Analysis of early auditory evoked potentials elicited by stimuli with directional information

Helmut Riedel

geb. am 23. Juli 1963

in Lübeck

Contents

Chapter 1

General introduction

The ability to localize sounds is one of the most intriguing capabilities of the auditory system. The amazement about this feature is emphasized by a comparison with the visual system. Adjacent points in the visual field are represented by the activity of vicinal cells in the retina and higher brain structures and therefore naturally stage a topographical mapping. In contrast, the solely information available to the auditory system is the sound pressure fluctuation at both eardrums. These two signals suffice to receive an acoustic impression of the surrounding three-dimensional space.

A large body of psychophysical and physiological studies has thrown light on the principles of binaural signal processing in the auditory system (for reviews see, e.g., Yost and Gourevitch, 1987; Edelman et al., 1988; Webster et al., 1992; Popper and Fay, 1992; Moore, 1995). The most important cues used in spatial hearing are the interaural differences in time and level (Jeffress and McFadden, 1971; Durlach and Colburn, 1978). The first stage along the auditory pathway, where these differences are processed, is the brain stem (Irvine, 1992).

Various noninvasive methods are available to investigate the living human brain. Positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) provide a good spatial resolution to map active brain areas, but lack to have a sufficient time resolution for the analysis of binaural hearing. On the other hand, magnetoencephalography (MEG) and electroencephalography (EEG) achieve an arbitrarily good time resolution, but attain only a limited spatial resolving power. Unfortunately, none of the methods developed so far simultaneously allows both good spatial and temporal resolution. The activity of deep, central brain structures is hardly detectable with MEG which is only sensitive to tangential, but blind to radial currents. Therefore, EEG appears to be the

method of choice for the analysis of the first stages of human binaural processing.

From the discovery of spontaneous electric brain signals (Berger, 1929) it took more than two decades until signal averagers allowed the acquisition of cortical event related potentials (Dawson, 1951, 1954). Another two decades later the auditory brain stem responses (ABRs) were detected (Jewett et al., 1970; Jewett and Williston, 1971). Thousands of stimulus repetitions are necessary to uncover the tiny ABR, usually exhibiting amplitudes smaller than a microvolt, from the background noise.

The objective of the present thesis is to elucidate a neural representation or correlate of spatial hearing. This aim is approached by the analysis of binaural ABRs and so-called binaural difference potentials (BDs) to stimuli with interaural disparities. The BD is computed as the difference between binaural and summed monaural response and believed to represent a correlate of the specific binaural processing of a stimulus. Its signal-to-noise ratio (SNR) is about an order of magnitude smaller than that of the ABR. Inadequate control of the residual noise and deficient estimation of the SNR led to conflicting findings in BD studies. The current dissertation therefore concentrates on methods to overcome these methodological problems and to provide reliable data about binaural neural processing in the human brain stem.

In chapter 2 the advantage of using information from the raw data of a measurement to accurately estimate the residual noise and the SNR of an evoked potential is demonstrated. Twenty years ago it was necessary to rely on online-averagers discarding the single sweeps during recording. However, with the computer facilities available nowadays it is relatively easy to record the complete raw data of an ABR measurement. Apparently, for historical reasons this is not commonly done. Based on the noise information gathered from the raw data, different methods to build the evoked potential from the single sweeps are compared. An optimal averaging method is proposed by means of a simulation study.

ABRs to binaurally presented clicks with different ITDs and ILDs are analyzed in chapter 3. The effects of synergistic and antagonistic combinations of the interaural parameters are also described. It is investigated if characteristics of the waveforms reflect properties of the stimuli. This allows to determine if the evoked potentials rather correspond to the interaural cues or more to the lateralization caused by the cues.

Chapter 4 deals with dipole source analysis of multi-channel EEG data. The methodological improvements introduced in chapter 2 directly result in refinements in the algorithms used for the solution of the inverse problem. The inclusion of the noise covariance matrix gained from the data allows generalized maximum likelihood estimation instead of least-

squares fitting. Furthermore, confidence regions for dipole locations and moments and the goodness-of-fit can be accessed accurately. However, these improvements cannot vanquish the non-uniqueness of the inverse problem already mentioned by Helmholtz (1853). The use of physiological a-priori information is believed to be an appropiate method to overcome this ambiguity. Three alternative source models are presented and compared in their capability to explain the multi-channel ABR data.

The click is the traditional stimulus in ABR recordings. Contributions to the click-evoked ABR are known to originate predominantly from high frequencies. Compared to the click, a chirp with rising instantaneous frequency allows a more synchronous activation of auditory fibers over the entire frequency range and leads to enlarged monaural ABR (Dau et al., 2000). The fifth chapter focuses on the comparison of clicks and chirps in their respective capability of eliciting binaurally evoked ABRs and binaural difference potentials.

Finally, the last chapter gives a summary, some concluding remarks and an outlook to the intended future research in the field of binaurally evoked potentials.

Chapter 2

Analysis of averaging methods to improve the quality of auditory brain stem responses[1]

Abstract

A systematic evaluation of the influence of different averaging methods on the quality of recorded auditory brain stem responses is performed on a single sweep basis, i.e., on a post-hoc analysis of all unaveraged single epochs. The question of an optimal averaging method is addressed. Single sweeps of auditory brain stem responses were recorded to monaural and binaural click stimuli at levels of 20, 40 and 60 dB nHL. Recording sites were both mastoids and the forehead (Fz), the vertex (Cz) served as common reference electrode. Five averaging methods were applied to the same set of data to compare their capability to reduce the residual noise. In addition, the method of iterative averaging is introduced, which relies on an improved estimation of the noise of single epochs. A simulation allows a verification of the quality of the different averaging methods and estimators of signal, residual noise, and signal-to-noise-ratio given by those methods. It is shown that single-sweep-based estimation of residual noise and signal-to-noise ratio is superior to average-based estimation. Weighted averaging with one iteration step is the most favourable averaging method with respect to minimum residual noise and valid

---

[1]A modified version of this chapter was published in Z. Audiol. 40 (2), 62–85: Riedel et al. (2001): "Single-sweep-based methods to improve the quality of auditory brain stem responses. Part II: averaging methods"

estimation of the signal. For reliable quality estimation, single-sweep-based methods are preferable. Weighted averaging using iteration not only provides reliable signal and noise estimates, but also overcomes the arbitrariness of an artifact threshold.

2.1 Introduction

In the last two decades, recording of auditory brain stem responses (ABRs) has developed into a standard method for differential diagnostics of the ascending auditory pathway. Hence, high quality requirements have to be fulfilled within a minimum time to measure the responses. In the clinical routine, artifact rejection is the common technique to increase the signal-to-noise ratio (SNR) of the measurement (Gevins and Rémond, 1987, chap. 5).

The quality of ABR measurements is often only assessed by visual inspection of the responses averaged into two buffers. SNR estimates are provided using the sum and the difference of two averages as signal and noise estimate, respectively (Schimmel, 1967; Wong and Bickford, 1980). However, the estimation of the residual noise can be improved by the calculation of the so-called single-point variance, i.e., the variance across sweeps for a fixed time point (Elberling and Don, 1984; Don et al., 1984). Recently, Cebulla et al. (2000a,b), using Monte-Carlo simulations, showed that further improvement of the estimate of the residual noise can be achieved by calculating the single point variance for every time sample and averaging over time, if the number of sweeps is low.

Here, the quality of ABRs is investigated if not only one or two averages but all single epochs[2] of a recording are available. Although such methods imply higher computational cost, they provide the possibility to decide with optimized post-hoc criteria if each individual epoch should be included or excluded from the average and what kind of weighting should be used. With increased availability of high computing capacities we do not regard computational cost or online-realisation as an issue here.

There are typically two methods used to improve the SNR of the recorded signals: filtering and averaging. Digital linear-phase filters were investigated by Granzow et al. (2001). Here it is shown both theoretically and experimentally that the estimation of the residual noise based on single sweeps is superior to an estimation based on two averages.

In addition to the commonly used averaging method using an artifact criterion, alternative techniques are investigated: sorted averaging (Mühler and von Specht, 1999), weighted

---

[2]The terms 'sweep' and 'epoch' are used interchangeably.

averaging (Hoke et al., 1984; Lütkenhöner et al., 1985) and averaging by means of Bayesian inference (Elberling and Wahlgreen, 1985), which we refer to as block-weighted averaging.

Using an artifact threshold has a major drawback: it must be known prior to the measurement. In actual practice, researchers and clinicians rely on their experience and choose a value that has proven reasonable in the past. However, since background noise varies significantly between subjects, it is not always possible (and certainly not reasonable) to use a fixed artifact threshold. The researcher will notice after a while whether the value chosen is adequate for the subject under study and will adjust it accordingly. Clearly, this is a very unsatisfactory situation because the value of the artifact threshold is arbitrary and the result is not reproducible.

Mühler and von Specht (1999) have suggested the method of sorted averaging on a single sweep basis to determine the sweeps entering an average *a posteriori*. Before averaging, the recorded epochs are sorted according to their (estimated) contamination with noise. Only sweeps containing less than a certain degree of noise will be included in the average.

The method of weighted averaging (Hoke et al., 1984; Lütkenhöner et al., 1985) is an extension of the former techniques allowing an assignment of continuous positive weightings to individual epochs. The weightings are chosen according to the extent of contamination of the sweeps with noise. This should be useful in the case of non-stationarity of the EEG background noise, which can arise, for example, when the subject moves or his muscular activity changes.

The scheme of block weighted averaging (Elberling and Wahlgreen, 1985) tries to circumvent the problem of estimating the noise power of a single sweep by forming blocks of sweeps. From these blocks a more accurate noise power estimate can be obtained. This is important since a good estimate of the noise power of a single epoch (or a block of epochs) is essential for the weighted averaging schemes.

The goal of an improved noise estimate leads to the new technique of iterative averaging, which is introduced and assessed here in comparison to the techniques mentioned so far. The idea is to subtract the current signal estimate from each epoch to estimate the respective noise component. This estimate is used in the subsequent iteration step for a weighted average and results in a more accurate signal estimate.

Finally, the properties of the investigated averaging methods and SNR estimates are validated by a simulation study using a known signal and recordings of noise in the no-stimulus condition.

2.2 Methods

2.2.1 Subjects, stimuli and recordings

Nine male subjects aged from 25 to 35 years participated voluntarily in this study. They were clinically classified as normal hearing and had no history of audiological or neurological problems.

Rare-faction click stimuli were produced by applying rectangular voltage pulses of $100\,\mu$s duration to Etymotic Research ER-2 insert phones. The time interval between the onset of successive clicks was uniformly distributed between 60 and 80 ms, yielding an average stimulation rate of approximately 14.3 clicks per second.

ABRs were recorded from the left mastoid (M1), the right mastoid (M2), and the forehead (Fz) with respect to the common reference electrode at the vertex (Cz). Responses to monaural left, monaural right, and binaural stimulation at levels of 20, 40 and 60 dB normal hearing level were recorded for all subjects. For the simulation study, no-stimulus recordings were also made. For every stimulus condition $J = 10.000$ individual sweeps were collected and stored to hard disk.

2.2.2 Averaging methods

The processing of the raw data primarily comprises linear filtering and averaging. If no weighting or artifact rejection is applied, the order of filtering and averaging is interchangeable due to the linearity of both operations. Since a high DC value or drift of the epochs can thwart any meaningful weighting, all single epochs were filtered before a decision about exclusion or assignment of weightings was made. An FIR-bandpass-filter with 200 taps – designed with the window design method using a Hamming window – with corner frequencies 50 and 1500 Hz was used.

The averaging methods considered here can commonly be expressed to yield a signal estimate $s(t)$ by forming a weighted average of J (filtered) epochs $x_j(t)$ where t denotes the time:

$$s(t) = \frac{\sum_{j=1}^{J} w_j x_j(t)}{\sum_{j=1}^{J} w_j} \; . \tag{2.1}$$

The averaging methods differ in their strategies of assigning the weightings w_j to the epochs $x_j(t)$. The most simple average is obtained by setting $w_j = 1$ for all epochs. This is called the conventional average $s_c(t)$:

$$s_c(t) = \frac{1}{J} \sum_{j=1}^{J} x_j(t) \,.$$ (2.2)

Averaging using an artifact criterion

According to this strategy, epochs $x_j(t)$ with a peak-to-peak voltage A_j larger than a certain threshold value A (which has to be specified in advance) are considered non-physiological and are excluded from the average ($w_j = 0$). The remaining $J_a \leq J$ epochs enter the average with $w_j = 1$.

$$s_a(t) = \frac{1}{J_a} \sum_{j=1}^{J_a} x_{j,A_j \leq A}(t) \,.$$ (2.3)

J_a and the average itself critically depend on the choice of the artifact threshold A.

