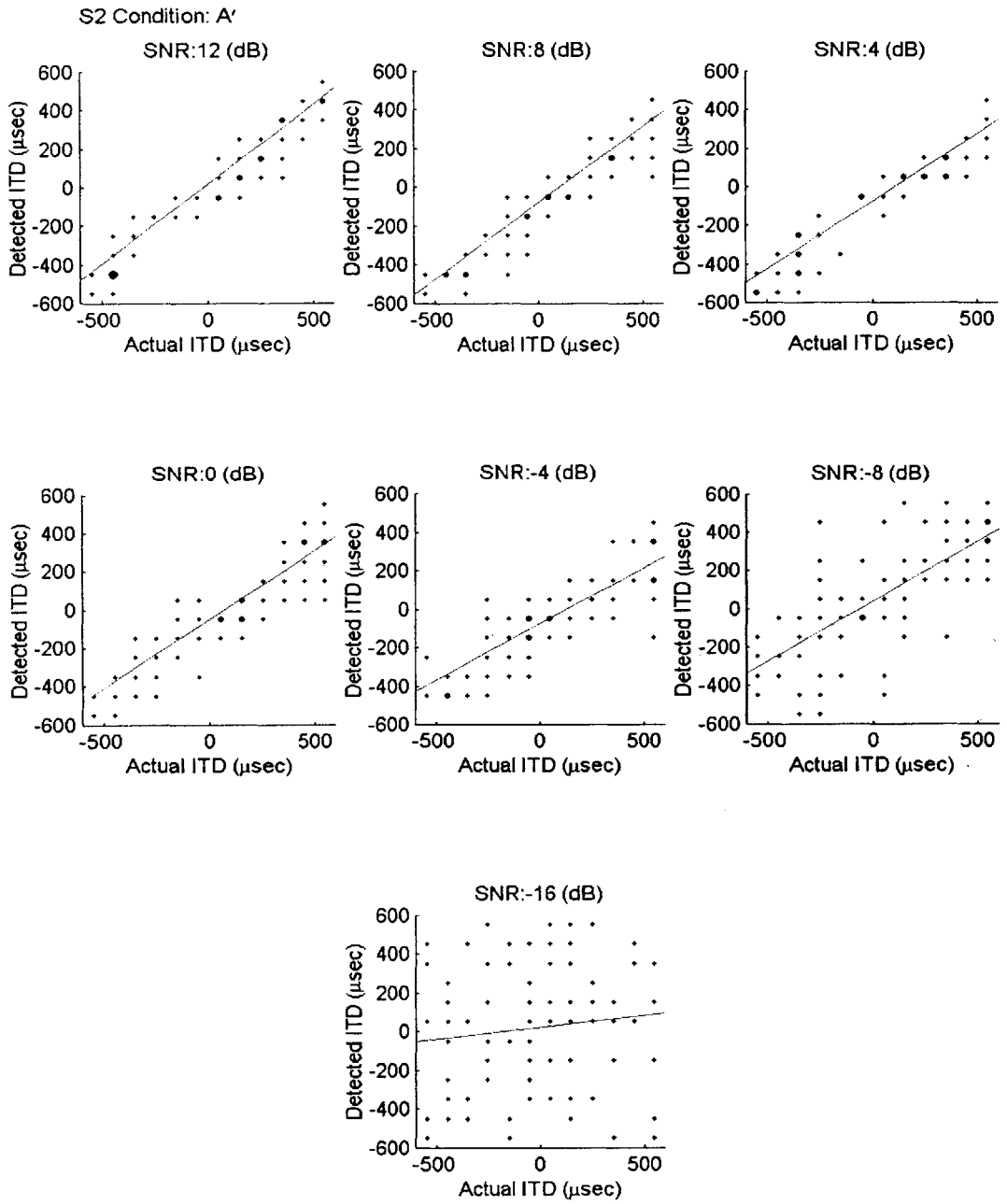
S2 Condition: A'
SNR:12 (dB)
SNR:8 (dB)
SNR:4 (dB)
SNR:0 (dB)
SNR:-4 (dB)
SNR:-8 (dB)
SNR:-16 (dB)
Detected ITD (μsec)
Actual ITD (μsec)
600
400
200
0
-200
-400
-600
-500
0
500

Fig. A- 11

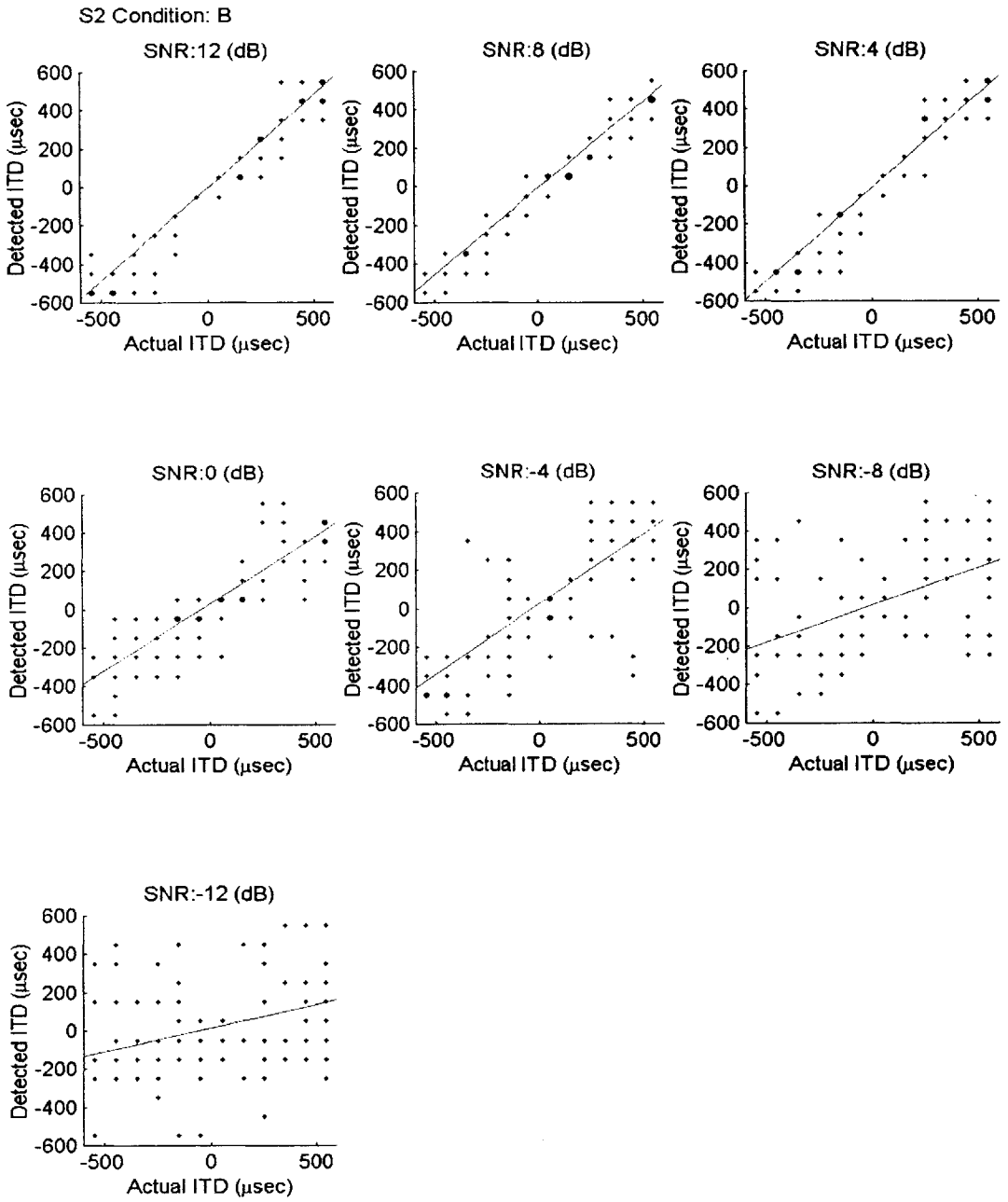
S2 Condition: B
SNR:12 (dB)
SNR:8 (dB)
SNR:4 (dB)
SNR:0 (dB)
SNR:-4 (dB)
SNR:-8 (dB)
SNR:-12 (dB)
Detected ITD (μsec)
Actual ITD (μsec)
600
400
200
0
-200
-400
-600
-500
0
500

Fig. A- 12

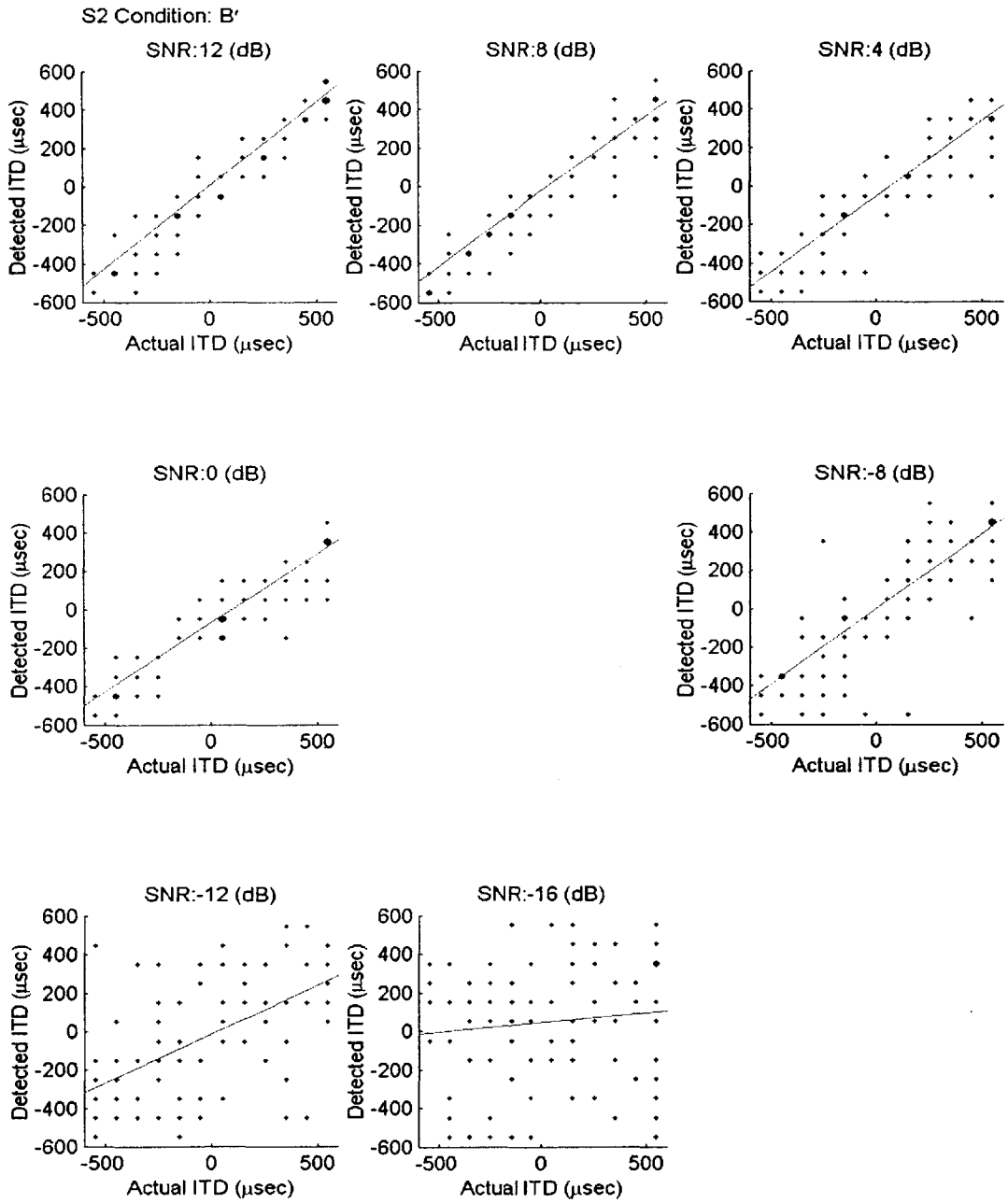
S2 Condition: B′
SNR:12 (dB)
SNR:8 (dB)
SNR:4 (dB)
SNR:0 (dB)
SNR:-8 (dB)
SNR:-12 (dB)
SNR:-16 (dB)
Detected ITD (μsec)
Actual ITD (μsec)
600
400
200
0
-200
-400
-600
-500
0
500