Sorted averaging

The idea of the sorted averaging method (Mühler and von Specht, 1999) is to sort the sweeps according to their contamination with noise and to classify them into two groups: sweeps with a small amount of noise are accepted to enter the average ($w_j = 1$) while sweeps with high noise values are excluded ($w_j = 0$). The critical noise value which separates accepted and rejected epochs is derived by the following consideration. Because in ABR recordings the single sweep SNR is very low (-20 to -30 dB), one can approximate

$$x_j(t) = S(t) + N_j(t) \approx N_j(t) \,, \ \ j = 1 \ldots J \,.$$ (2.4)

Capital letters S and N are used to denote "true" signal and noise quantities, respectively, in contrast to estimates which are designated by lower case letters (s and n). Eq. 2.4 states that the measured and filtered signal primarily consists of noise.

The power[3] P of any discrete signal $x(t)$ of length T is defined as

---

[3]Power is the variance across time. The term 'variance' is used for the statistical variance over the ensemble of epochs.

$$P(x) = \frac{1}{T} \sum_{t=1}^{T} (x(t))^2 \,. \tag{2.5}$$

Now the epochs are sorted in order of increasing power $P(x_j(t)) \equiv P_j$. The noise value dividing accepted and rejected sweeps is determined by minimizing the power of the mean cumulative normalized noise

$$P_{\text{cum}}(J') = \frac{1}{J'(J'-1)} \sum_{j'=1}^{J'} P(N_{j'}) \approx \frac{1}{J'(J'-1)} \sum_{j'=1}^{J'} P_{j'} \,, \tag{2.6}$$

where j' stands for the index of the sorted epochs. If all J' terms of the sum roughly have the same magnitude, the numerator will increase in proportion to J', whereas the denominator increases with J'^2. Consequently, the sum decreases proportionally to $1/J'$. In the case of non-stationary noise, however, not all the terms in the sum in eq. (2.6) have the same magnitude. If the cumulative noise for the inclusion of one sweep after the other is computed, i.e., J' is increased, a minimum can be found for a certain number of sweeps. This only holds if the increase of the noise power caused by the inclusion of a given sweep outweighs the increase of the denominator due to raising J' by one.

For most practical ABR measurements such a minimum can in fact be found. Hence an optimal number J_s of sweeps can be determined. The sorted average therefore becomes

$$s_s(t) = \frac{1}{J_s} \sum_{j'=1}^{J_s} x_{j'}(t) \,. \tag{2.7}$$

Weighted averaging

Hoke et al. (1984) have shown that the highest SNR is obtained if the inverse power of the noise of an epoch is assigned as weighting w_j to sweep $x_j(t)$.

$$w_j = \frac{1}{P(N_j(t))} \approx \frac{1}{P_j} = \frac{T}{\sum_{t=1}^{T} x_j^2(t)} \,, \tag{2.8}$$

where again the approximation of eq. (2.4) has been used. The weighted average over J sweeps is then

$$s_w(t) = \frac{\displaystyle\sum_{j=1}^{J} \frac{x_j(t)}{P_j}}{\displaystyle\sum_{j=1}^{J} \frac{1}{P_j}} \;. \tag{2.9}$$

Weighted averaging has the advantage that it is not necessary to set a somewhat arbitrary artifact threshold.

Block-weighted averaging

The block-weighted averaging technique or method of Bayesian inference was introduced to the field of ABR analysis by Elberling and Wahlgreen (1985). They showed that the power of the noise can better be estimated on the basis of a group or block of sweeps rather than from a single sweep. In this study block sizes $\beta = 1, 2, 4, \ldots, 256$ are compared. $\beta = 1$ represents the weighted average. A block of β consecutive sweeps is averaged conventionally resulting in intermediate 'sweeps' $x_{(b_\beta)}(t)$:

$$x_{(b_\beta)}(t) = \frac{1}{\beta} \sum_{j=(b-1)\beta+1}^{b\beta} x_j(t) \;, \tag{2.10}$$

with the block size β. The various blocks are identified by a subscript b, $b = 1 \ldots B$, and their respective average powers are

$$P_{(b_\beta)} = \frac{1}{\beta} \sum_{j=(b-1)\beta+1}^{b\beta} P_j \;. \tag{2.11}$$

The inverse powers of the quantities $x_{(b_\beta)}(t)$ serve as weightings for a weighted average of blocks of sweeps:

$$s_{b_\beta}(t) = \frac{\displaystyle\sum_{b=1}^{B} \frac{x_{(b_\beta)}(t)}{P_{(b_\beta)}}}{\displaystyle\sum_{b=1}^{B} \frac{1}{P_{(b_\beta)}}} \;. \tag{2.12}$$

The number of epochs entering this average is $J_b = \beta * B$.

The block diagram in Fig. 2.1 gives an overview of the averaging methods used in this paper.
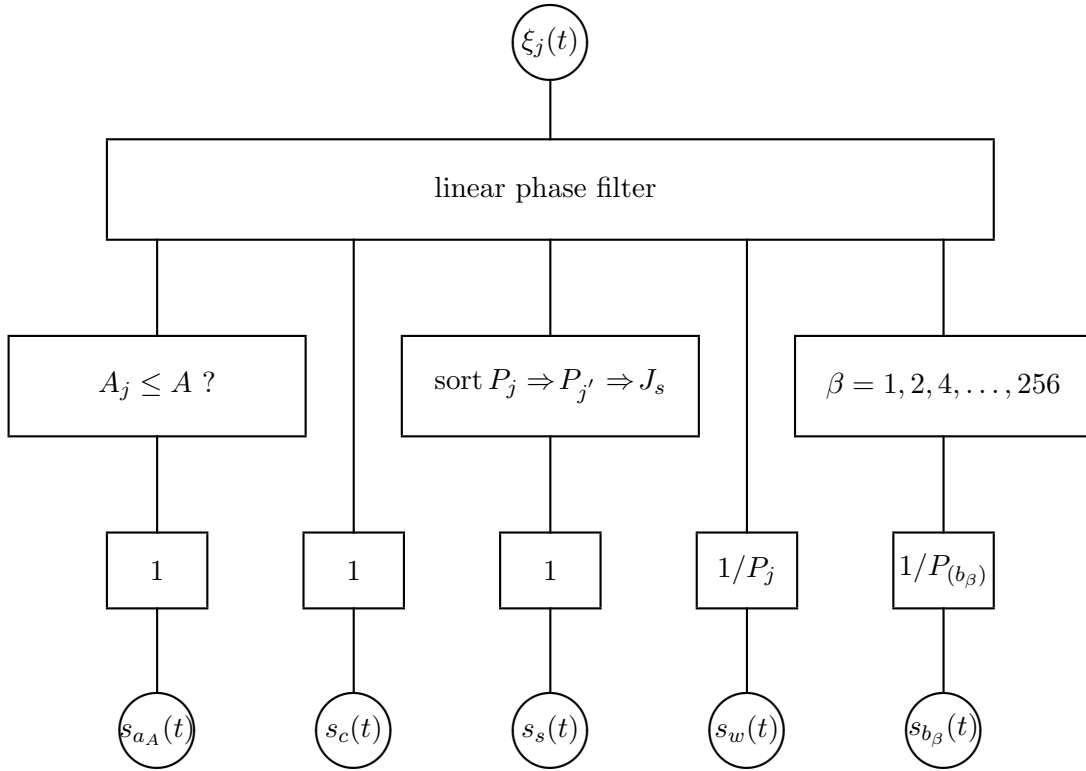
Fig. 2.1: Different averaging methods: The epochs $\xi_j(t)$ are passed through an FIR filter and than processed as indicated in the boxes in the second row (A_j: peak-to-peak voltage of sweep j, A artifact criterion, P_j powers of the epochs, $P_{j'}$ sorted powers, J_s number of sweeps entering the sorted average, β block size of the block weighted averaging scheme). The weighting of the processed sweeps is shown in the third row (P_j powers of the epochs, $P_{(b_\beta)}$ powers of blocks of epochs). The resulting averages are depicted in the circles at the bottom of the figure: $s_{a_A}(t)$ average using artifact rejection with threshold A, $s_c(t)$ conventional average, $s_s(t)$ sorted average, $s_w(t)$ weighted average, $s_{b_\beta}(t)$ block weighted average with block size β.

2.2.3   Evaluation of data quality

Two methods of data quality estimation are compared: one based on averages, the other on single sweeps.

It is common clinical practice to store only a (non-weighted) average. An estimate of the signal-to-noise ratio (SNR) is then obtained by repeating the measurement or by averaging alternately into two buffers yielding two sub-averages, each consisting of $J/2$ epochs. The

more similar they appear the better the quality of the data and the higher the SNR will be. By denoting the filtered epochs with odd number x_o and epochs with even number x_e the sub-averages as a function of time t are:

$$s_o(t) = \frac{1}{J/2} \sum_{o=1,3,\ldots}^{J-1} x_o(t), \qquad s_e(t) = \frac{1}{J/2} \sum_{e=2,4,\ldots}^{J} x_e(t) \,. \tag{2.13}$$

An attempt to estimate the residual noise σ_{oe} can be done in terms of the difference of the two sub-averages:

$$\sigma_{\mathrm{oe}}(t) = \frac{1}{2}(s_o(t) - s_e(t)) \,. \tag{2.14}$$

This concept was proposed as the 'plus-minus-reference' by Schimmel (1967) and later used by Wong and Bickford (1980) as the 'plus-minus-average'.

In the case of non-weighted averaging, the signal estimate $s(t)$ from all J epochs is identical with the estimate from the two sub-averages:

$$s(t) = \frac{1}{2}(s_o(t) + s_e(t)) = \frac{1}{J} \sum_{j=1}^{J} x_j(t) \,. \tag{2.15}$$

The advantage of storing single epochs becomes apparent in the estimation of the noise. Given all epochs the residual noise $\sigma(t)$ can be rigorously defined and calculated as the standard error of the average $s(t)$:

$$\sigma(t) = \sqrt{\frac{1}{J(J-1)} \sum_{j=1}^{J} (x_j(t) - s(t))^2} \,. \tag{2.16}$$

For data analysis the time-averaged quantities $s \equiv \mathrm{rms}(s(t))$, $\sigma_{\mathrm{oe}} \equiv \mathrm{rms}(\sigma_{\mathrm{oe}}(t))$, and $\sigma \equiv \mathrm{rms}(\sigma(t))$ as estimates of signal and noise rms are used, where $\mathrm{rms}(\cdot)$ denotes the root-mean-square-value, i.e., the square root of the average across time of all squared samples.

The expected values of the estimates of the noise variances σ^2 and σ_{oe}^2 are the same:

$$\mathrm{E}(\sigma_{\mathrm{oe}}^2) = \mathrm{E}(\sigma^2) = \frac{\sigma_0^2}{J} \,, \tag{2.17}$$

with σ_0^2 denoting the noise variance of the single epochs.

However, the variance of σ_{oe}^2 is greater by a factor $\frac{7}{4}J$ than the variance of σ^2 :

$$\text{var}(\sigma_{\text{oe}}^2) \quad = \quad \frac{7}{2}\frac{\sigma_0^4}{J^2}\,, \tag{2.18}$$

$$\text{var}(\sigma^2) \quad = \quad 2\frac{\sigma_0^4}{J^3}\,. \tag{2.19}$$

A derivation of equations (2.17), (2.18), and (2.19) is given in appendix A.

The single-sweep-based SNR estimate γ is given by

$$\gamma = \frac{s}{\sigma}\,, \tag{2.20}$$

while the average-based SNR estimate γ_{oe} is analogously defined by

$$\gamma_{\text{oe}} = \frac{s}{\sigma_{\text{oe}}}\,. \tag{2.21}$$

Note that γ and γ_{oe} are dimensionless.

Another single-sweep-based quality measure, the so-called F_{sp}-value, was introduced by Elberling and Don (1984). It is basically the square of γ as defined here with the denominator evaluated at a particular time sample t_{sp} only: $F_{sp} = \gamma^2(t_{sp}) = s^2/\sigma^2(t_{sp})$. Elberling and Don have shown that F_{sp} does not depend significantly on the choice of the single time point. Therefore it is frequently used to save computation time.

On the other hand, the average across time yields an estimate of the noise power with higher accuracy (variance divided by the number of samples per epoch) than a single-point variance. It is therefore used in this study and is also recommended for use in practical applications. Cebulla and collaborators, using Monte Carlo simulations, have recently shown an averaged "single point" variance to be superior to the classical F_{sp}, especially if the number of sweeps is low (Cebulla et al., 2000a,b).

The single-sweep-based estimate for the residual noise of the conventional average is defined by 2.16. For averaging using an artifact criterion and sorted averaging $\sigma_a(t)$ and $\sigma_s(t)$ are defined similarly. The standard error of the weighted average is defined as

$$\sigma_w(t) = \sqrt{\frac{\displaystyle\sum_{j=1}^{J}\frac{1}{P_j}(x_j(t) - s_w(t))^2}{(J-1)\displaystyle\sum_{j=1}^{J}\frac{1}{P_j}}}\,. \tag{2.22}$$

The residual noise estimate for the block-weighted averaging scheme is analogously defined by

$$\sigma_b(t) = \sqrt{\frac{\displaystyle\sum_{b=1}^{B} \frac{1}{P_{(b_\beta)}\,\beta} \sum_{j=(b-1)\beta+1}^{b\beta} \left(x_{(b_\beta)}(t) - s_b(t)\right)^2}{(J_b - 1)\displaystyle\sum_{b=1}^{B} \frac{1}{P_{(b_\beta)}}}} \, . \tag{2.23}$$

To compare both estimation methods experimentally, the average-based noise estimate $\sigma_{\mathrm{oe}}(t)$ was also computed for all averaging methods. For sorted and block-weighted averaging it is not possible to apply a recording technique using alternating buffers. Therefore each raw epoch-file was split into two files, one containing $J/2$ sweeps with odd, the other $J/2$ sweeps with even epoch numbers. Afterwards all averaging methods were applied to both files, but only the averages were used in the subsequent analysis.

2.2.4   Iterative averaging

The iterative averaging technique was developed to avoid problems with weighted averaging described, e.g., by Lütkenhöner et al. (1985). The correct weighting of an epoch for the weighted average is the inverse power of its noise. However, with the approximation of eq. (2.4) the weighting is determined as the inverse power of the measured epoch consisting of signal and noise. This leads to an undesired underestimation of the overall magnitude of the signal.

With the approximation of eq. (2.4) an estimate of the noise in a single epoch is defined as

$$n_j^{(0)}(t) = x_j(t) \, . \tag{2.24}$$

As indicated by the superscript in parentheses, this quantity is called the noise estimate of order zero.

Since an estimate of the noise in the single sweep is the basis for all averaging schemes except for the conventional average, an improvement of this estimate will affect the sorted, weighted, and block-weighted average as well as the average involving an artifact criterion. Only the equations for the case of weighted averaging are presented here, because it is straightforward to apply the following considerations to the other methods.

The residual noise of the weighted average in order zero[4] is given by

$$\sigma_w^{(0)}(t) = \sqrt{\frac{\displaystyle\sum_{j=1}^{J} \frac{1}{P_j}(x_j(t))^2}{(J-1)\displaystyle\sum_{j=1}^{J}\frac{1}{P_j}}} \ . \tag{2.25}$$

The signal estimate corresponding to eq. (2.24) is

$$s^{(0)}(t) \equiv 0 \ . \tag{2.26}$$

In the computation of the weighted average, this approximation was used to determine the weightings (cf. eq. (2.8)). The result of this computation, however, is a better signal estimate $s_w(t) = s_w^{(1)}(t)$ of first order (cf. eq. (2.9)), which in turn can be used to improve the noise estimate of the single sweeps:

$$n_{j,w}^{(1)}(t) = x_j(t) - s_w^{(1)}(t) \tag{2.27}$$

If the inverse power of the noise estimate $n_{j,w}^{(1)}(t)$ is used as new weight

$$w_j^{(1)} = \frac{1}{P(n_{j,w}^{(1)}(t))} = \frac{1}{\displaystyle\sum_{t=1}^{T}(x_j(t) - s_w^{(1)}(t))^2} \ , \tag{2.28}$$

an improved signal estimate, the weighted average of second order, can now be calculated:

$$s_w^{(2)}(t) = \frac{\displaystyle\sum_{j=1}^{J} w_j^{(1)} x_j(t)}{\displaystyle\sum_{j=1}^{J} w_j^{(1)}} \ . \tag{2.29}$$

The residual noise of the improved weighted average is analogously defined as

$$\sigma_w^{(2)}(t) = \sqrt{\frac{\displaystyle\sum_{j=1}^{J} w_j^{(1)}(x_j(t) - s_w^{(2)}(t))^2}{(J-1)\displaystyle\sum_{j=1}^{J} w_j^{(1)}}} \ . \tag{2.30}$$

---

[4]This quantity is a noise estimate of the measurement rather than a standard error.

Of course it is possible to repeat the process by stating that if $s_w^{(2)}(t)$ is a better signal estimate than $s_w^{(1)}(t)$, then $n_{j,w}^{(2)}(t) = x_j(t) - s_w^{(2)}(t)$ should also be a better noise estimate than $n_{j,w}^{(1)}(t)$. This method is therefore called iterative weighted averaging.

One can easily generalize the process of iteration to the case of block-weighted averaging. On the other hand it is not so obvious that iteration can also be applied in the cases of averaging with artifact criterion and sorted averaging. However, these methods can also be considered as weighted averaging schemes with only weightings zero and one allowed. Improving the noise estimate of the single epochs will affect the result. Providing the better noise estimate $n_{j,a_A}^{(1)}(t) = x_j(t) - s_{j,a_A}^{(1)}(t)$, one has to reinvestigate the artifact criterion A for $n_{j,a_A}^{(1)}(t)$ instead of the $x_j(t)$. The same holds of course for the case of sorted averaging. Only conventional averaging is not influenced by iteration since no rejection or weighting takes place.

2.2.5 Simulations

The properties of the respective averaging methods can only be assessed on the basis of *estimates* of the "true" signal and noise components rather than directly being based on these components themselves. Hence, the apparent advantage of one of the methods tested above in comparison to another might perhaps be due to an overly optimistic signal estimate or a too low residual noise estimate. To overcome this problem simulations using *a priori* known signal and noise components were performed as follows: From each subject sweeps were recorded without presenting a stimulus. To each of these sweeps a known signal $S(t)$ was added. (The mean over subjects of one of the previously calculated averages was chosen.) All averaging methods described above were applied to these derived test signals. From the results the following questions can be addressed:

1. Which averaging method is superior to the others in terms of the best reconstruction $s(t)$ of the true signal $S(t)$?

2. Which averaging method yields the best estimation of the signal, the residual noise and the SNR of the average it produces, i.e., is most reliable in assessing its own performance ?

The true residual noise Σ is defined as the rms-value of the difference between $s(t)$ and $S(t)$. Note that $\mathrm{rms}(s(t)) = \mathrm{rms}(S(t))$ does not imply $\Sigma = 0$ because true and estimated signal may have the same rms-value without being identical. Therefore the quality of the

averaging methods must be evaluated by computing Σ. However, the comparison of s and S, σ and Σ as well as γ and Γ, i.e., of estimated and true quantities, provides the answers to the second question.

2.3   Results

2.3.1   Averaging methods

Fig. 2.2 gives an example of the different approaches underlying the single-sweep-based and average-based signal and noise estimation. Data from binaural stimulation at $60\,\mathrm{dB\,nHL}$ from one subject are shown. The time interval from 0 to $10\,\mathrm{ms}$ after stimulus onset was chosen to determine all quantities described in the following.

In the upper left graph of Fig. 2.2 the conventional average (10000 sweeps) $s_c(t)$ is depicted as a solid line, while the dotted lines refer to $s_c(t) \pm \sigma_c(t)$, i.e., signal estimate $\pm$ residual noise estimate. The upper right graph shows the same data averaged into two buffers in an alternating way. Both averages result from 5000 sweeps, the first from the sweeps with odd sweep number, the second from the sweeps with even sweep number.

In the lower left graph the time dependent standard error $\sigma_c(t)$ and its rms-value σ_c are depicted. In the lower right graph the average-based estimate of the residual noise $\sigma_{\mathrm{oe}}(t)$ equal to half the difference between the subaverages (see Sect. 2.2.3) and its rms-value σ_{oe} are shown.

While $\sigma_c(t)$ only shows small variation over time, $\sigma_{\mathrm{oe}}(t)$ exhibits large fluctuations. Note that the voltage scale for $\sigma_c(t)$ is 20 times smaller than for $\sigma_{\mathrm{oe}}(t)$. Since $\sigma_{\mathrm{oe}}(t)$ vanishes where the sub-averages intersect, it does not provide a realistic time course of the noise. Therefore only the rms-value σ_{oe} can be considered as a meaningful noise estimate. Thus, the higher accuracy of σ_c in comparison to σ_{oe} as theoretically derived in appendix A is experimentally supported here.

Fig. 2.3 shows the estimates of signal rms s, noise rms σ, and their quotient γ which serves as SNR estimate depending on the various averaging methods for channel M2 (right mastoid versus vertex). Data for individual subjects are plotted with open circles and connected with broken lines. Mean values across subjects are represented by solid circles and connected with solid lines. The indices at the abscissa refer to the averaging methods as explained in section 2.2.2 and are used as labels in the subsequent figures, too. The variance of s, σ and γ across subjects is considerably higher than across averaging schemes.

From the left graph in Fig. 2.3 it can be seen that the signal estimates for weighted averaging s_w (label 'w') and sorted averaging s_s (label 's') are lowest whereas the value of s_b for the block weighted averaging scheme with block size 256 is similar to s_a (i.e., employing different artifact criteria) and s_c (i.e., conventional averaging). The middle graph shows that estimates of noise rms values σ increase with increasing artifact threshold. The only exception – for one subject the artifact criterion $\pm 6\,\mu\text{V}$ results in the highest σ-value – is due to the small number of accepted epochs in that particular average. Except for this case, conventional averaging leads to the highest levels of residual noise. Hence, switching off the artifact rejection or choosing a too strict criterion for it raises the noise level. σ_s and σ_w are the lowest noise values reflecting the known underestimation of signal and noise in weighted averaging and suggesting that there is a similar, but weaker effect for sorted averaging. The noise level of block weighted averaging σ_b is in the same range as for averaging with artifact criterion.

The right graph of Fig. 2.3 indicates that the estimated SNR values γ are lowest for the case of conventional averaging, but no method can unambiguously be identified as superior to the others. For some subjects the non-weighted averaging schemes give higher γ-values, for other subjects the weighted averaging schemes are advantageous.

Fig. 2.4 demonstrates the effect of different block sizes of the block weighted averaging method on s, σ and γ. For each subject, the same 9984 (39×256) epochs entered the nine different averages. Block sizes were chosen in steps of powers of two ranging from 1 to 256. Averages corresponding to block sizes $\beta = 1$ and $\beta = 256$ already appeared in Fig. 2.3 with the labels 'w' and 'b'. Fig. 2.4 therefore provides a fine resolution between weighted averaging and block weighted averaging using the largest block size available. Both s and σ increase monotonically with increasing block size. This reflects the monotonic decrease of the underestimation of signal and residual noise. Because σ increases faster with β than s, a maximum of γ is found at block size eight on the average across subjects. Note that the resulting value ($\gamma_{b_8} = 13.4$) is larger than that obtained for all averaging methods considered in the previous figure.

In Fig. 2.5 the estimates of residual noise based on single epochs and based on two sub-averages are compared for the various averaging methods. Data is normalized to the individual estimate obtained for each individual subject with averaging using an artifact criterion of $\pm 10\,\mu\text{V}$. This eliminates the high variance between subjects. The average-based noise estimate depicted in the right graph of Fig. 2.5 shows noticeable higher variation than the single-sweep-based noise estimate.

Monaural stimulation results in s-values of about half the magnitude compared to binaural stimulation. Since the residual noise estimates are nearly unaffected by the mode of stimulation, the γ-values also halve in the case of monaural stimulation. Lowering the presentation level did not affect the σ-values either, but lowered the s- and γ-values. For the stimulus levels of 40 and 20 dB normal hearing level, mean SNR estimates of about 10 and 6, respectively, are obtained. Inspection of channel M1 (left mastoid versus vertex) and channel FZ (forehead versus vertex) shows similar dependence on averaging schemes. For channel M1, the s-, σ-, and γ-values are nearly identical to the values of channel M2. In the case of channel FZ, the s-values range about one third of the values of channels M1 and M2. Residual noise level is at about 75 % of the mastoidal channels resulting in γ-values of about 45 % of the values of channels M1 and M2.

2.3.2   Iterative averaging

To demonstrate the effect of iteration, an example of weighted averaging is shown for the case of diotic stimulation at 60 dB nHL for one subject in Fig. 2.6. In the upper graph the non-iterated signal estimate $s_w^{(1)}(t)$ is compared to the iterated estimates $s_w^{(2)}(t)$ and $s_w^{(3)}(t)$ as well as the trivial signal estimate $s_w^{(0)}(t) = 0$. Apparently, $s_w^{(1)}(t)$ has a smaller amplitude than the other two curves demonstrating the underestimation of the signal by weighted averaging, which is eliminated by the iteration process. The fact that $s_w^{(2)}(t)$ and $s_w^{(3)}(t)$ are nearly identical shows that the iteration process leads to a significant change in the waveform only in the first step, resulting in $s_w^{(2)}(t)$. Similar results with the data from the other subjects show that the iteration procedure converges quickly and is stable.

In the lower graph the corresponding residual noise estimates $\sigma_w^{(0)}(t)$, $\sigma_w^{(1)}(t)$, $\sigma_w^{(2)}(t)$, and $\sigma_w^{(3)}(t)$ are depicted. The dashed line representing $\sigma_w^{(0)}(t)$ clearly contains signal information because the measured signal is considered as noise alone (cf. eq. (2.4)). Apparently, a more accurate weighted average is achieved after at least one iteration step. Further iteration steps do not significantly lower the estimated residual noise.