Fig. A- 13

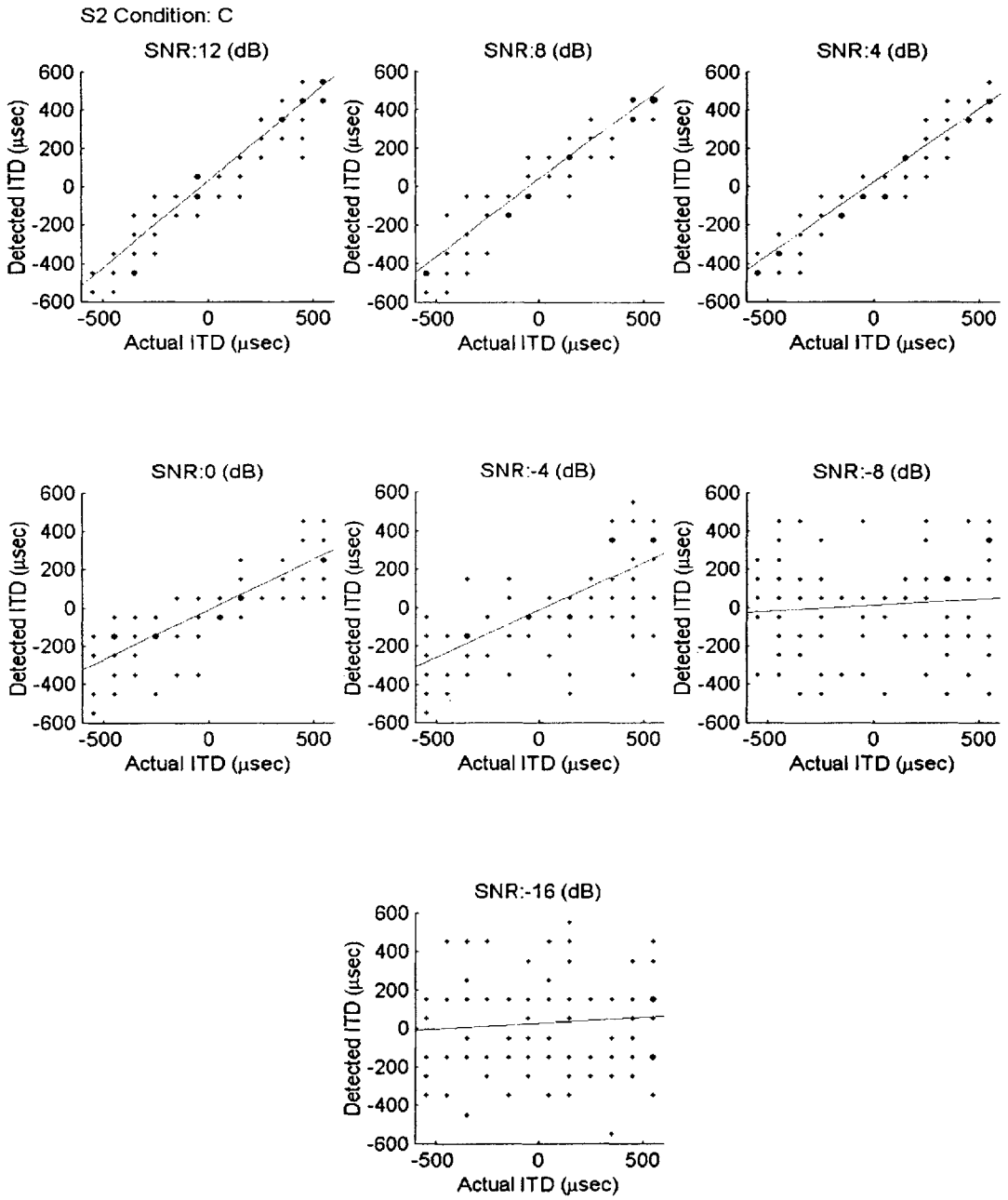
S2 Condition: C
SNR:12 (dB)
SNR:8 (dB)
SNR:4 (dB)
SNR:0 (dB)
SNR:-4 (dB)
SNR:-8 (dB)
SNR:-16 (dB)
Detected ITD (μsec)
Actual ITD (μsec)
600
400
200
0
-200
-400
-600
-500
0
500

Fig. A- 14

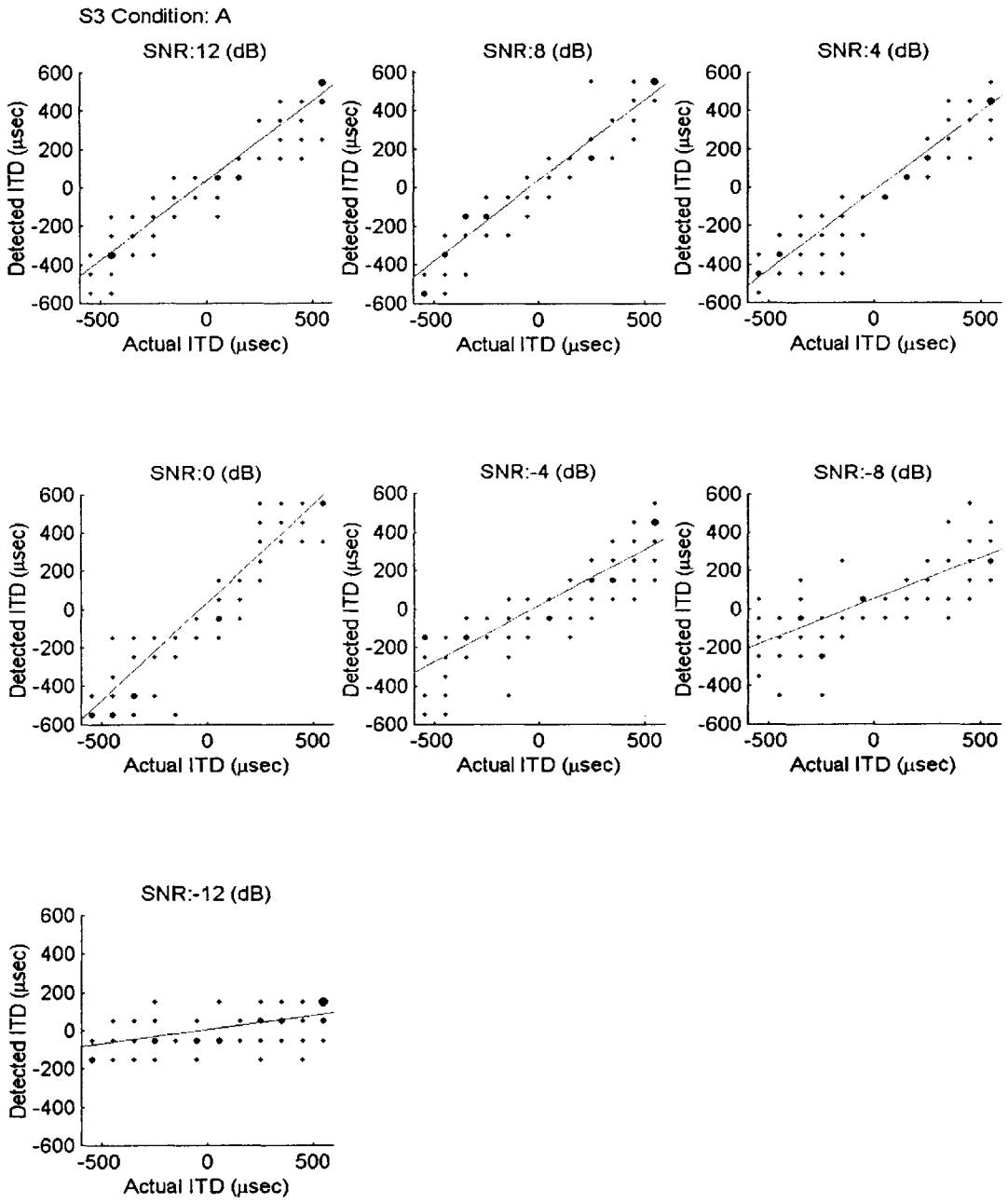
S3 Condition: A
SNR:12 (dB)
SNR:8 (dB)
SNR:4 (dB)
SNR:0 (dB)
SNR:-4 (dB)
SNR:-8 (dB)
SNR:-12 (dB)
Detected ITD (μsec)
Actual ITD (μsec)
600
400
200
0
-200
-400
-600
-500
0
500

Fig. A- 15

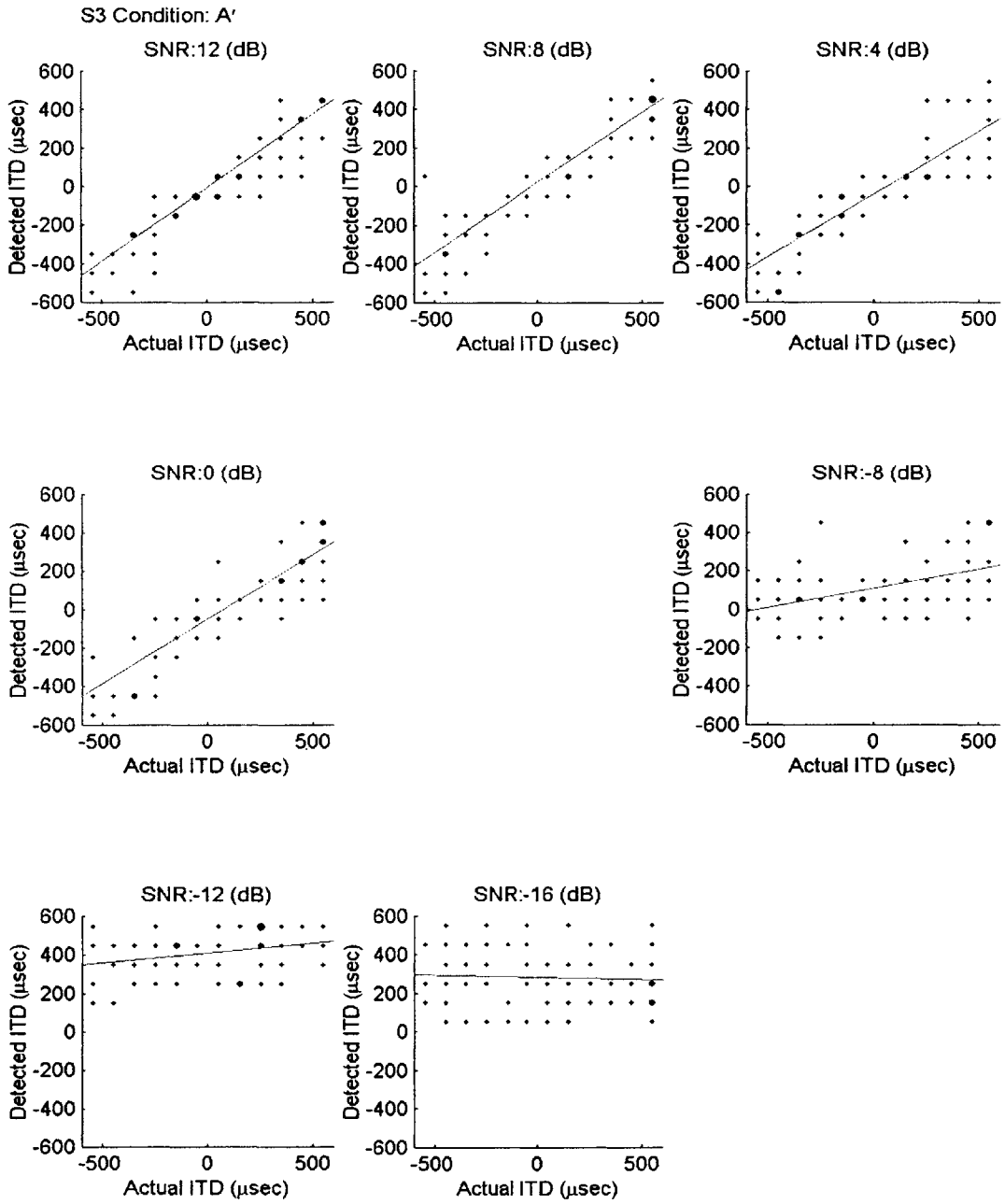
S3 Condition: A'
SNR:12 (dB)
SNR:8 (dB)
SNR:4 (dB)
SNR:0 (dB)
SNR:-8 (dB)
SNR:-12 (dB)
SNR:-16 (dB)
Detected ITD (μsec)
Actual ITD (μsec)
600
400
200
0
-200
-400
-600
-500
0
500