In Fig. 2.7 the effect of iteration on the estimates is shown for the various averaging methods. Again data is normalized to the individual estimate obtained for each individual subject with averaging using an artifact criterion of $\pm 10 \, \mu$V. For each averaging method a pair of values is shown, on the left, the non-iterated quantities (order 1) by means of a triangle pointing to the right, on the right, the iterated quantities (order 2) by means of a triangle pointing to the left. From the left graph of Fig. 2.7 it can be seen that the $s^{(2)}$-values are approximately independent of the averaging method. A pronounced increase in

signal power due to a single iteration is observed for $s_w^{(2)}$, $s_s^{(2)}$, and $s_6^{(2)}$, respectively. This underlines the notion that the underestimation of the signal in the case of weighted and sorted averaging can be overcome by one iteration step.

The middle graph shows that residual noise estimates $\sigma^{(1)}$ and $\sigma^{(2)}$ do not differ significantly. This was already observed in the lower graph of Fig. 2.6 for one subject. The residual noise estimates $\sigma^{(0)}$ – not shown here – are slightly larger than $\sigma^{(1)}$ and $\sigma^{(2)}$ for all averaging methods. This reflects the exclusion of signal components from the noise estimates if iteration is used.

The SNR-estimates depicted in the right graph indicate that weighted averaging with iteration is superior to all the other methods. Sorted averaging is also improved by iteration. Except for the strictest artifact criterion there is no strong effect of iteration on the other averaging methods.

The effect of iteration for the case of block weighted averaging is shown in Fig. 2.8. As before, data is normalized to the individual data from averaging using an artifact threshold of $\pm 10\,\mu$V. Iteration raises s-values significantly for small block sizes, but yields smaller effects with increasing block size. As for all other averaging schemes, the σ-values do not vary due to iteration. The "optimal" block size with the highest SNR-estimate is lowered from eight in the non-iterated case to one when iteration is applied. γ-values in iteration decrease monotonically with increasing block size. Hence, the optimal averaging scheme in the sense of the highest γ is weighted averaging with one iteration.[5]

2.3.3   Simulations

In Fig. 2.9 the true residual noise rms-value Σ, definded as the rms-value of the difference of true signal S and estimate s, is plotted for each subject (open symbols) for the various averaging methods in the non-iterated as well as the iterated case. Again, mean values (solid symbols) and standard deviation across subjects are given. The resulting residual noise Σ shows larger values for non-iterated than for iterated averaging in the case of weighted averaging, to a smaller extent also for sorted averaging and averaging with artifact threshold $\pm 6\,\mu$V.

Except for conventional averaging, the mean true residual noise after one iteration $\Sigma^{(2)}$

---

[5]Again the analysis of the other channels recorded leads to analogous dependencies of the estimates on averaging methods and iteration. For channel M1 (left mastoid versus vertex) the magnitudes of the data are nearly identical, while for channel Fz (forehead versus vertex) the s- and γ-values are lower than in the other channels.

ranges from 15 to 19 nV. From the averaging methods shown in Fig. 2.9 the lowest value is obtained for block weighted averaging with the largest block size ($\Sigma_{b_{256}}^{(2)} = 15.0$ nV), followed by averaging with artifact threshold $\pm 10\,\mu$V ($\Sigma_{a_{10}}^{(2)} = 15.9$ nV) and weighted averaging ($\Sigma_w^{(2)} = 16.4$ nV).

The rms values of the true residual noise Σ are plotted for the block weighted averaging method in Fig. 2.10 as a function of the block size. Without iteration, the residual noise values $\Sigma^{(1)}$ for small block sizes are comparatively large. For block sizes greater than eight $\Sigma^{(1)}$ stays almost constant. $\Sigma^{(2)}$-values (iterated case) are nearly constant for all block sizes, and there is a shallow minimum for block size four with $\Sigma_{b_4}^{(2)} = 13.9$ nV. For block sizes greater than 16 the differences between $\Sigma^{(1)}$ and $\Sigma^{(2)}$ are negligible.

Fig. 2.11 gives the quotients of the estimated and true quantities for the various averaging methods with and without iteration. The left graph shows that without iteration weighted averaging underestimates the signal by 15% while a smaller effect occurs for sorted averaging (-7%) and averaging with artifact threshold $\pm 6\,\mu$V (-4%). Increasing the artifact threshold to $\pm 10\,\mu$V or $\pm 14\,\mu$V yields a slight overestimation of S.

Iteration considerably improves sorted and weighted averaging: signal estimates deviate by less than 1% above and 2% below the true value S. In conclusion, with iteration signal estimation works well for all methods. Without iteration, the classical technique using a not too strict artifact criterion provides good signal-rms estimates.

With regard to the residual noise estimation depicted in the middle graph of Fig. 2.11, rather large deviations between the non-iterated and iterated case are observed. Without iteration, residual noise is heavily underestimated by weighted averaging (38% of the true value) and sorted averaging (69%). On the other hand, averaging using artifact criterion and block weighted averaging overestimate Σ by up to 17%. Surprisingly, the best non-iterated residual-noise-estimation is provided by conventional averaging and averaging with the strictest artifact criterion.

Iteration does not correct the overestimation for those methods that overestimate the residual noise before iteration. However, the huge underestimation for sorted and weighted averaging is reduced to 5% and 8% respectively. This indicates that sorted and weighted averaging in iteration are capable of providing accurate estimates of residual noise.

In the right graph of Fig. 2.11, mean and standard deviation of the normalized SNR estimates across subjects are given. Without iteration, the SNR is overestimated by sorted averaging by 45% and by more than a factor of two by weighted averaging. Averaging using artifact rejection yields estimates close to the true Γ, but both signal and residual noise

are overestimated. The overestimation of Γ by sorted and weighted averaging is strongly reduced by iteration and amounts to 9% and 13%, respectively. The other averaging methods are nearly unaffected by iteration.

In Fig. 2.12 the same type of plot is presented as in the previous figure but for different block sizes of the block weighted averaging method. Generally, the effect of iteration is negligible for block sizes greater than eight. A trend of overestimating the signal with increasing block size can be observed. Non-iterated signal estimators underestimate S for block sizes smaller than 32. For greater block sizes and for all iterated estimators, signal rms S is well estimated within a range of ± 2 %.

Again, the effect of iteration is larger for the residual noise estimates σ than for the signal estimates. Iteration increases σ-values for block sizes up to $\beta = 8$. For greater block sizes, Σ is overestimated by about 15% for both the non-iterated and iterated case. Without iteration, Σ is underestimated for $\beta = 1$ by 62%, for $\beta = 2$ by 35% , and best estimated for $\beta = 4$ and $\beta = 8$. With iteration, $\beta = 1$ (weighted averaging) fits the true Σ best (-9%), while bigger block sizes overestimate Σ by at least 10%.

The quality of the SNR estimates is again shown in the right graph. In the non-iterated case a strong overestimation by more than a factor two and 52% occurs for block size one (weighted averaging) and two, respectively. In the iterated case, Γ is best estimated by weighted averaging ($\beta = 1$) and $\beta = 2$. A considerable underestimation of Γ independent of iteration of about 10% is observed for block sizes larger than four.

With the simulation technique employed here, the true residual noise Σ can be computed as a function of the number of sweeps included in the average. This was done by calculating the rms-value of the difference between the known signal $S(t)$ and the average after the inclusion of $j, j = 1 \ldots J$ sweeps. In Fig. 2.13 the average $\Sigma_j^{(2)}$ across subjects is shown for four averaging methods using iteration: conventional averaging, averaging using the $\pm 10 \, \mu$V artifact criterion, weighted and sorted averaging. For a given number of sweeps, $\Sigma_j^{(2)}$ decreases in the order of the above list, i.e., conventional averaging yields the highest residual noise values for a given number of sweeps. The classical method using the $\pm 10 \, \mu$V artifact threshold considerably lowers residual noise. A further improvement is possible with weighted averaging. Sorted averaging has the lowest residual noise values over a wide range of sweep numbers. However, it has to be taken into account that for sorted averaging $\Sigma_{j'}$ is shown, i.e., the residual noise as a function of the epoch index after sorting. Hence, the sweeps with the lowest noise contamination enter the average first. This results in a lower residual noise at the beginning than for the other methods. However, for sweep

numbers greater than about 9400, Σ_s remains constant. About 600 sweeps are rejected by this averaging scheme, because their inclusion would raise $\Sigma_{j'}$ again. Weighted averaging on the other hand can further decrease Σ by using all sweeps, i.e., also the highly noise contaminated ones, to which small weightings are assigned.

From Fig. 2.13 it can also be seen how many epochs have to be included into an average to bring the residual noise below a given criterion. Tab. 2.1 lists those values for residual noise criteria between 20 and 50 nV.

Although the number of sweeps required is lowest in the case of sorted averaging, this does not allow a reduction of the number of epochs to be recorded before sorting because all sweeps have to be collected. To reach the 25-nV-criterion, for example, artifact rejection reduces the number of epochs to 85%, weighted averaging to 76%, compared to conventional averaging.[6]

Σ [nV]	$J_c^{(2)}$	$J_{a_{10}}^{(2)}$	$J_w^{(2)}$	$J_s^{'(2)}$
50	1775	1586	1229	1069
45	2101	1864	1746	1239
40	3621	2132	2092	1679
35	5290	3196	2498	2084
30	6248	3817	3420	2466
25	9513	6052	4683	3114
20	> 10000	8039	7215	6449

Tab. 2.1: *Sweep numbers necessary to reach a given residual noise criterion for four averaging methods using the iteration procedure proposed here, i.e., conventional averaging ($J_c^{(2)}$), averaging using an artifact threshold of $\pm 10\,\mu V$ ($J_{a_{10}}^{(2)}$), weighted averaging ($J_w^{(2)}$) and sorted averaging ($J_s^{'(2)}$). Mean over subjects. Channel M2 (right mastoid versus vertex).*

---

[6]The results of the simulations look very similar if channel M1 (left mastoid versus vertex) is considered. For the third channel with poor signal, of course, true and estimated signal and SNR are lower, but the dependence of the averaging methods and iteration is the same.
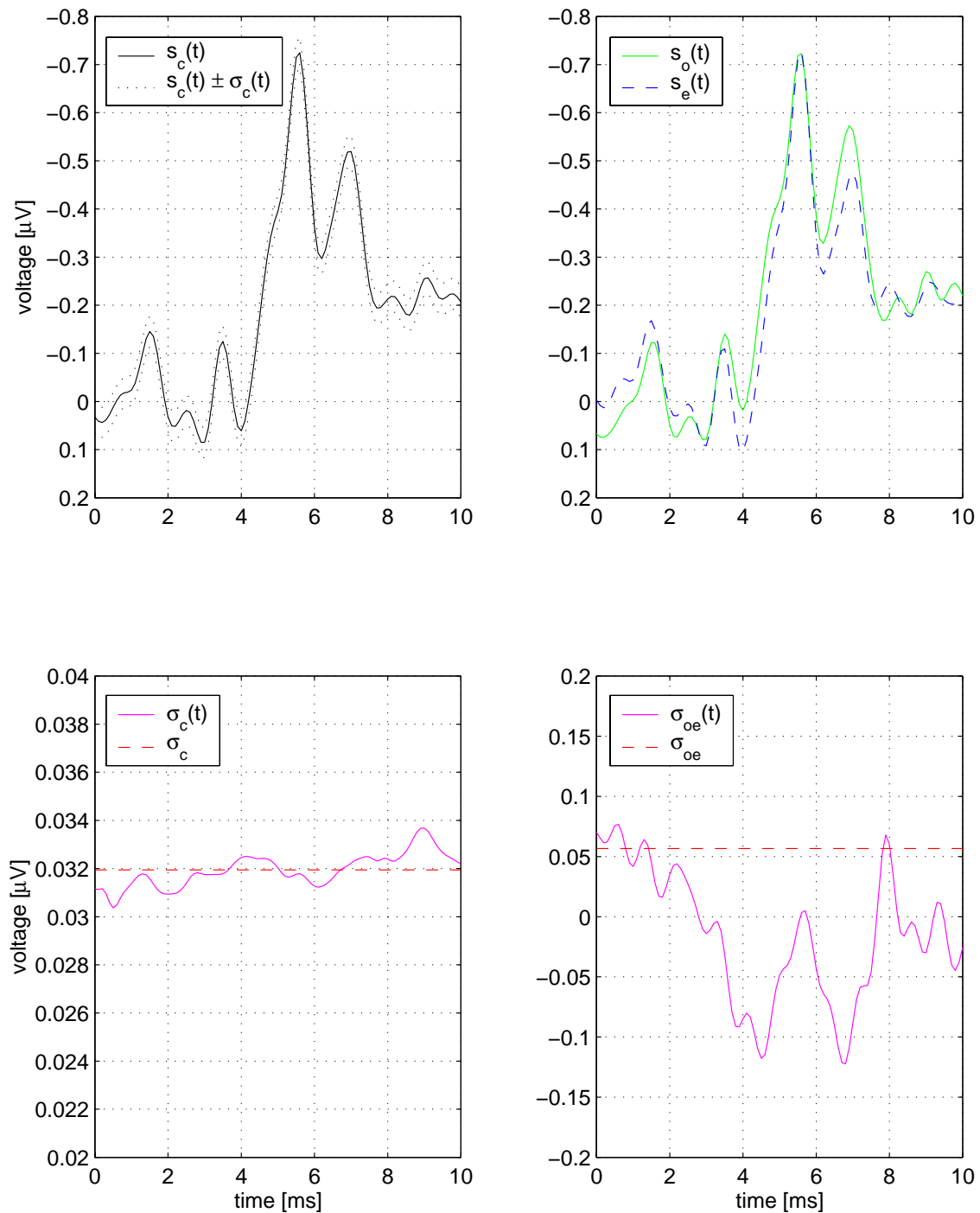
Fig. 2.2: *Comparison between single-sweep-based and average-based signal and noise estimates, example for binaural stimulation at 60 dB nHL for subject vk.* **Top row:** *Signal estimates.* **Upper left graph:** *Conventional average of 10000 epochs with one standard error.* **Upper right graph:** *Two sub-averages of 5000 epochs each.* **Bottom row:** *Corresponding noise estimates.* **Lower left graph:** *The single-sweep-based time dependent estimate of the standard error $\sigma_c(t)$ and its rms-value σ_c.* **Lower right graph:** *The average-based time dependent noise estimate $\sigma_{oe}(t)$ and its rms-value σ_{oe}.*

Fig. 2.3: Signal rms estimate s (left), residual noise rms estimate σ (middle), and SNR estimate γ (right) for various averaging methods. s: sorted averaging, 6,10,14: different artifact thresholds, c: conventional averaging without artifact criterion, w: weighted average, b: block-weighted average with block size 256. Data from individual subjects are connected by dotted lines, solid lines with filled symbols represent the mean across subjects. Data from diotic stimulation at 60 dB normal hearing level for channel M2 (right mastoid versus vertex).
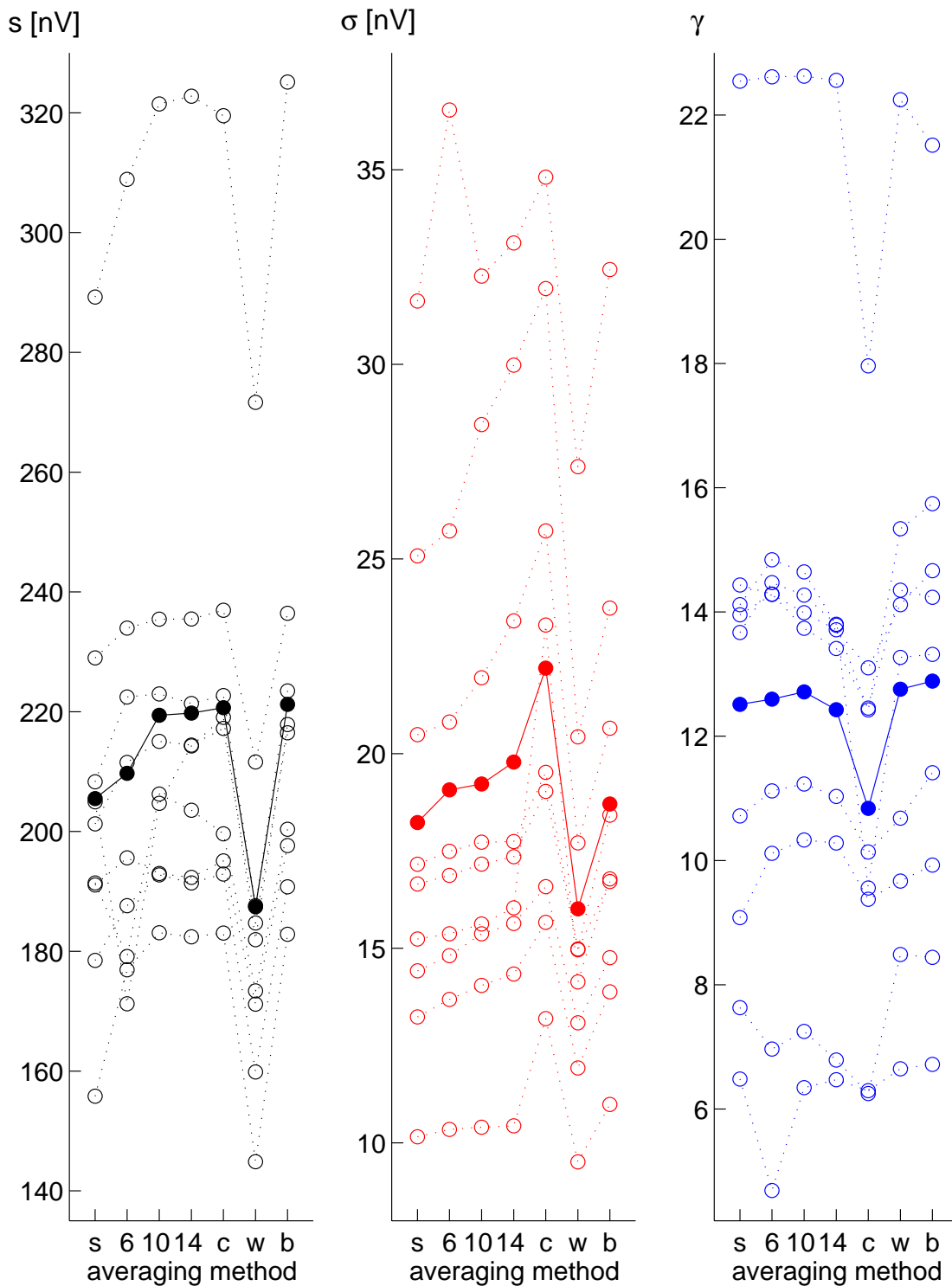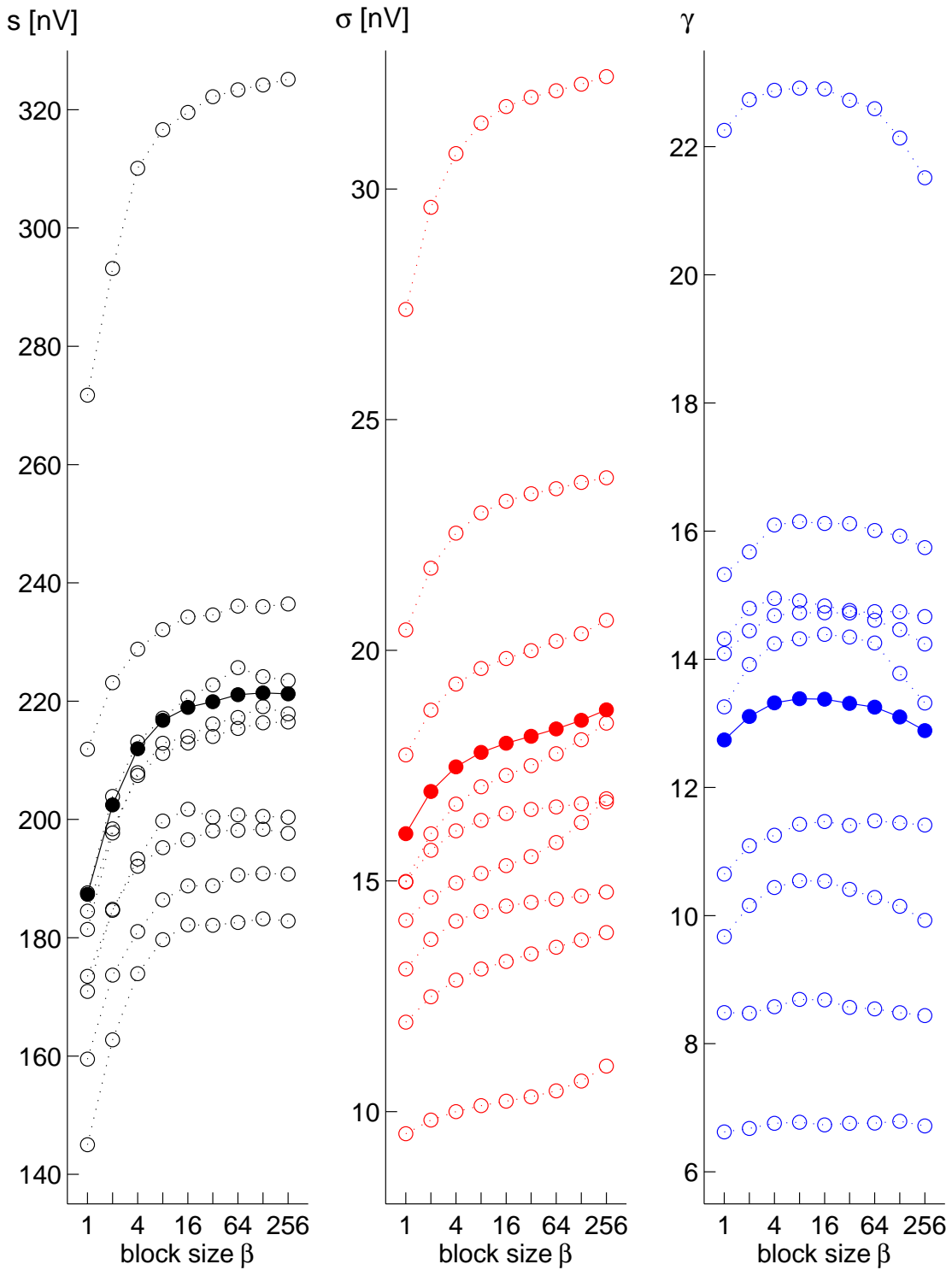
Fig. 2.4: Signal rms estimate s (left), residual noise rms estimate σ (right), and SNR estimate γ (right) as a function of the block size β for the block weighted averaging method. Data from individual subjects are connected by dotted lines, solid lines with filled symbols represent the mean across subjects. Data from diotic stimulation at 60 dB normal hearing level for channel M2 (right mastoid versus vertex).
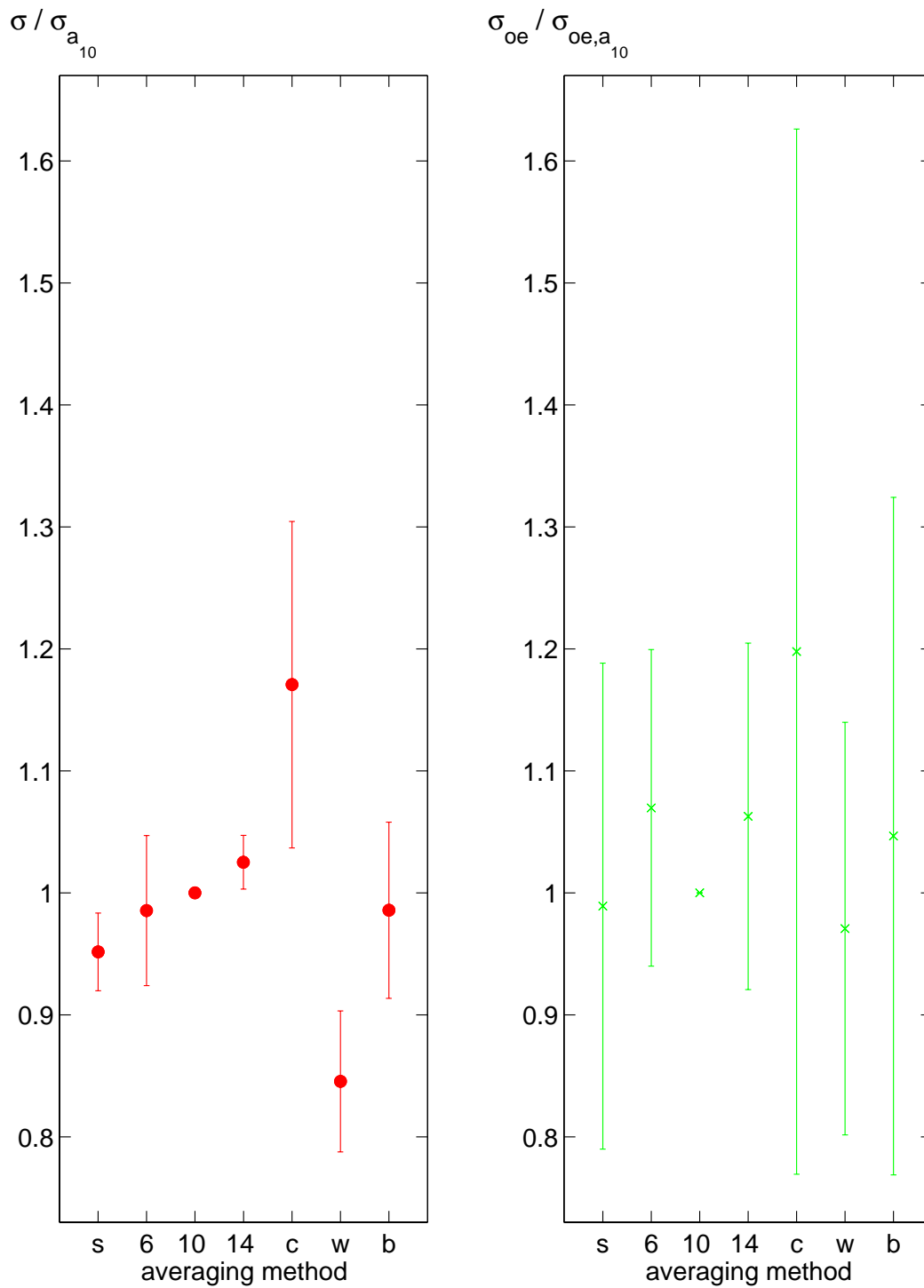
Fig. 2.5: *Comparison between single-sweep-based and average-based noise estimation for the various averaging methods. To reduce the high interindividual variance data are normalized to the averaging method using an artifact criterion of $\pm 10\,\mu V$.* **Left graph:** *noise estimate σ based on single sweeps.* **Right graph:** *noise estimate σ_{oe} based on two averages. The standard deviation (error bars) of σ_{oe} is significantly higher than that of σ for all averaging methods.*
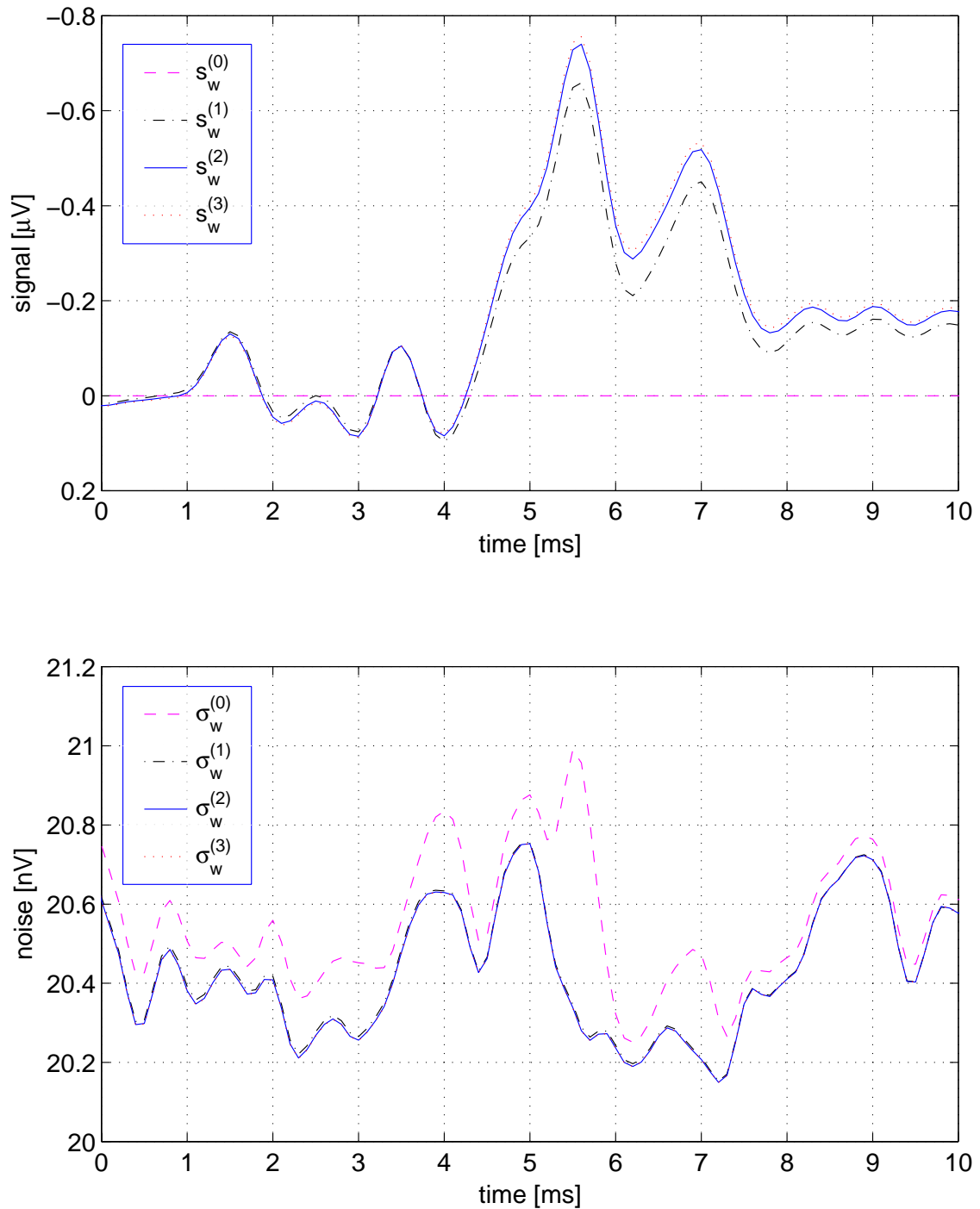
Fig. 2.6: Iterative averaging: **Upper graph:** Signal estimates $s_w^{(0)}(t) \equiv 0$, $s_w^{(1)}(t)$, $s_w^{(2)}(t)$, and $s_w^{(3)}(t)$. $s_w^{(1)}$ has a smaller magnitude than $s_w^{(2)}$. The difference between $s_w^{(2)}$ and $s_w^{(3)}$ is negligible. **Lower graph:** Noise estimates $\sigma_w^{(0)}(t)$, $\sigma_w^{(1)}(t)$, $\sigma_w^{(2)}(t)$, and $\sigma_w^{(3)}(t)$. $\sigma_w^{(0)}$ contains signal components which disappear in the iterated estimates. The differences between $\sigma_w^{(1)}$, $\sigma_w^{(2)}$, and $\sigma_w^{(3)}$ are very small indicating the quick convergence of the iteration process. Data are taken from diotic stimulation at $60\,\mathrm{dB\,nHL}$ for one subject.