Fig. A- 16

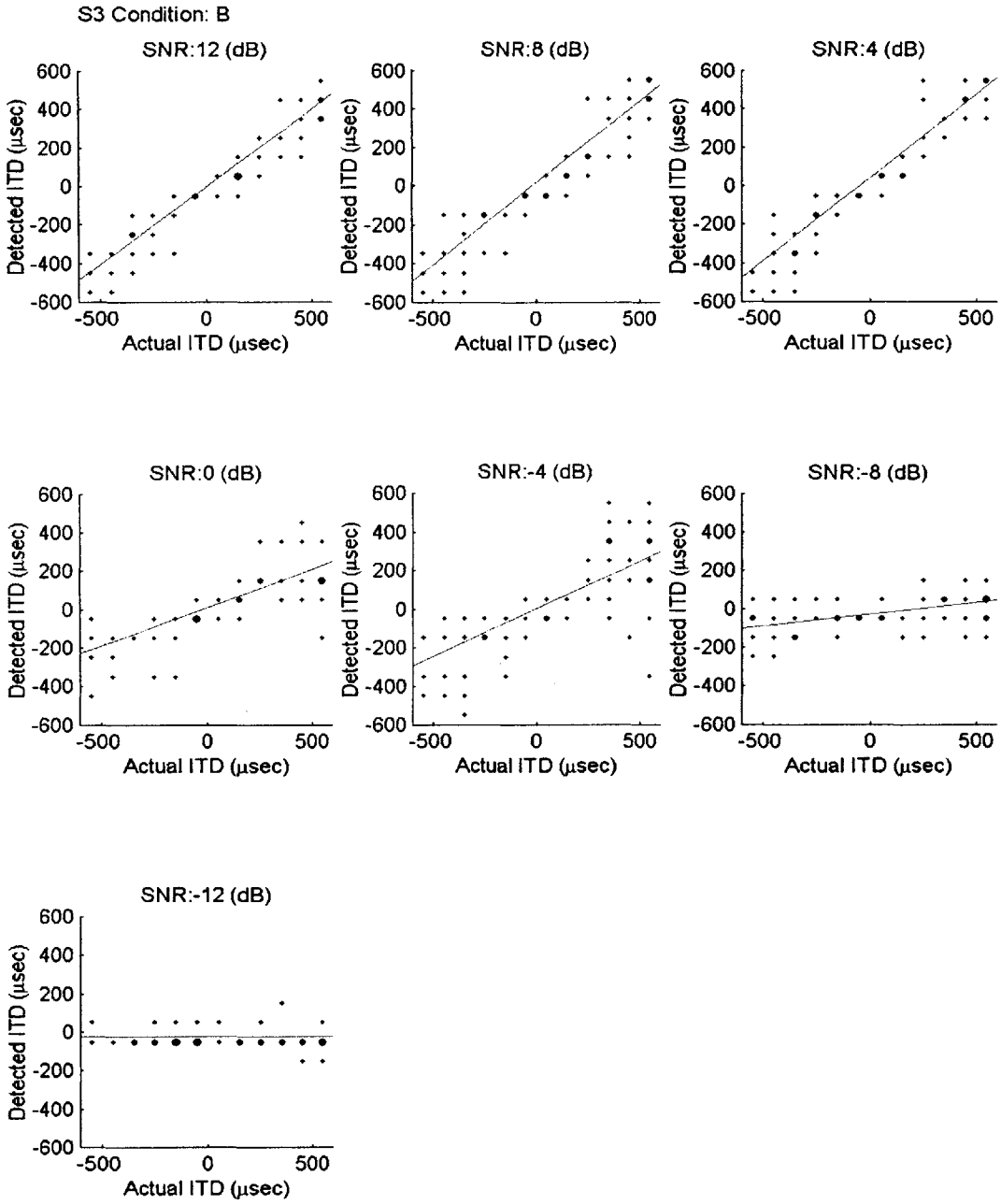
S3 Condition: B
SNR:12 (dB)
SNR:8 (dB)
SNR:4 (dB)
SNR:0 (dB)
SNR:-4 (dB)
SNR:-8 (dB)
SNR:-12 (dB)
Detected ITD (μsec)
Actual ITD (μsec)
600
400
200
0
-200
-400
-600
-500
0
500

Fig. A- 17

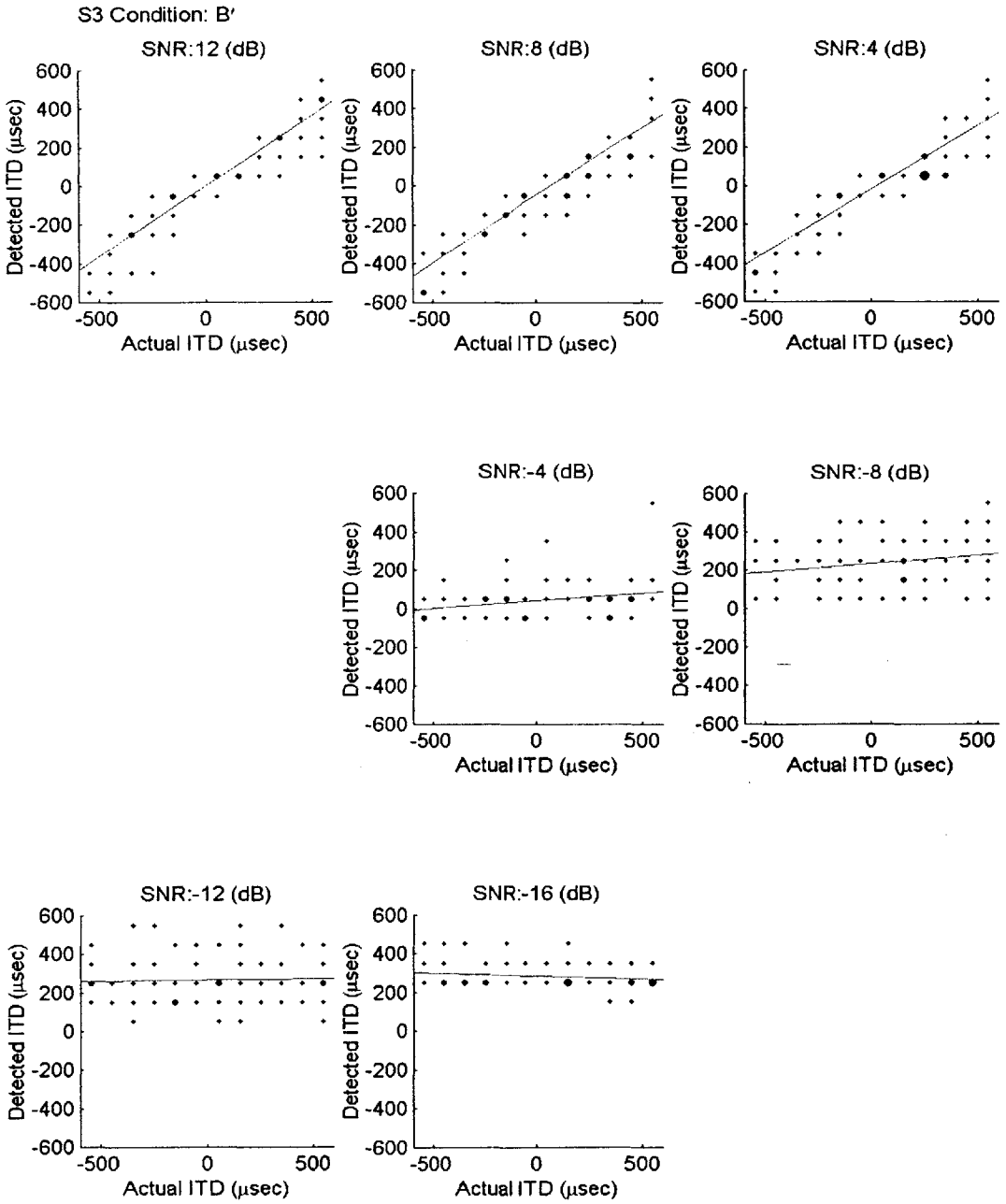
S3 Condition: B′
SNR:12 (dB)
SNR:8 (dB)
SNR:4 (dB)
SNR:-4 (dB)
SNR:-8 (dB)
SNR:-12 (dB)
SNR:-16 (dB)
Detected ITD (μsec)
Actual ITD (μsec)
600
400
200
0
-200
-400
-600
-500
0
500

Fig. A- 18

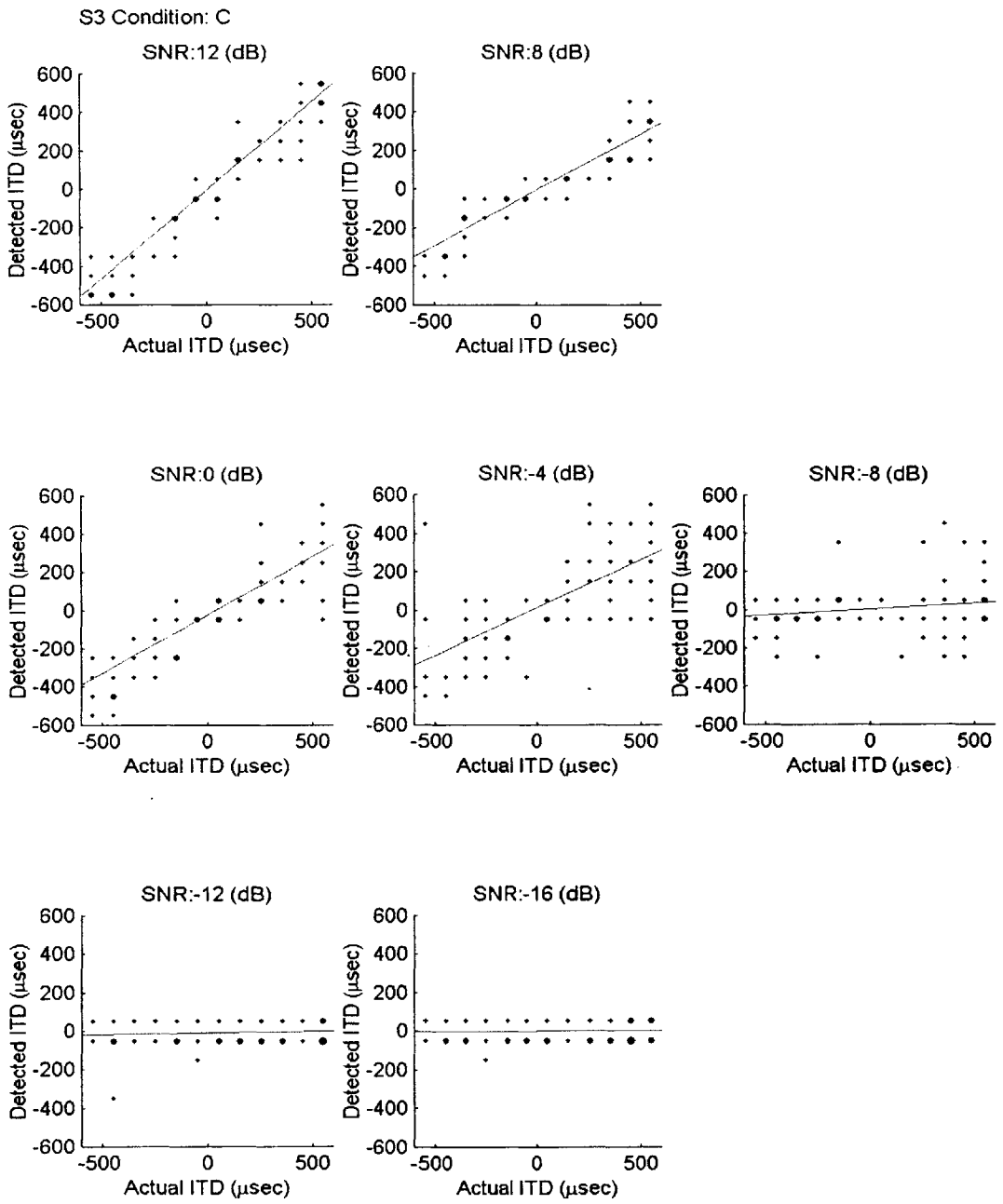
S3 Condition: C
SNR:12 (dB)
SNR:8 (dB)
SNR:0 (dB)
SNR:-4 (dB)
SNR:-8 (dB)
SNR:-12 (dB)
SNR:-16 (dB)
Detected ITD (μsec)
Actual ITD (μsec)
600
400
200
0
-200
-400
-600
-500
0
500

Fig. A- 19

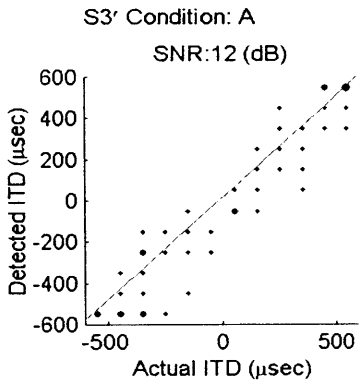
S3′ Condition: A
SNR:12 (dB)
600
400
200
0
-200
-400
-600
Detected ITD (μsec)
-500
0
500
Actual ITD (μsec)