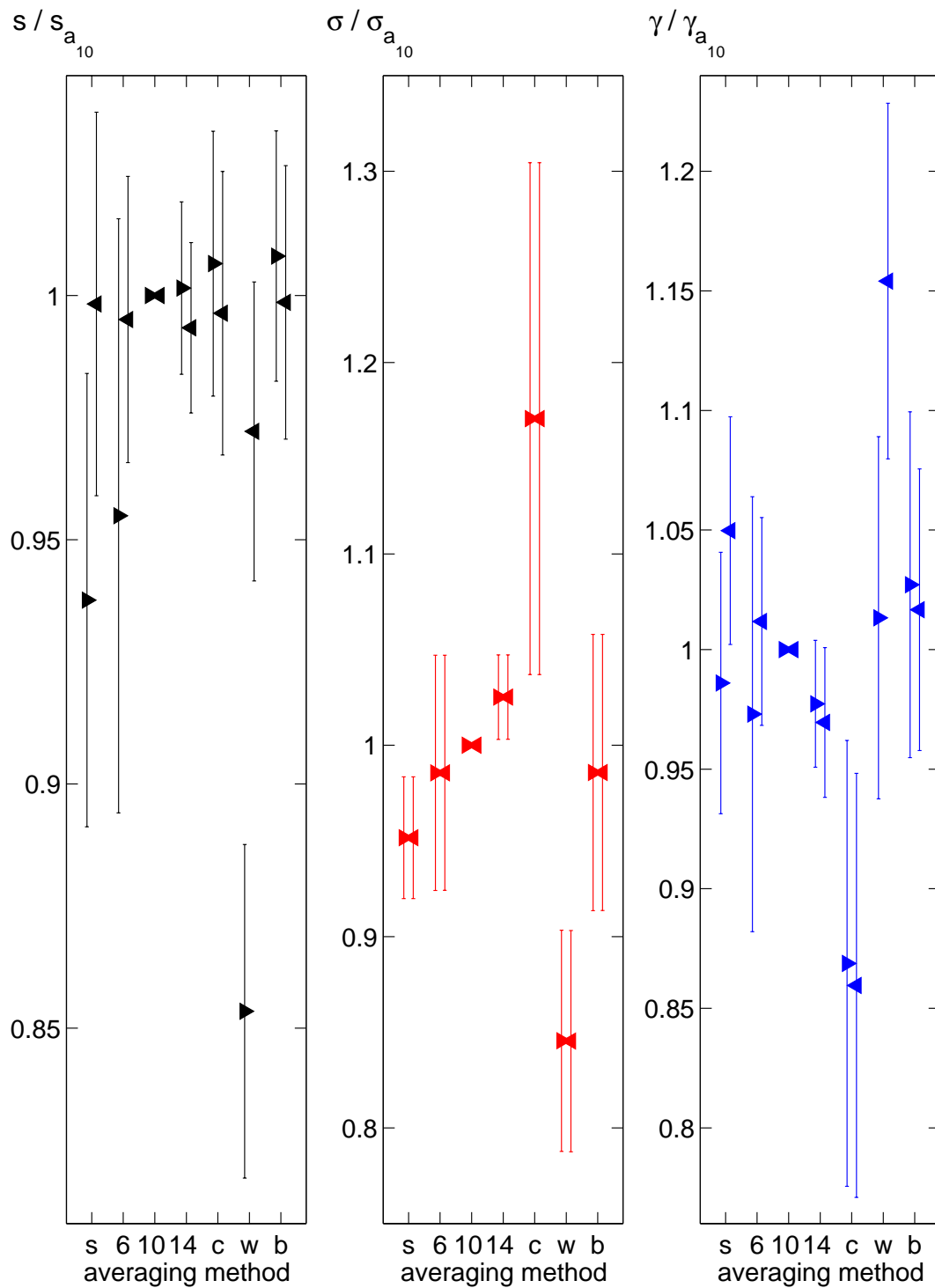
Fig. 2.7: Mean and interindividual standard deviation of the relative signal rms estimate $s/s_{a_{10}}$ (left), residual noise rms estimate $\sigma/\sigma_{a_{10}}$ (middle), and SNR estimate $\gamma/\gamma_{a_{10}}$ (right) for various averaging methods (labelled as in Fig. 2.3). Data are individually normalized to the values for averaging using an artifact threshold of $\pm 10\,\mu V$. For each method a pair of data is depicted: the symbols on the left side represent the non-iterated estimates, the symbols on the right side the estimates after one iteration step. Data from diotic stimulation at $60\,dB$ normal hearing level for channel M2 (right mastoid versus vertex).
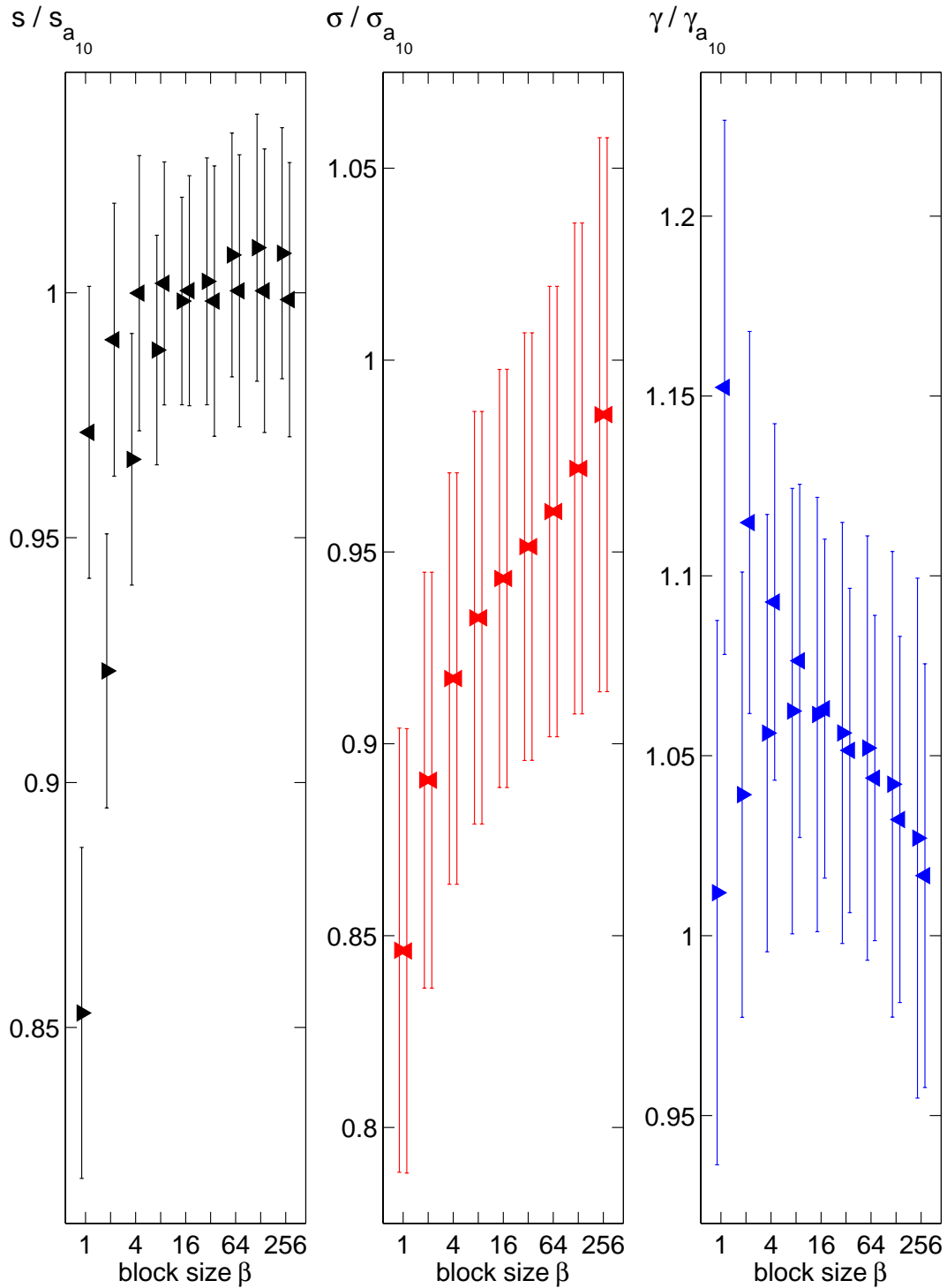
Fig. 2.8: Mean and interindividual standard deviation of the relative signal rms estimate $s/s_{a_{10}}$ (left), residual noise rms estimate $\sigma/\sigma_{a_{10}}$ (middle), and SNR estimate $\gamma/\gamma_{a_{10}}$ (right) as a function of the block size β for the block weighted averaging method. Data are divided by the values for averaging using an artifact threshold of $\pm 10\,\mu V$ to eliminate the variance across subjects. For each method a pair of data is depicted: the symbols on the left side represent the non-iterated estimates, the symbols on the right side the iterated estimates. Data from diotic stimulation at $60\,dB$ normal hearing level for channel M2 (right mastoid versus vertex).
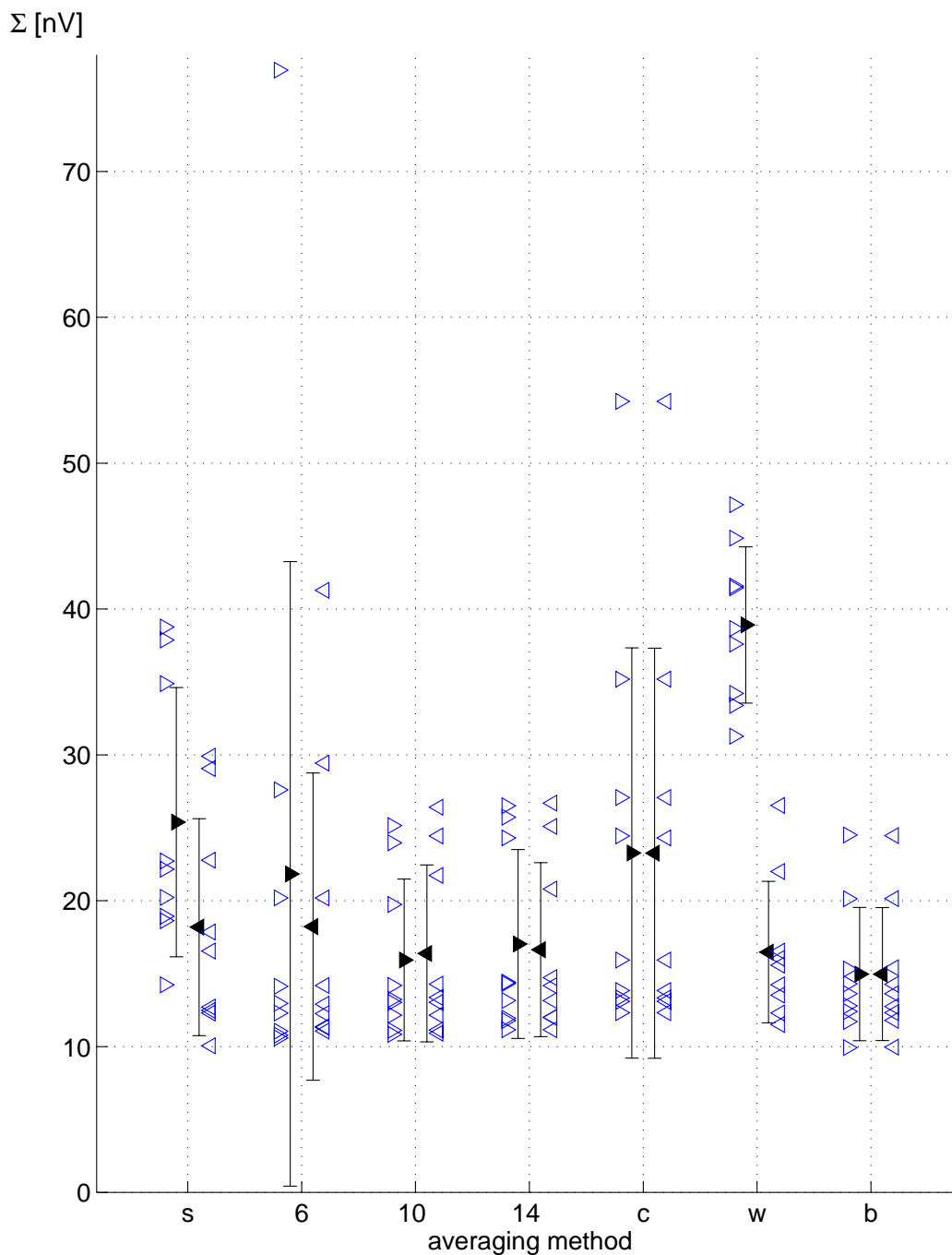
Fig. 2.9: Simulation results. True residual noise Σ for various averaging methods (labelles as in Fig. 2.3) and iterations. Mean values (over subjects) are shown with filled symbols and error bars indicating $\pm$ one standard deviation, the open symbols represent the data for individual subjects. Non-iterated data are plotted at the left side of the vertical grid lines, iterated data at the right side. Data from diotic stimulation at 60 dB normal hearing level for channel M2 (right mastoid versus vertex).
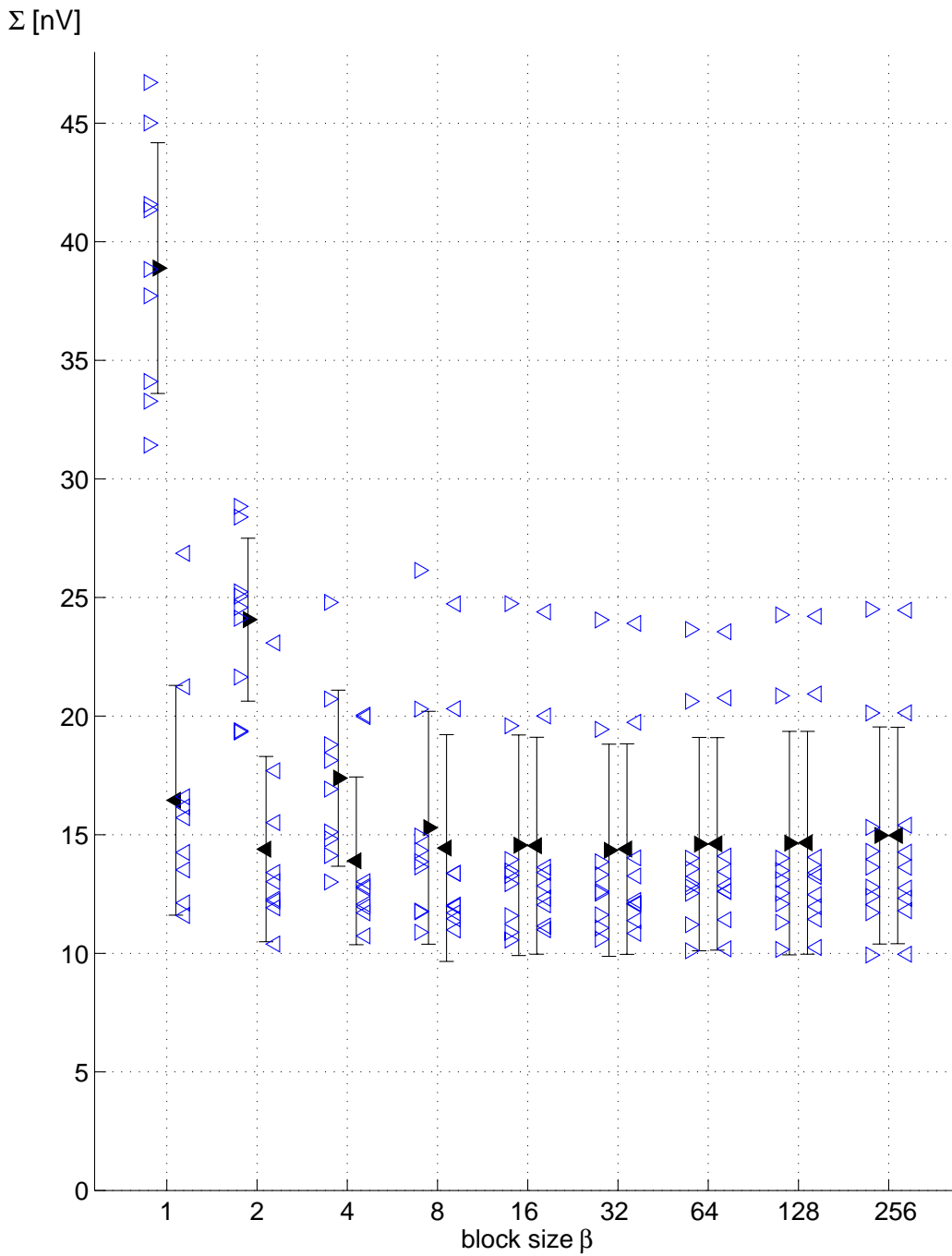