Fig. A- 20

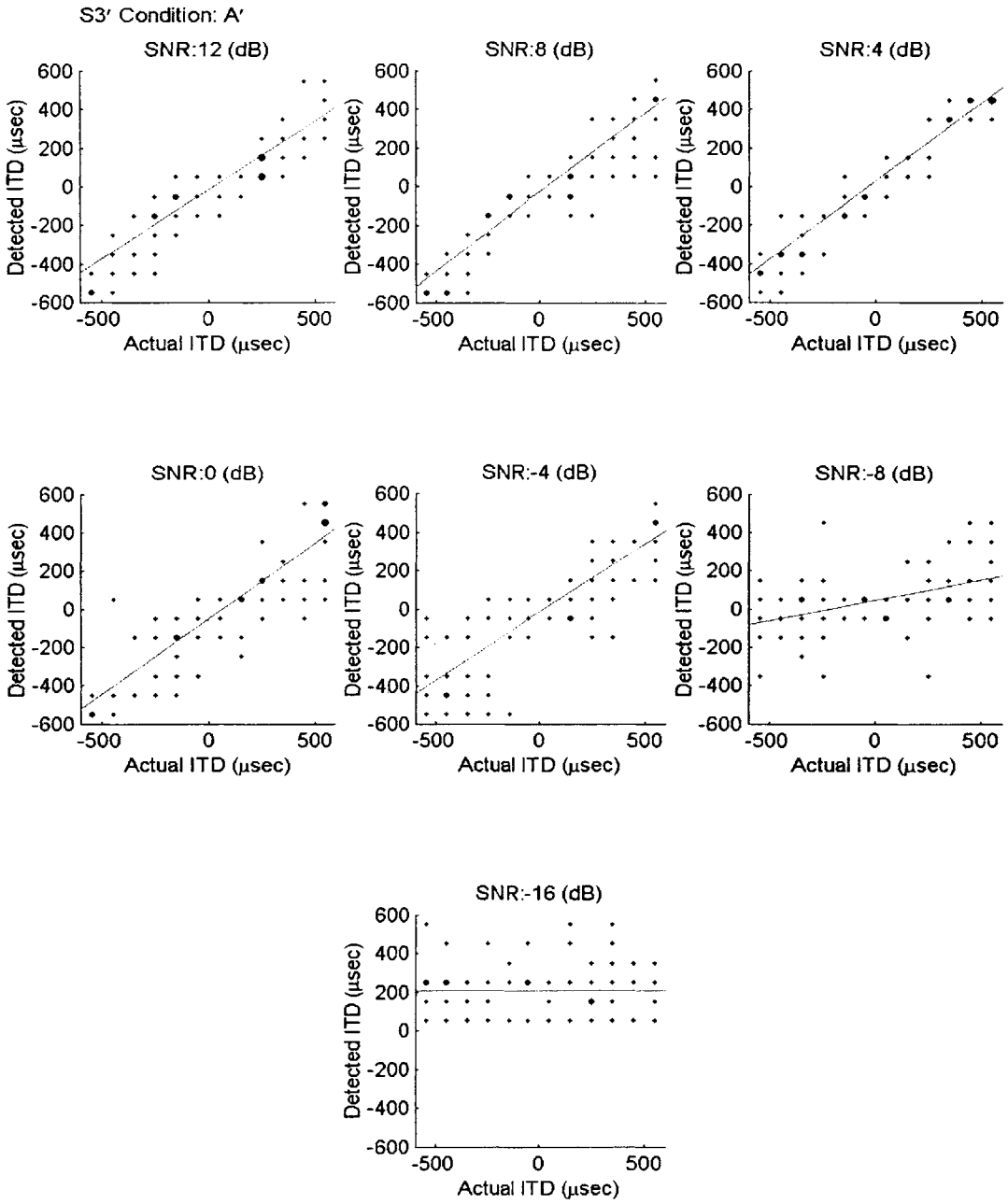
S3′ Condition: A′
SNR:12 (dB)
SNR:8 (dB)
SNR:4 (dB)
SNR:0 (dB)
SNR:-4 (dB)
SNR:-8 (dB)
SNR:-16 (dB)
Detected ITD (μsec)
Actual ITD (μsec)
600
400
200
0
-200
-400
-600
-500
0
500

Fig. A- 21

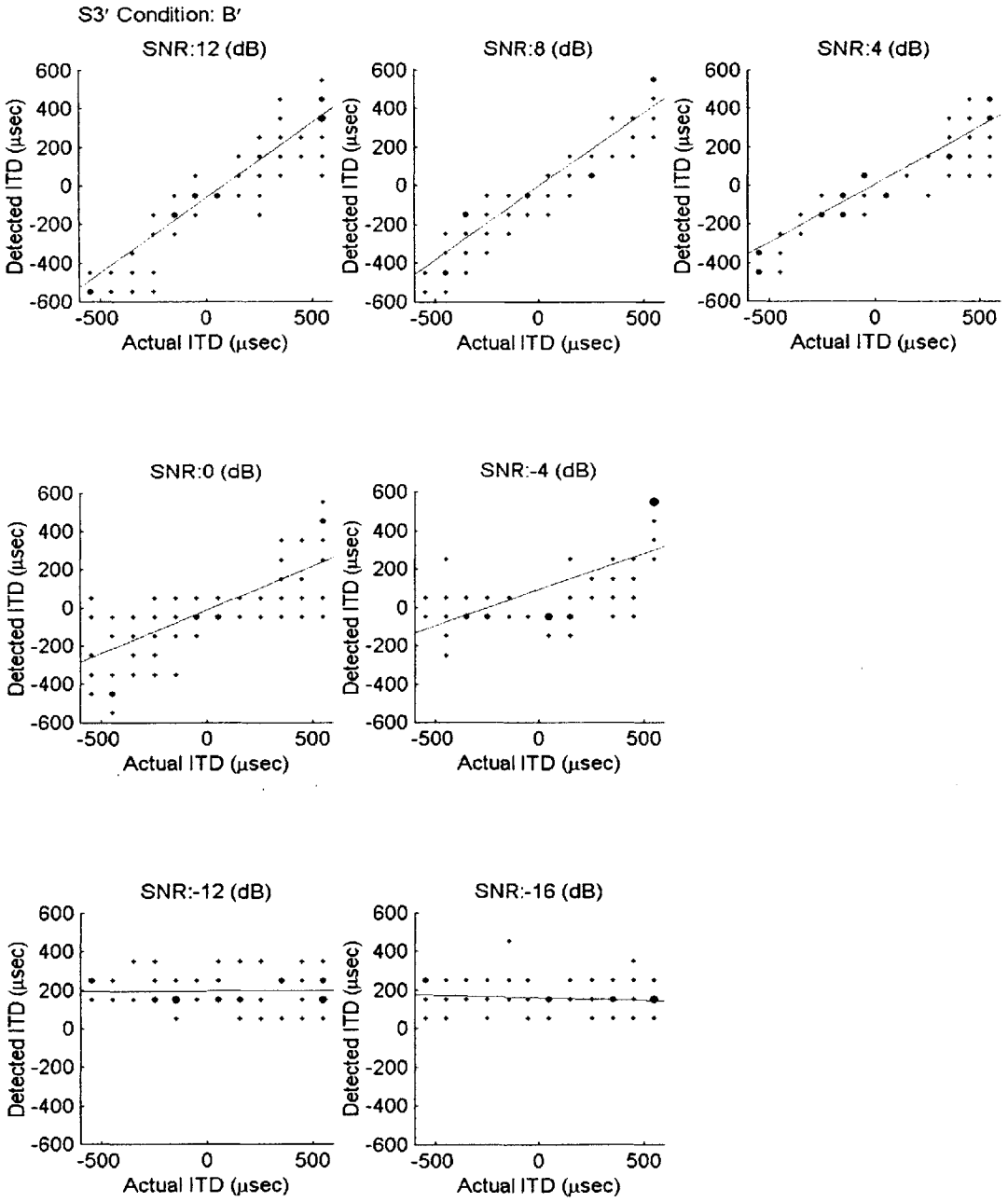
S3′ Condition: B′
SNR:12 (dB)
SNR:8 (dB)
SNR:4 (dB)
SNR:0 (dB)
SNR:-4 (dB)
SNR:-12 (dB)
SNR:-16 (dB)
Detected ITD (μsec)
Actual ITD (μsec)
600
400
200
0
-200
-400
-600
-500
0
500

Fig. A- 22

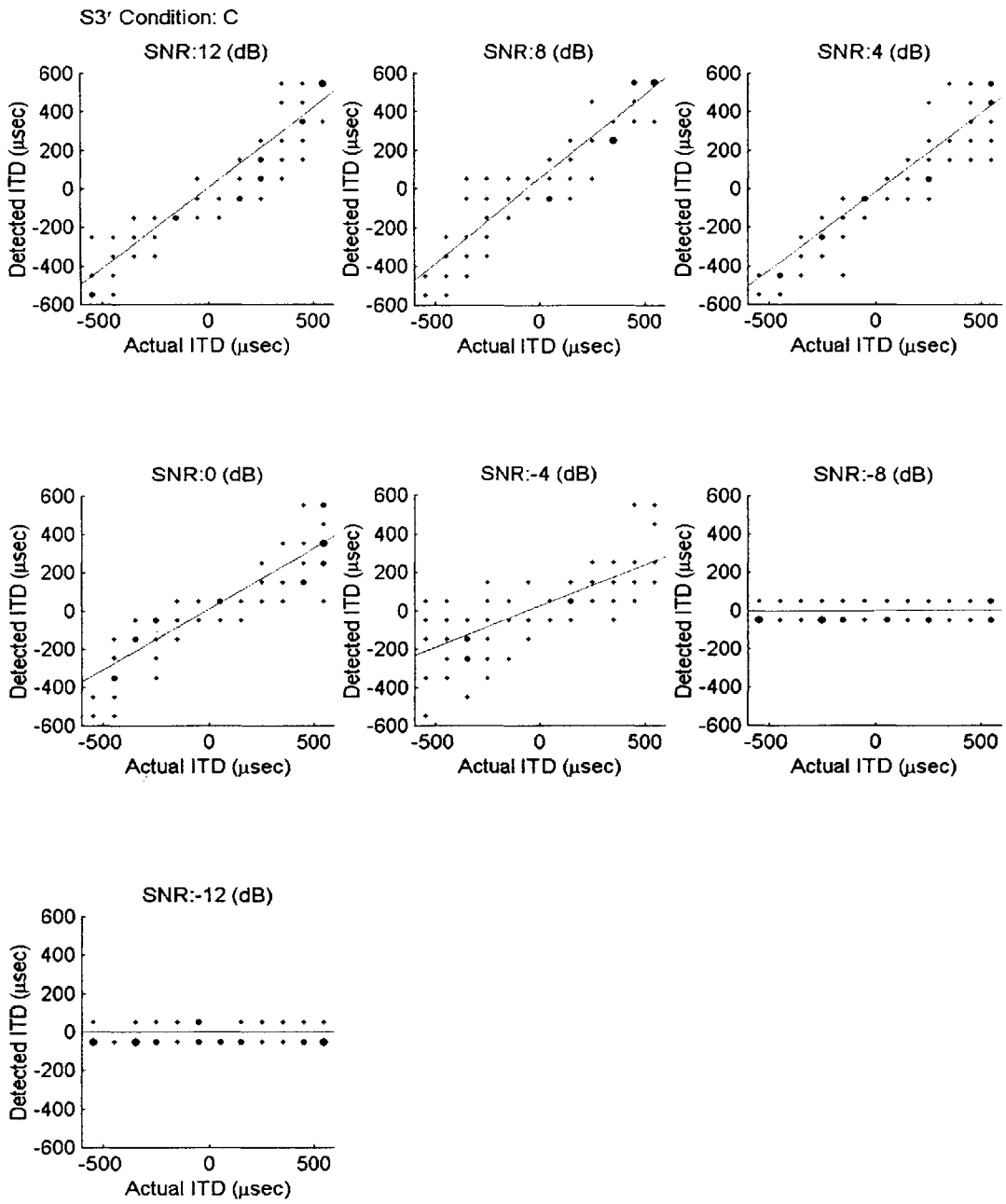
S3′ Condition: C
SNR:12 (dB)
SNR:8 (dB)
SNR:4 (dB)
SNR:0 (dB)
SNR:-4 (dB)
SNR:-8 (dB)
SNR:-12 (dB)
Detected ITD (μsec)
Actual ITD (μsec)
600
400
200
0
-200
-400
-600
-500
0
500

Fig. A- 23

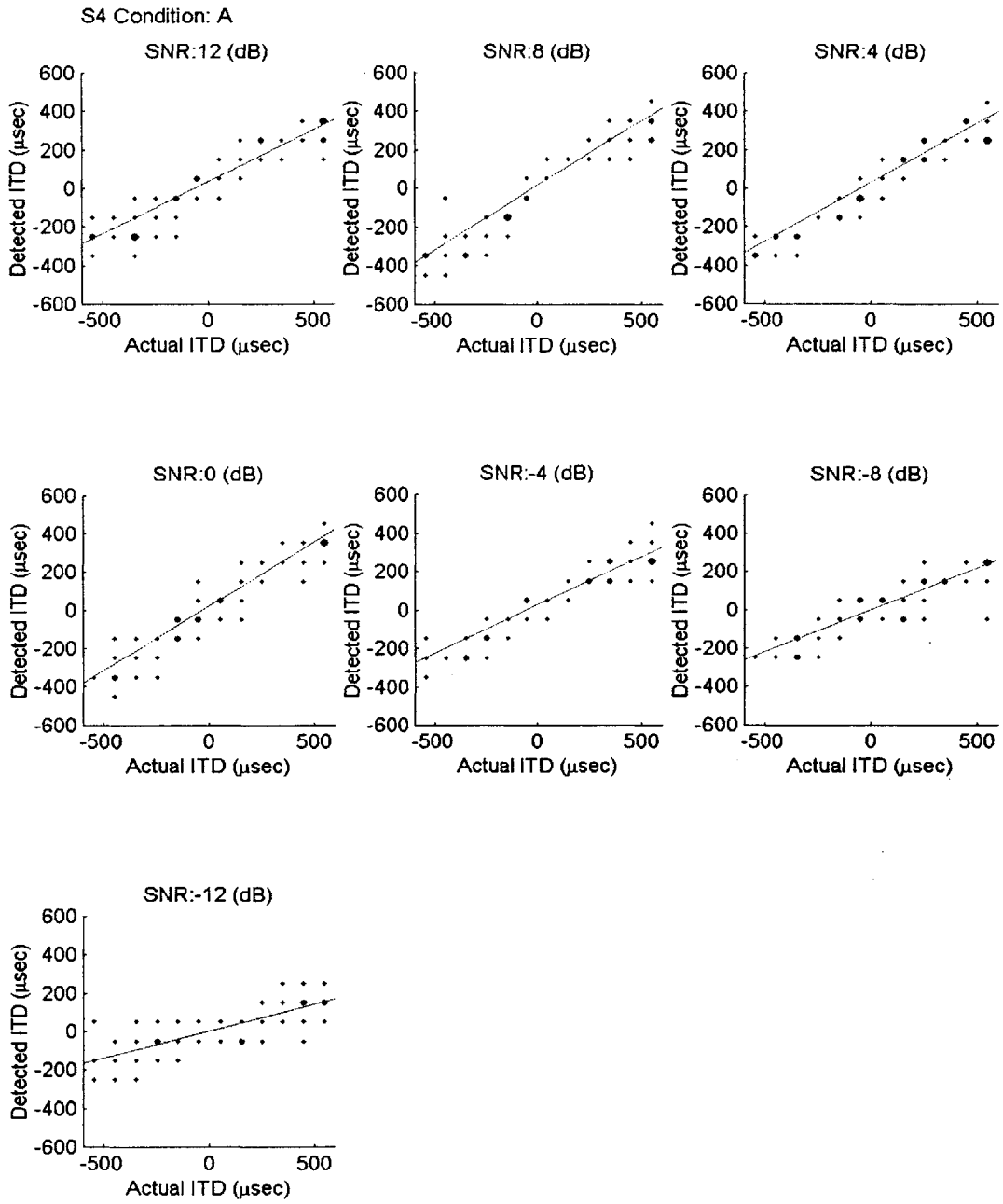
S4 Condition: A
SNR:12 (dB)
SNR:8 (dB)
SNR:4 (dB)
SNR:0 (dB)
SNR:-4 (dB)
SNR:-8 (dB)
SNR:-12 (dB)
Detected ITD (μsec)
Actual ITD (μsec)
600
400
200
0
-200
-400
-600
-500
0
500