Fig. 2.10: Simulation results. True residual noise Σ as a function of the block size β for the block weighted averaging method. Mean values (over subjects) are shown with filled symbols and error bars indicating ± one standard deviation, the open symbols represent the data for individual subjects. Non-iterated data are plotted at the left side of the vertical grid lines, iterated data at the right side. Data from diotic stimulation at 60 dB normal hearing level for channel M2 (right mastoid versus vertex).
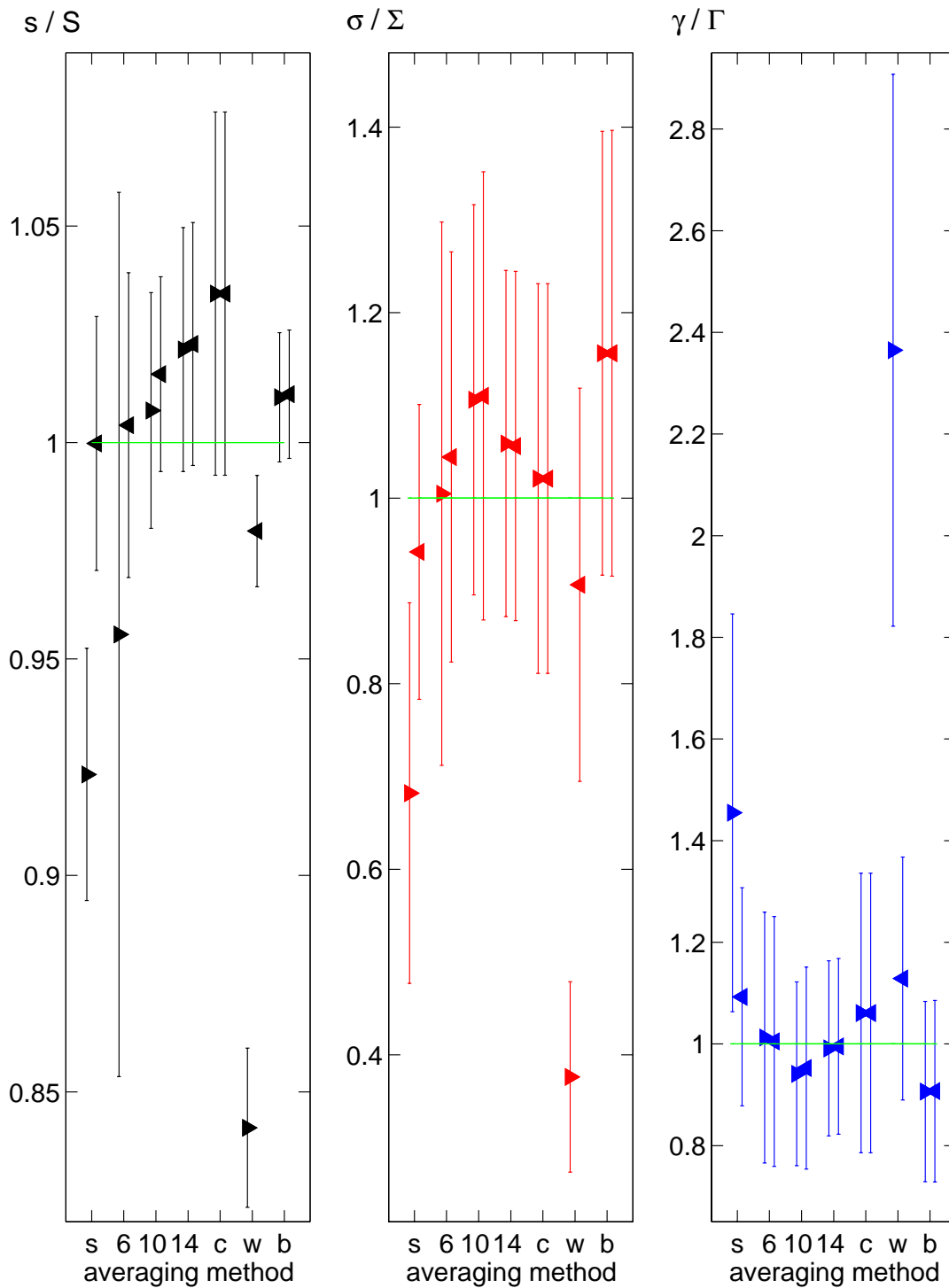
Fig. 2.11: Quotients of estimated signal rms s and true signal rms S (left), estimated residual noise rms σ and true residual noise Σ (middle), and estimated SNR γ and true SNR Γ (right) for various averaging methods. Labels of averaging methods are the same as in Fig. 2.3. As in the previous figure, the data left of the respective abscissa value are not iterated, while the data to the right side are iterated. Triangles indicate mean values, error bars denote one standard deviation. Data from diotic stimulation at 60 dB normal hearing level for channel M2 (right mastoid versus vertex).
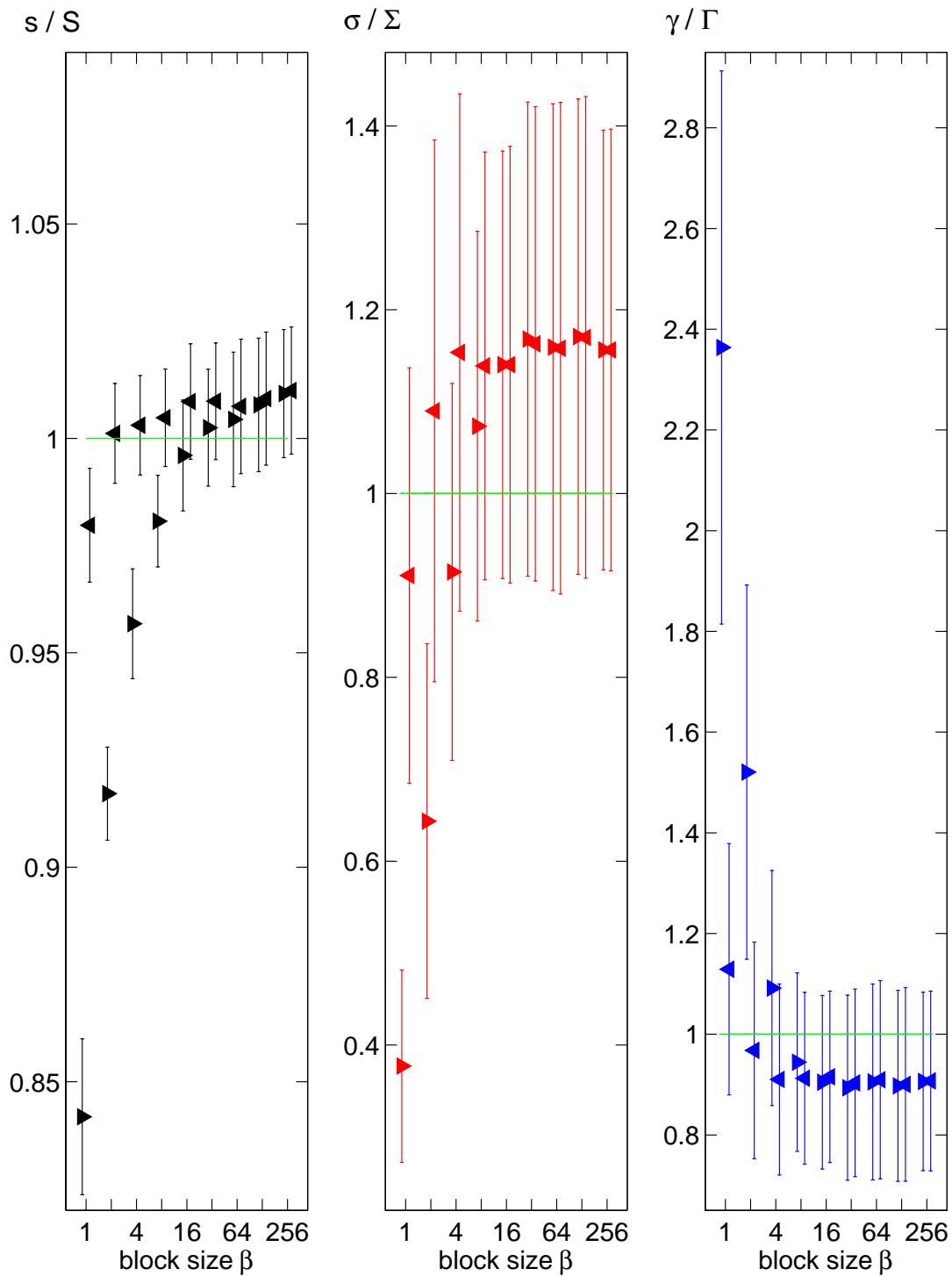
Fig. 2.12: Quotients of estimated signal rms s and true signal rms S (left), estimated residual noise rms σ and true residual noise Σ (middle), and estimated SNR γ and true SNR Γ (right) as a function of the block size β for the block weighted averaging method. Presentation and scales are identical to the previous figure to facilitate comparison. Data from diotic stimulation at 60 dB normal hearing level for channel M2 (right mastoid versus vertex).
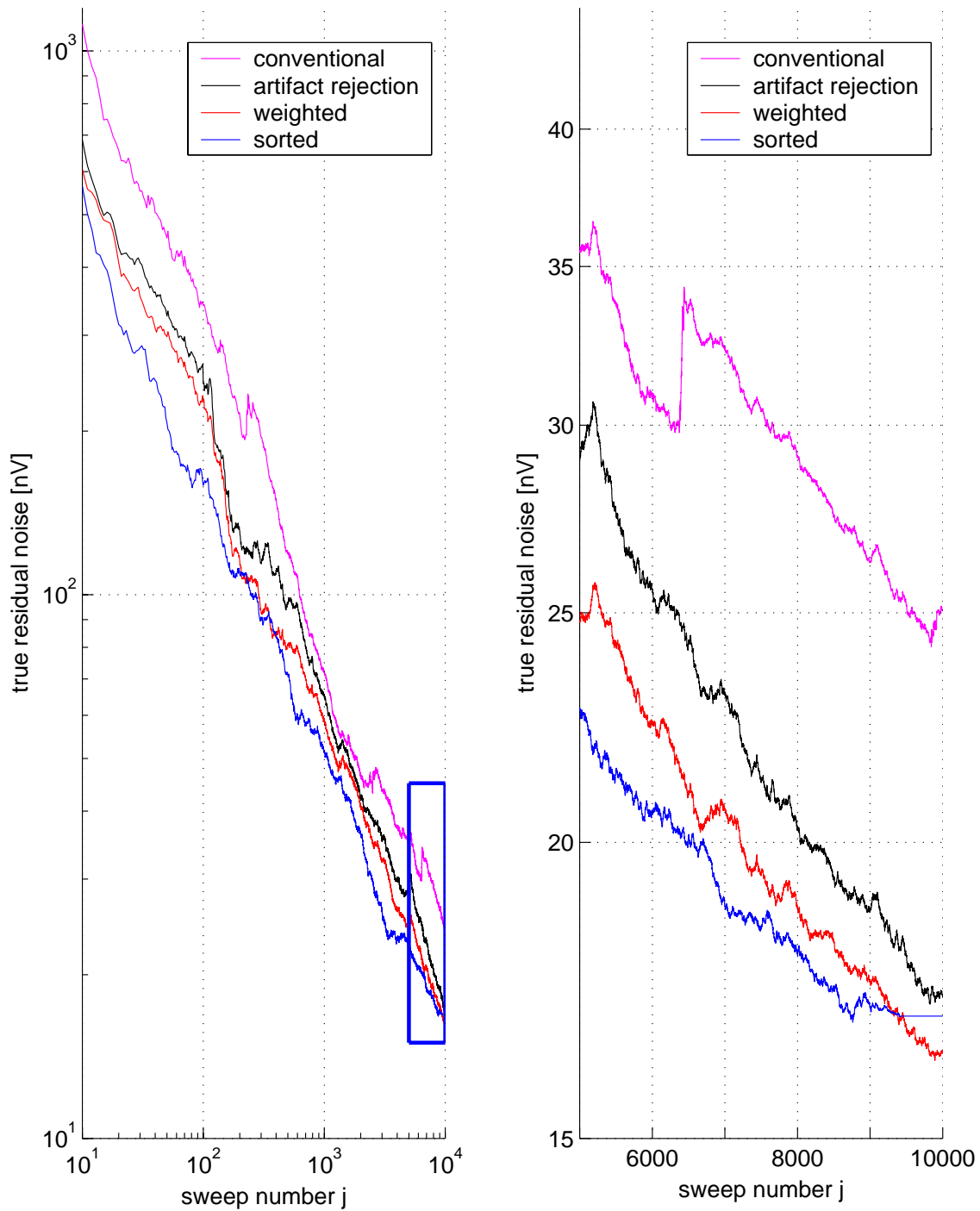
*Fig. 2.13:* **Left graph:** *True residual noise averaged over subjects after one iteration step* $(\Sigma^{(2)})$ *in dependence on the sweep number j (j' for sorted averaging) for four averaging methods From top to bottom: conventional $(\Sigma_c^{(2)})$, artifact rejection $\pm 10\,\mu V$ $(\Sigma_{a_{10}}^{(2)})$, weigthed $(\Sigma_w^{(2)})$, and sorted $(\Sigma_s^{(2)})$. Data from diotic stimulation at $60\,dB$ normal hearing level for channel M1 (left mastoid versus vertex).* **Right graph:** *Magnification of the left graph for sweep numbers 5.000 to 10.000. At about $j = 9400$ iterated sorted averaging cannot further decrease residual noise and is outperformed by iterated weighted averaging.*