Fig. A- 24

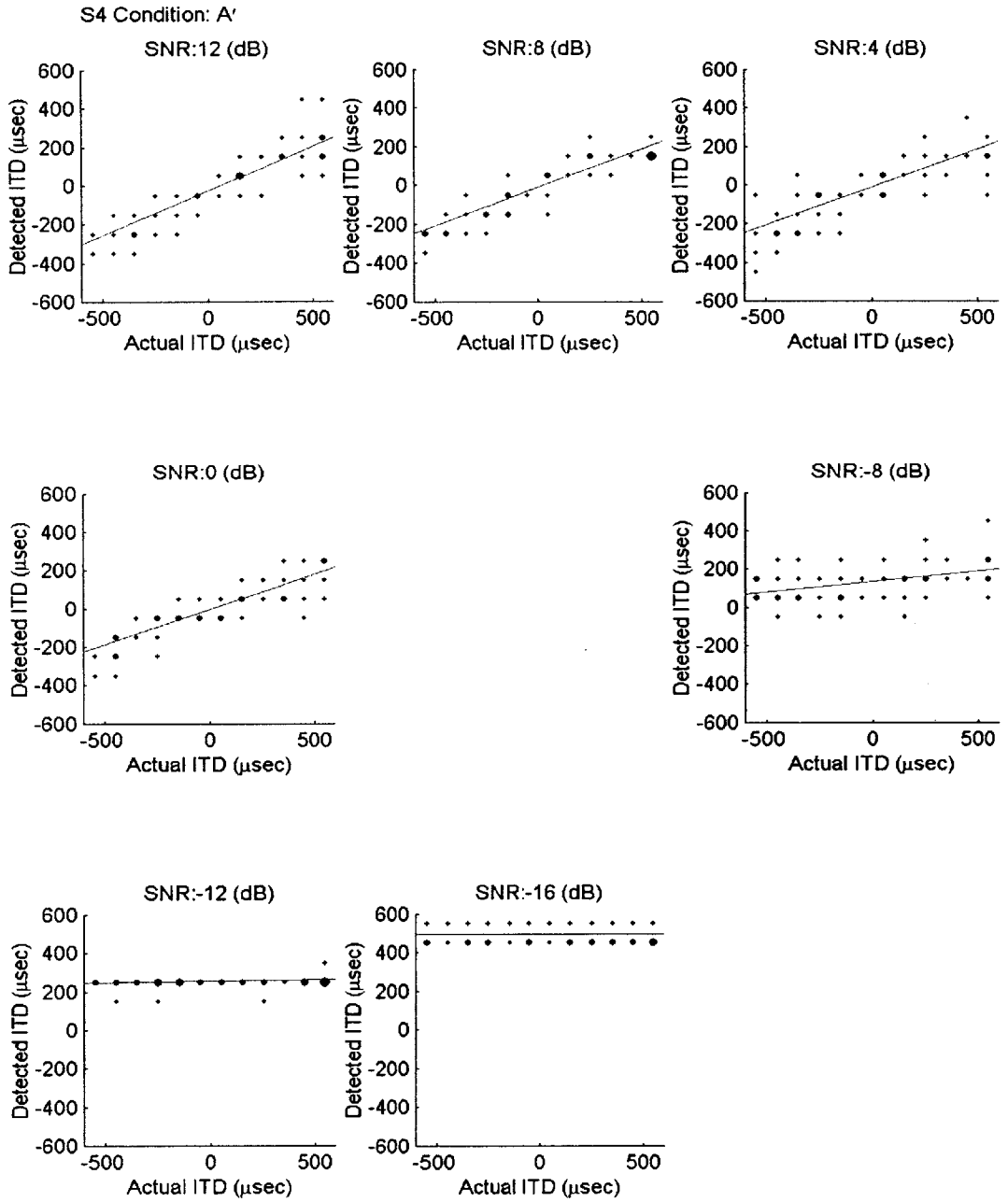
S4 Condition: A'
SNR:12 (dB)
SNR:8 (dB)
SNR:4 (dB)
SNR:0 (dB)
SNR:-8 (dB)
SNR:-12 (dB)
SNR:-16 (dB)
Detected ITD (μsec)
Actual ITD (μsec)
600
400
200
0
-200
-400
-600
-500
0
500

Fig. A- 25

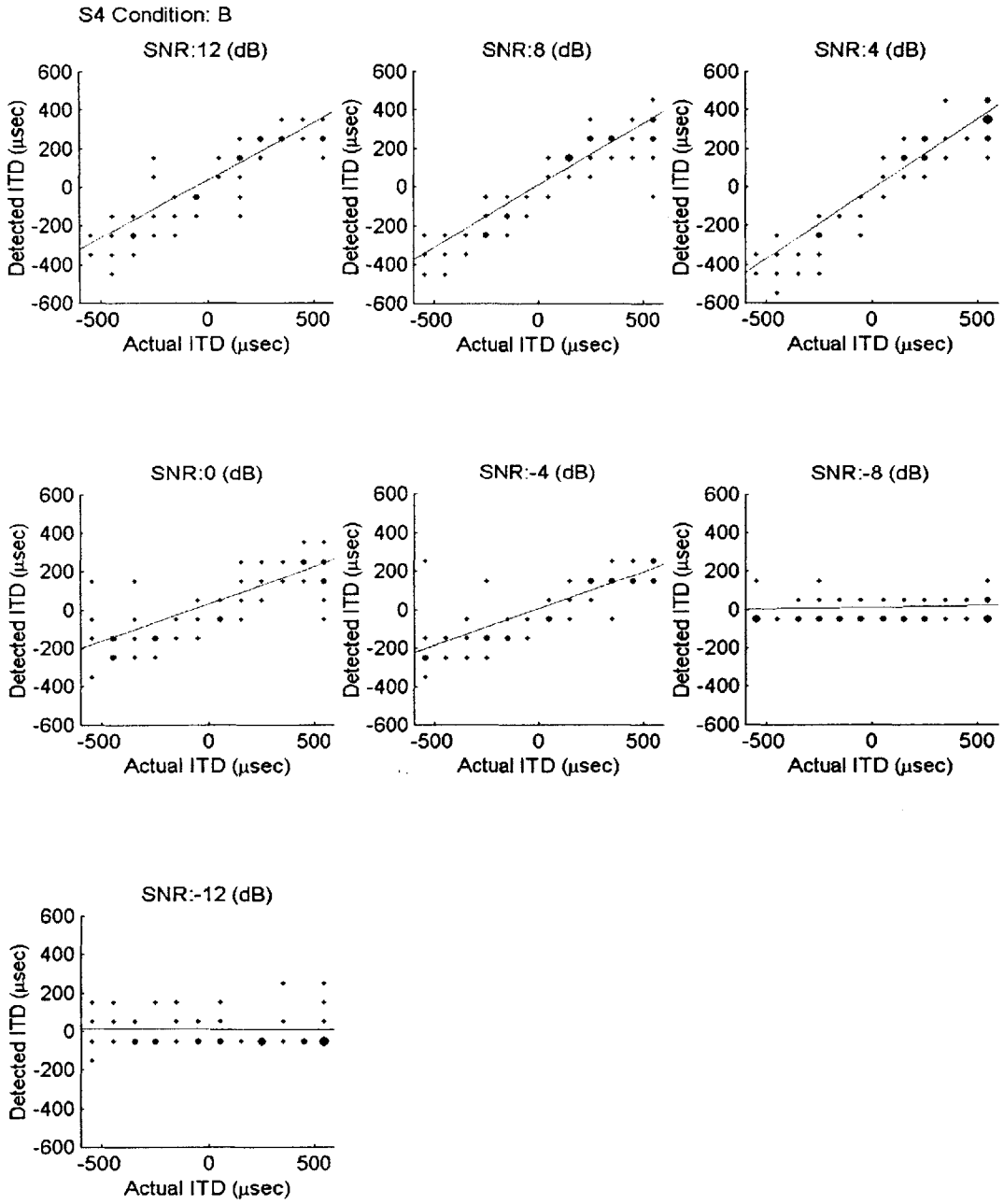
S4 Condition: B
SNR:12 (dB)
SNR:8 (dB)
SNR:4 (dB)
SNR:0 (dB)
SNR:-4 (dB)
SNR:-8 (dB)
SNR:-12 (dB)
Detected ITD (μsec)
Actual ITD (μsec)
600
400
200
0
-200
-400
-600
-500
0
500

Fig. A- 26

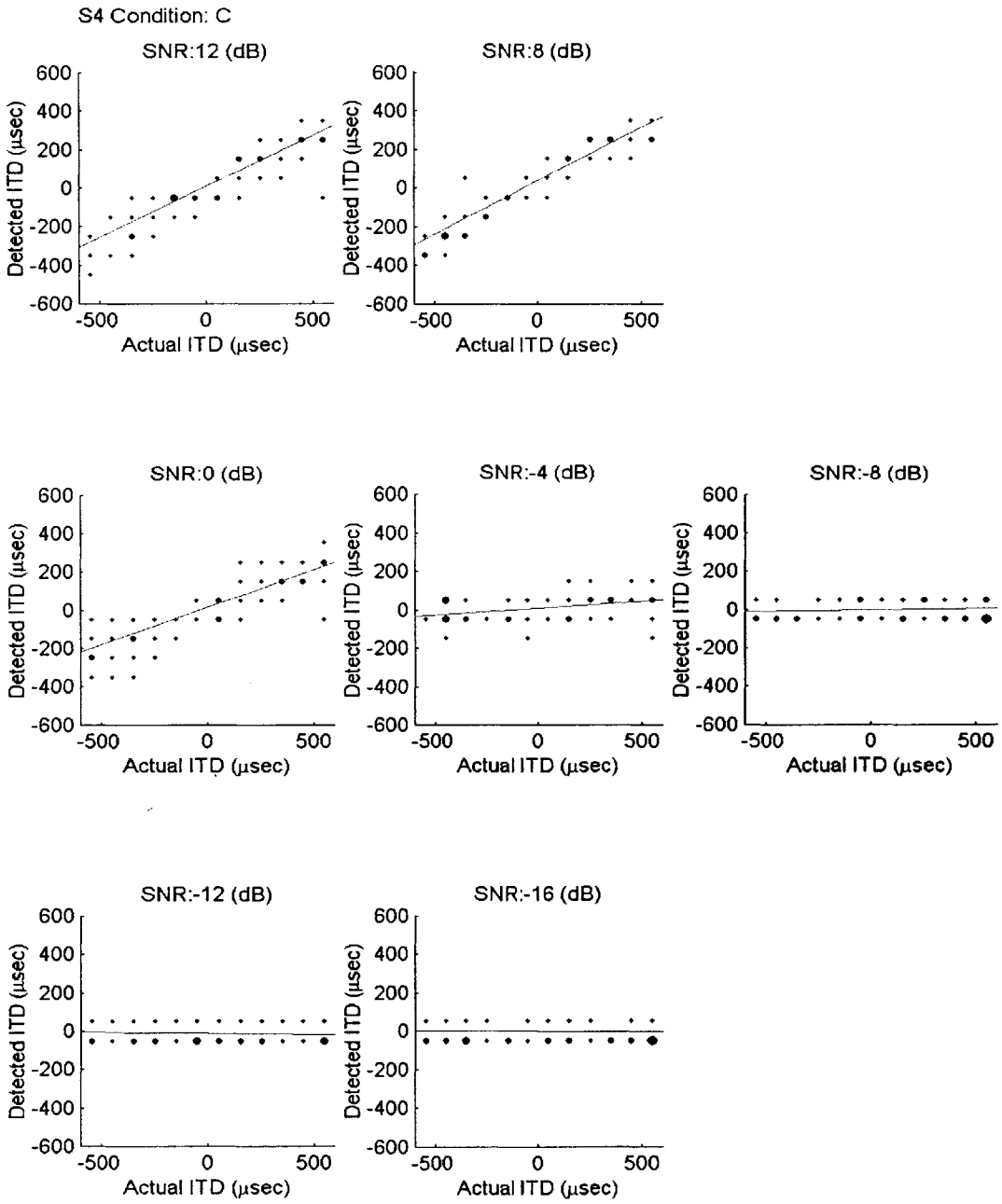
S4 Condition: C
SNR:12 (dB)
SNR:8 (dB)
SNR:0 (dB)
SNR:-4 (dB)
SNR:-8 (dB)
SNR:-12 (dB)
SNR:-16 (dB)
Detected ITD (μsec)
Actual ITD (μsec)
600
400
200
0
-200
-400
-600
-500
0
500

Fig. A- 27

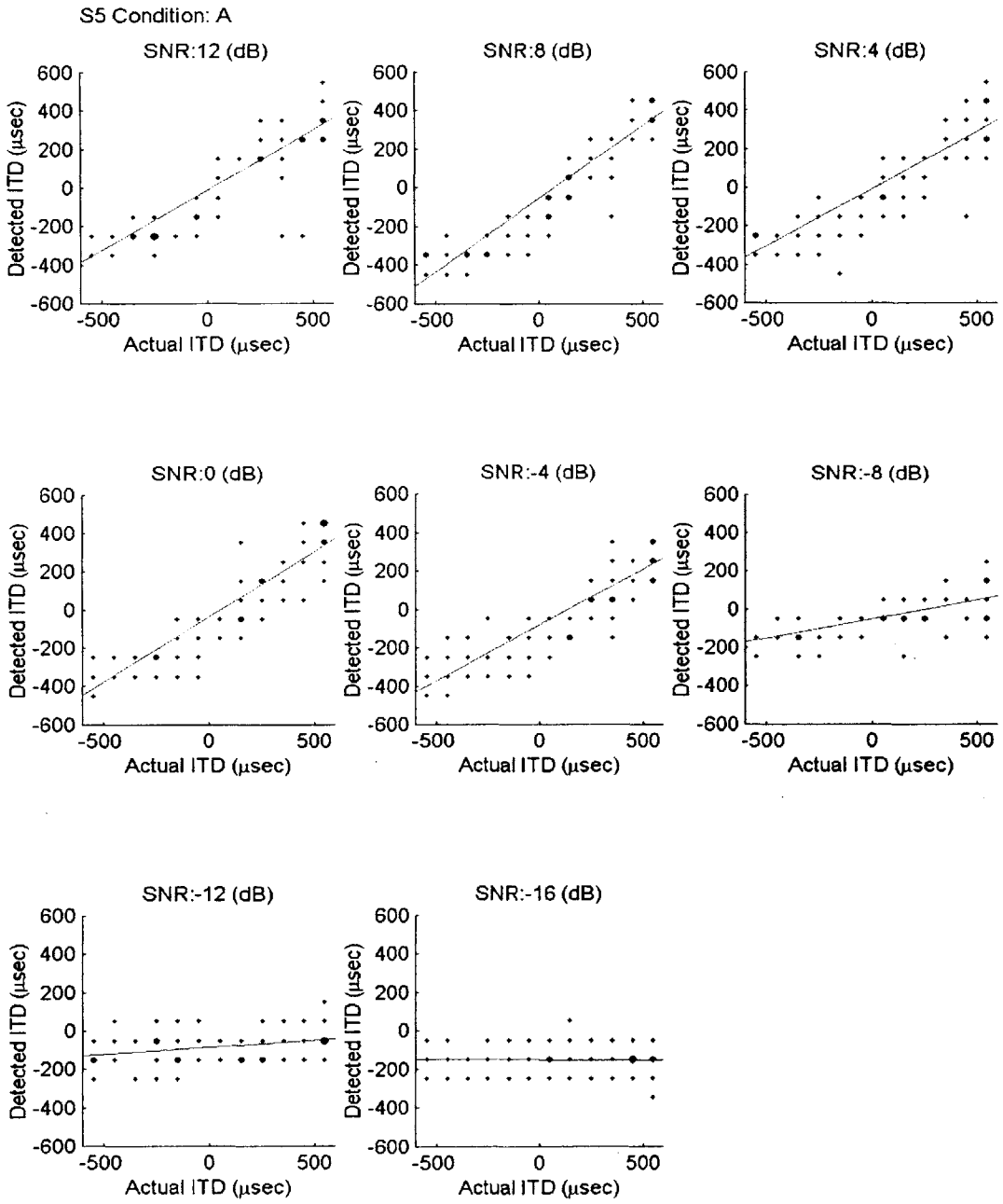
S5 Condition: A
SNR:12 (dB)
SNR:8 (dB)
SNR:4 (dB)
SNR:0 (dB)
SNR:-4 (dB)
SNR:-8 (dB)
SNR:-12 (dB)
SNR:-16 (dB)
Detected ITD (μsec)
Actual ITD (μsec)
600
400
200
0
-200
-400
-600
-500
0
500

Fig. A- 28

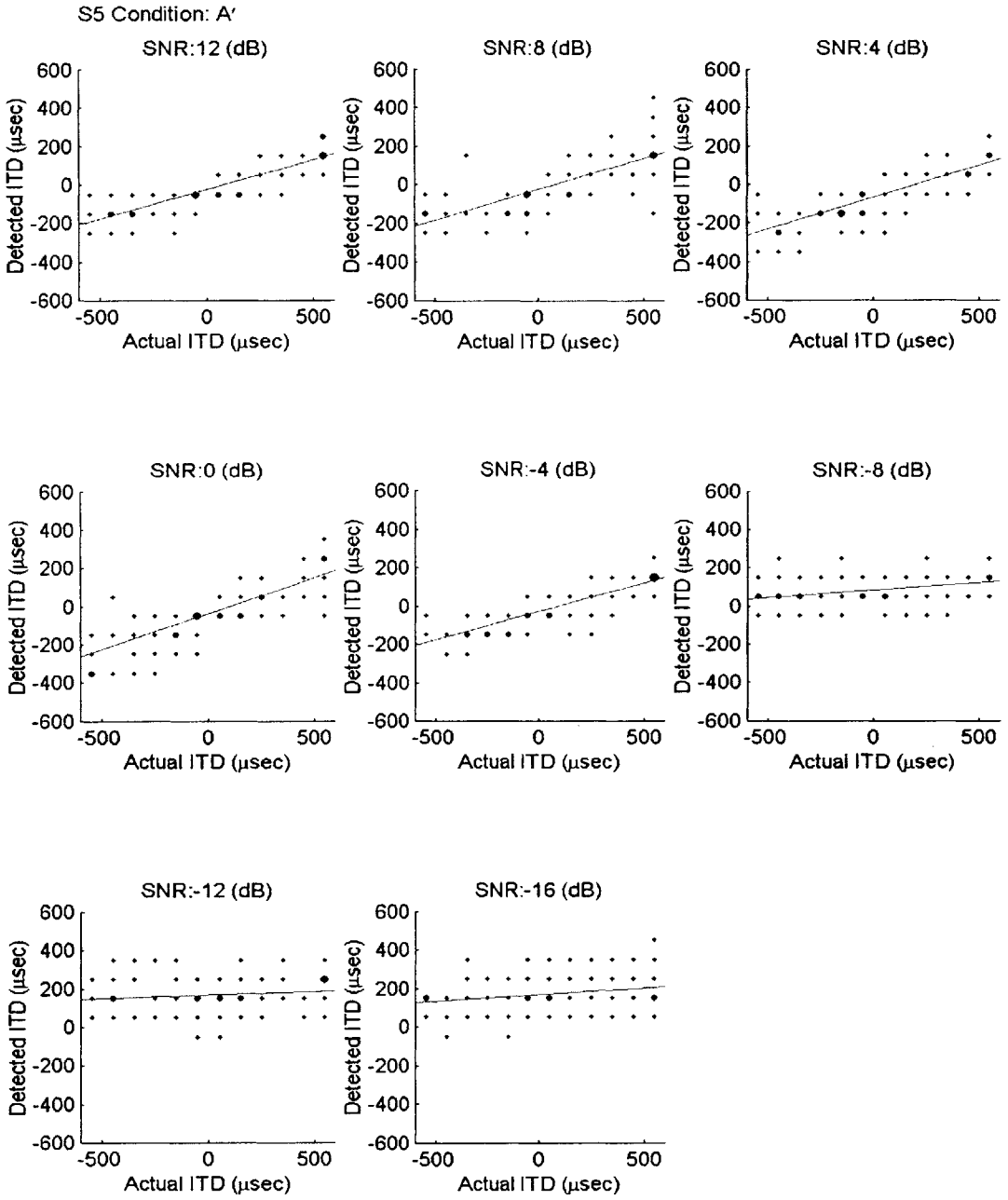
S5 Condition: A'
SNR:12 (dB)
SNR:8 (dB)
SNR:4 (dB)
SNR:0 (dB)
SNR:-4 (dB)
SNR:-8 (dB)
SNR:-12 (dB)
SNR:-16 (dB)
Detected ITD (μsec)
Actual ITD (μsec)
600
400
200
0
-200
-400
-600
-500
0
500

Fig. A- 29

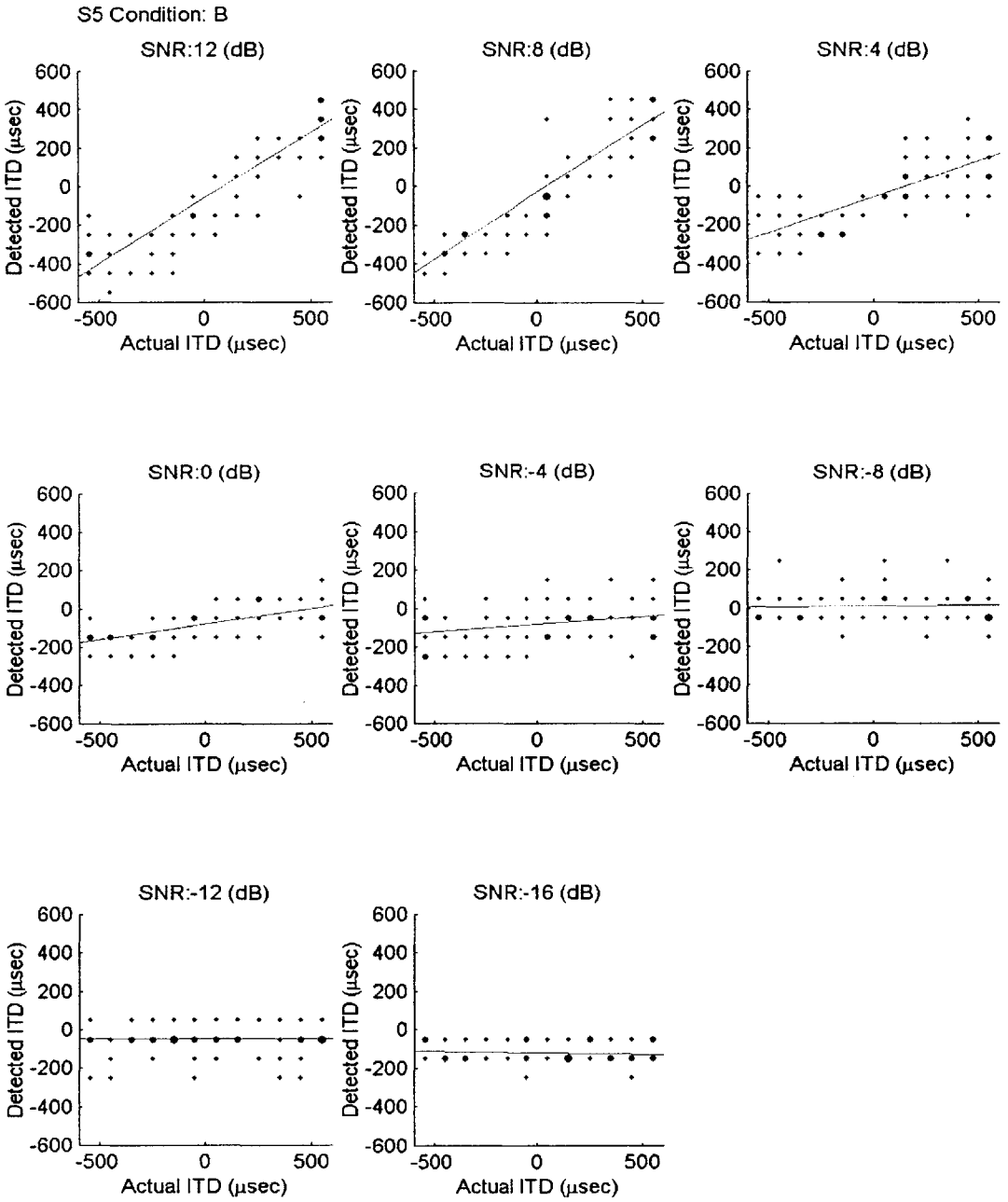
S5 Condition: B
SNR:12 (dB)
SNR:8 (dB)
SNR:4 (dB)
SNR:0 (dB)
SNR:-4 (dB)
SNR:-8 (dB)
SNR:-12 (dB)
SNR:-16 (dB)
Detected ITD (μsec)
Actual ITD (μsec)
600
400
200
0
-200
-400
-600
-500
0
500

Fig. A- 30

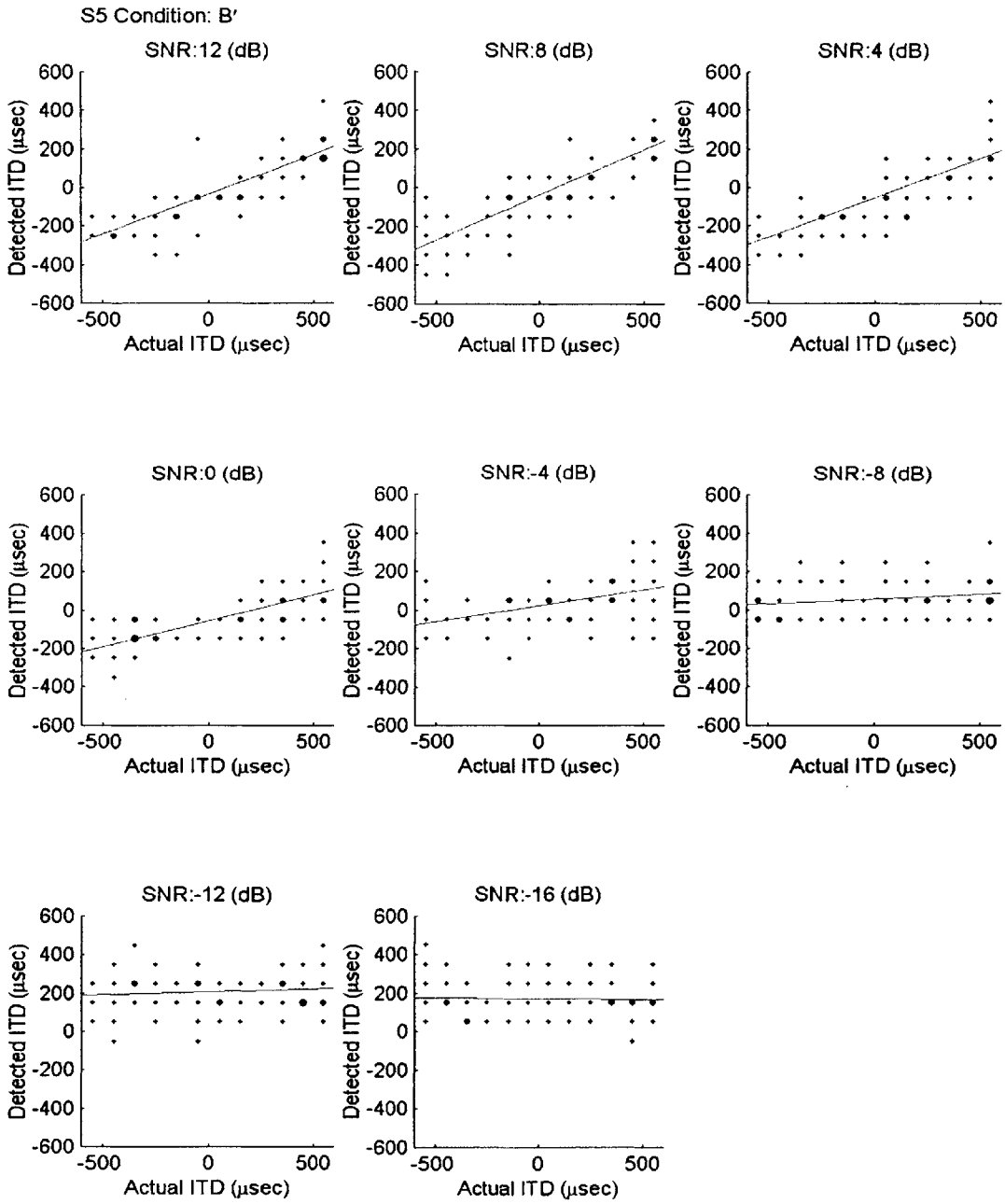
S5 Condition: B′
SNR:12 (dB)
SNR:8 (dB)
SNR:4 (dB)
SNR:0 (dB)
SNR:-4 (dB)
SNR:-8 (dB)
SNR:-12 (dB)
SNR:-16 (dB)
Detected ITD (μsec)
Actual ITD (μsec)
600
400
200
0
-200
-400
-600
-500
0
500

Fig. A- 31

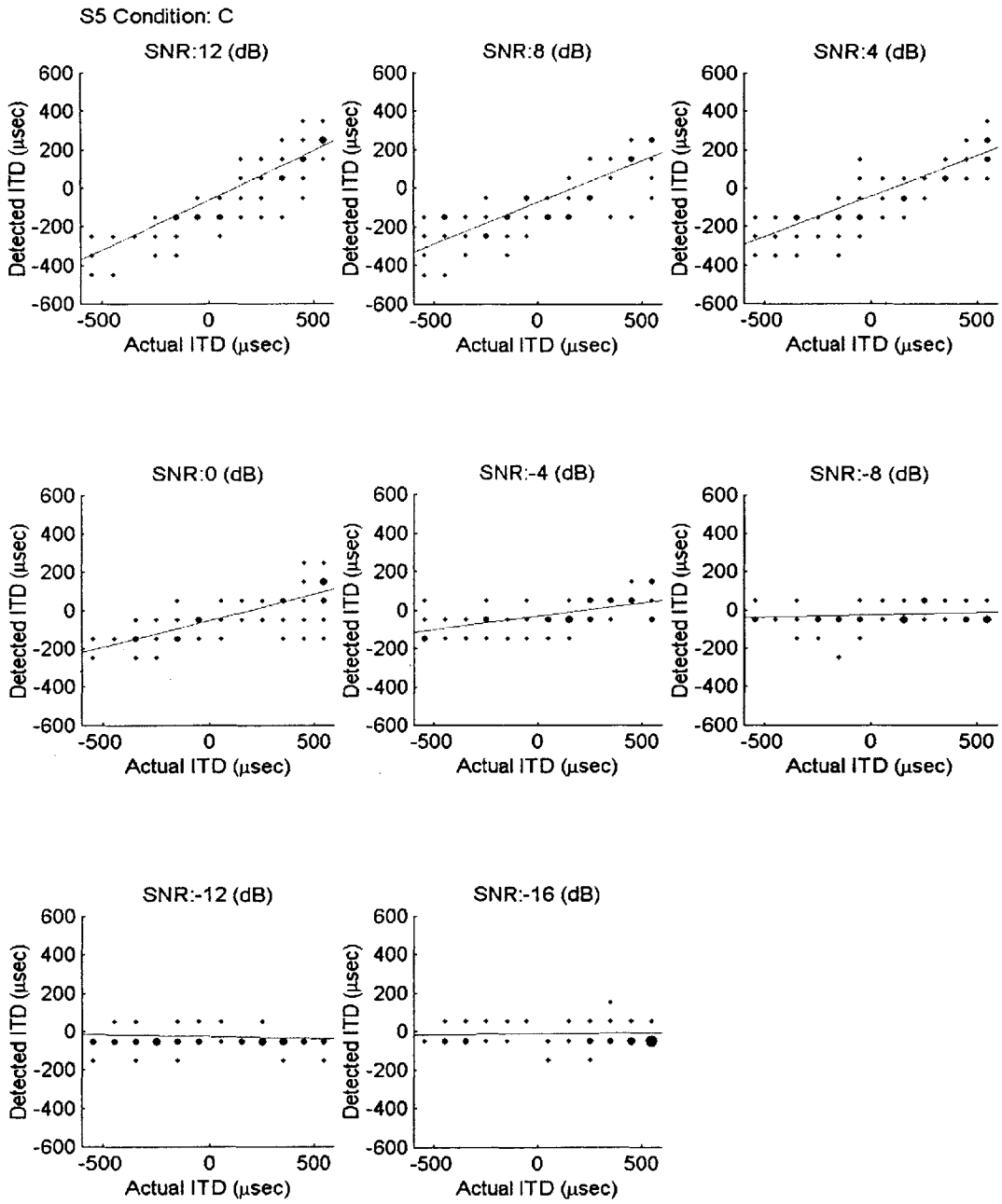
S5 Condition: C
SNR:12 (dB)
SNR:8 (dB)
SNR:4 (dB)
SNR:0 (dB)
SNR:-4 (dB)
SNR:-8 (dB)
SNR:-12 (dB)
SNR:-16 (dB)
Detected ITD (μsec)
Actual ITD (μsec)
600
400
200
0
-200
-400
-600
-500
0
500

Fig. A- 32

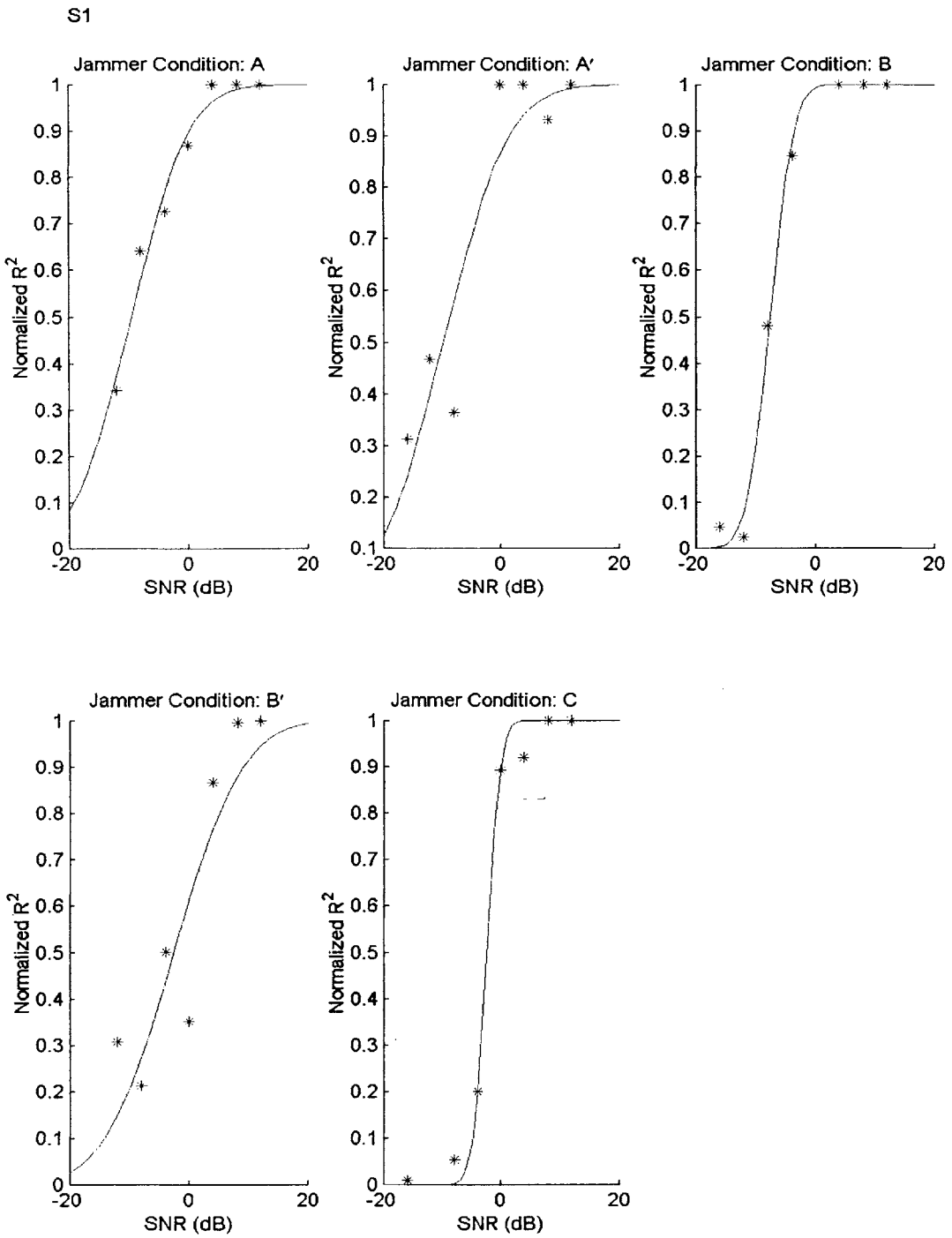
S1
Jammer Condition: A
Jammer Condition: A′
Jammer Condition: B
Jammer Condition: B′
Jammer Condition: C
Normalized R2
SNR (dB)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
-20
0
20

Fig. A- 33

S3

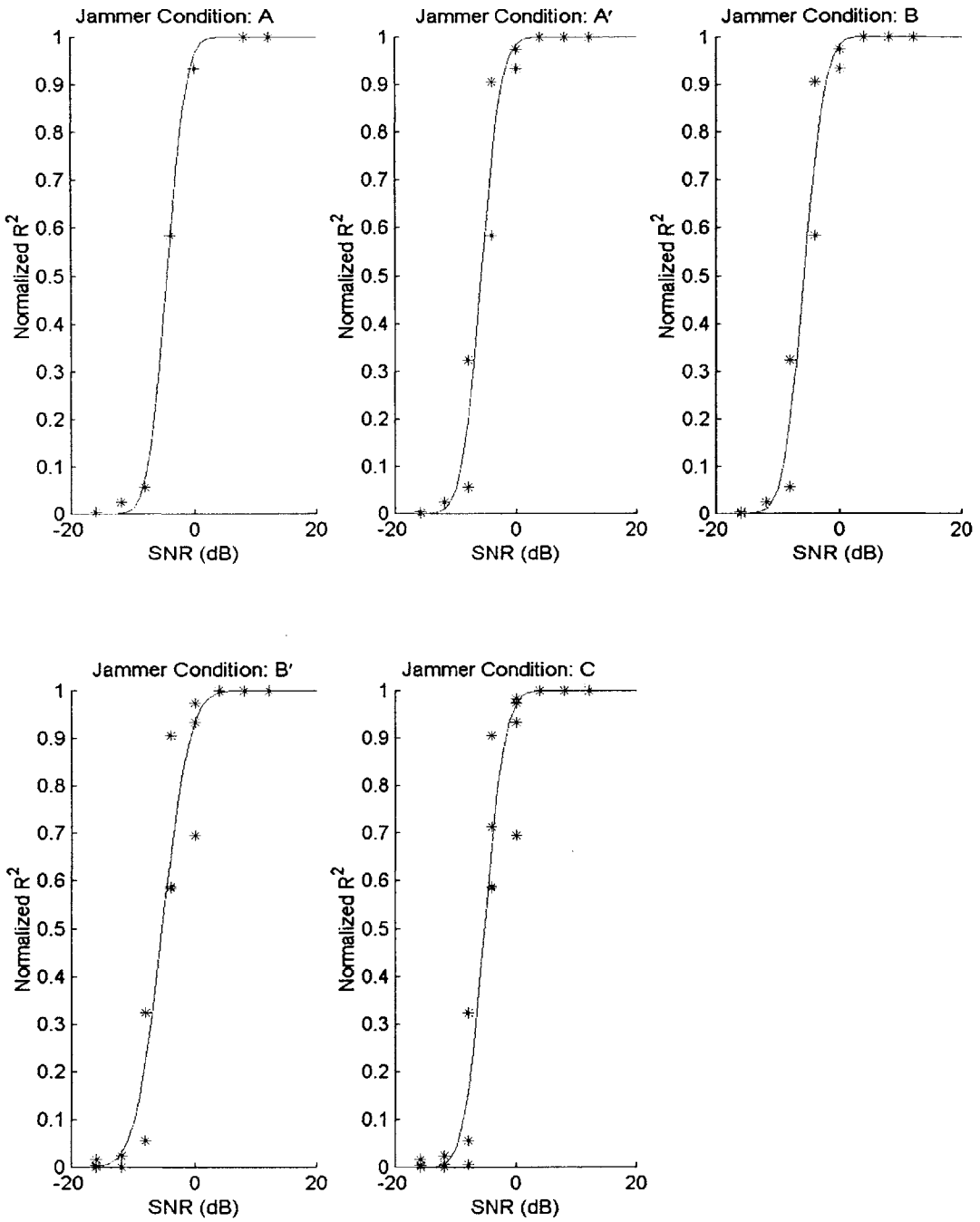
Jammer Condition: A
Jammer Condition: A'
Jammer Condition: B
Jammer Condition: B'
Jammer Condition: C
Normalized $R^2$
SNR (dB)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
-20
0
20

Fig. A- 34

S4

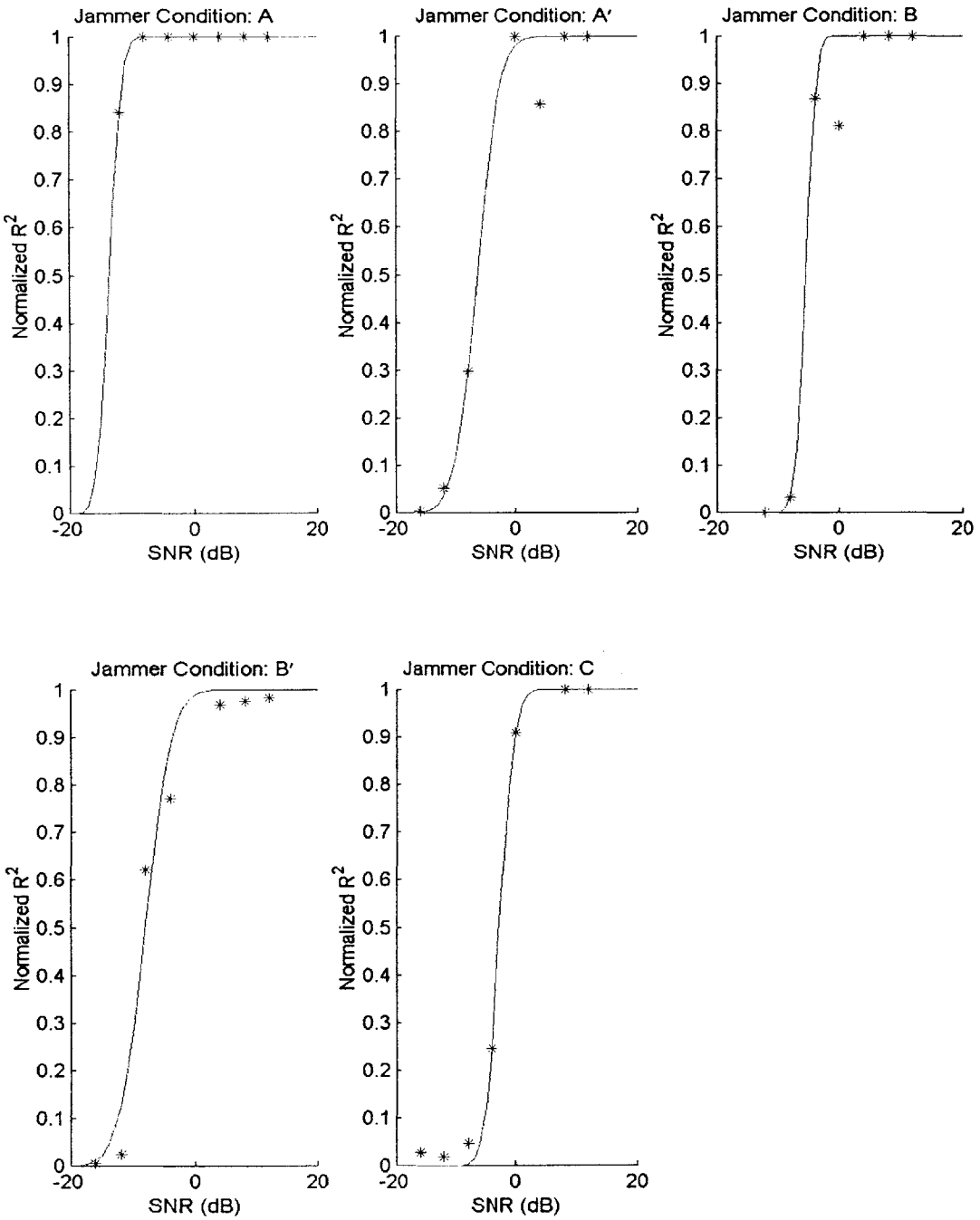
Jammer Condition: A
Jammer Condition: A'
Jammer Condition: B
Jammer Condition: B'
Jammer Condition: C
Normalized $R^2$
SNR (dB)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
-20
0
20

Fig. A- 35

S5

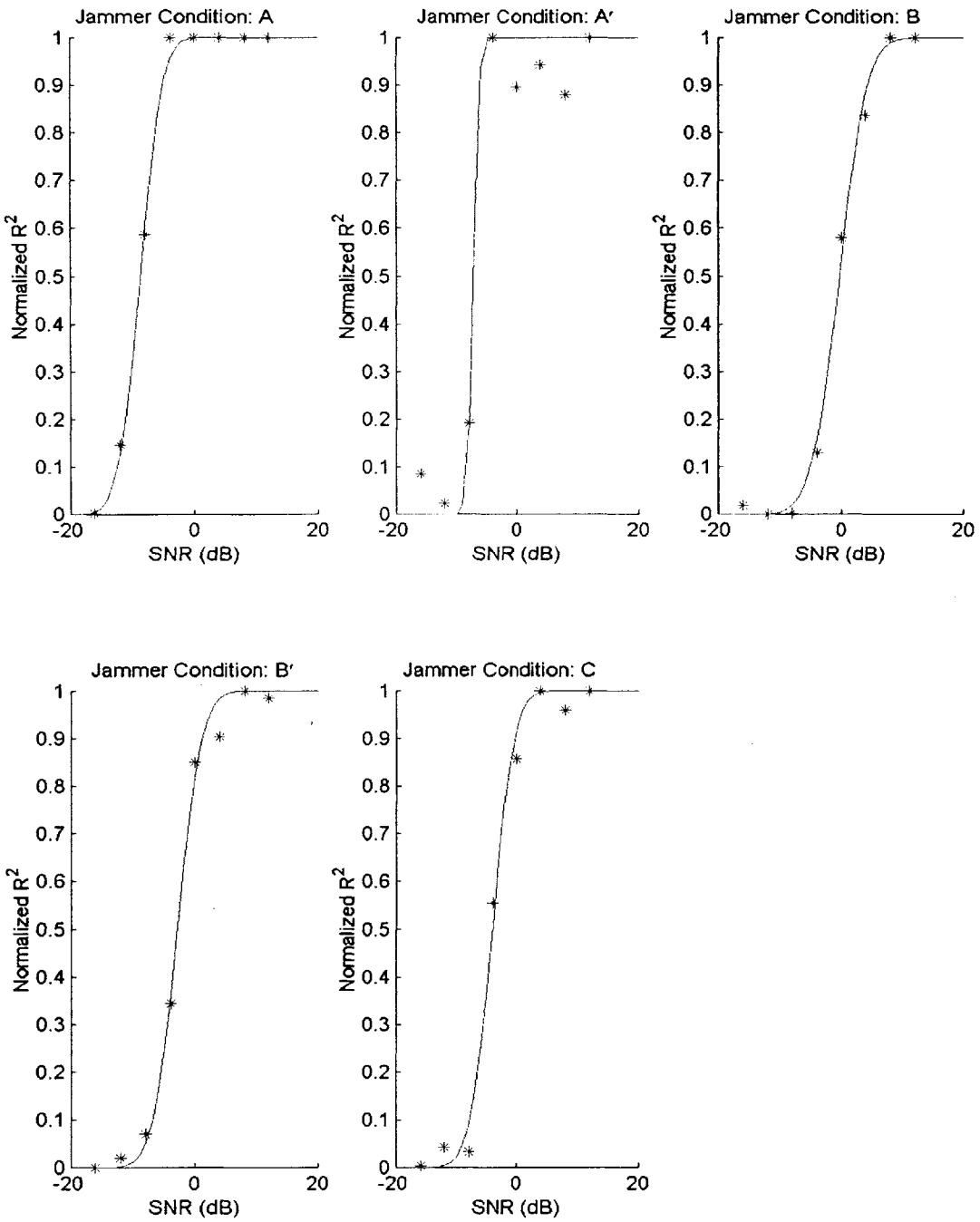
Jammer Condition: A
Jammer Condition: A′
Jammer Condition: B
Jammer Condition: B′
Jammer Condition: C
Normalized R²
SNR (dB)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
-20
0
20

Fig. A- 36

## References


Gardner, B. and K. Martin (1994). "Anechoic KEMAR HRTF recordings." Media Lab Machine Listening Group of the Massachusetts Institute of Technology.

Giguere, C. and S. M. Abel (1993). "Sound localization: effects of reverberation time, speaker array, stimulus frequency, and stimulus rise/decay." J Acoust Soc Am **94**(2 Pt 1): 769-76.

Good, M. D. and R. H. Gilkey (1996). "Sound localization in noise: the effect of signal-to-noise ratio." J Acoust Soc Am **99**(2): 1108-17.

Hartmann, W. M. (1983). "Localization of sound in rooms." J Acoust Soc Am **74**(5): 1380-91.

Kuhn, G. F. (1977). "Model for the interaural time differences in the azimuthal plane." J Acoust Soc Am **62**(1): 157-167.

Litovsky, R. Y., H. S. Colburn, W. A. Yost and S. J. Guzman (1999). "The precedence effect." J Acoust Soc Am **106**(4 Pt 1): 1633-54.

Lorenzi, C., S. Gatehouse and C. Lever (1999). "Sound localization in noise in normal-hearing listeners." J Acoust Soc Am **105**(3): 1810-20.

Martin, K. (1993). A Computational Model of Spatial Hearing. Electrical Engineering. Cambridge, Massachusetts Institute of Technology: 62.

Mills, A. W. (1958). "On the minimum audible angle." J. Acoust. Soc. Am. **30**: 237-246.

Rakerd, B. and W. M. Hartmann (1985). "Localization of sound in rooms, II: The effects of a single reflecting surface." J Acoust Soc Am **78**(2): 524-33.

Vedula, M. (2000). Localization of Sound Sources in the Horizontal Plane Using Interaural Time Difference. Electrical Engineering. Boston, Boston University: 87.

Wightman, F. L. and D. J. Kistler (1992). "The dominant role of low-frequency interaural time differences in sound localization." J Acoust Soc Am **91**(3): 1648-61.

Zurek, P. M. (1987). The precedence effect. Directional Hearing. W. A. Yost and G. Gourevitch. New York, Springer-Verlag: 85-105.