2.4 Discussion

Various averaging methods known from the literature were applied and compared with respect to their ability to estimate ABR waveforms accurately. It was shown, both theoretically in the companion paper (Granzow et al., 2001) and empirically here, that single sweep based estimation of signal and residual noise is superior to average based estimation. For all averaging methods the new concept of iterative averaging was investigated. The iteration technique does not strongly influence the results for conventional averaging or averaging using an artifact criterion. However, the improved estimation of the power of a single epoch results in much better signal, noise, and SNR estimates in the case of sorted and weighted averaging. Table 2.1 shows that there is a considerable advantage of iterated weighted averaging. For all residual noise criteria, the number of sweeps that have to be included is lowest for this method ($J_w^{(2)} < J_{a_{10}}^{(2)} < J_c^{(2)}$). For subjects with strongly inhomogeneous noise (background EEG) the advantage of the iterated weighted averaging scheme is more pronounced, while the difference from the classical method using the $\pm 10\,\mu$V artifact criterion becomes negligible for subjects with more homogeneous noise.

The first approach to estimate SNR and residual noise on a single-sweep basis was the single-point variance introduced by Elberling and Don (1984) and Don et al. (1984). Fig. 2.6 of the present paper shows that the standard error $\sigma^{(1)}(t)$, i.e., the residual noise, does not vary much over time, i.e., within the epoch. There are only small differences in the residual noise estimate if analysis is performed at different instances of time. Therefore, the approach underlying the single-point variance method can be justified on the basis of our data. However, on the basis of Monte-Carlo simulations, Cebulla et al. (2000a,b) showed that a residual noise estimate based on all samples of the epochs is advantageous, especially for small numbers of epochs entering the average.

The estimation of the residual noise using single sweep information has a large impact on the accuracy of peak detection. The typical clinical question is to decide if a given ABR-component is a response or not. If we assume Gaussian measurement errors, the residual noise based on single sweeps, i.e., the standard error σ, allows for far more precise statements about the significance of peaks than the residual noise estimated on averages.

The SNR improvement of weighted averaging was investigated by Lütkenhöner et al. (1985). Their conclusion was that for a homogeneous ensemble of epochs weighted averaging does not improve the SNR because both signal and noise are reduced by the same factor. They showed that for inhomogeneous sweep ensembles weighted averaging is superior to conventional averaging in terms of a better SNR, but leads to a systematic

underestimation of the signal. To overcome this problem and maintain the advantage of weighted averaging, the present study demonstrates that the undesired underestimation is effectively inhibited by the iteration procedure employed here (cf. Fig. 2.11).

Don and Elberling (1994) analyzed the effect of variation of the block size β in the block-weighted averaging scheme. They compared the residual variance using 256, 128, 64, and 32 sweeps per block and 1, 2, 4, and 8 'single' time points per sweep, respectively, yielding always 256 data points to estimate the variance/power of a block. The number of time points per sweep used for noise estimation was not increased above 8, because they argued that there is only a limited number of degrees of freedom in a single epoch, i.e., a small number of independent samples in the band-limited noise signal.

In accordance with their results, our findings for the block-weighted averaging method confirm that a block-weighted average with a small number of epochs for the estimation of the weightings is not optimal. Without iteration, the lowest acceptable block size was found to be eight in our study. However, using the iterated noise estimation, we observed that the true residual noise (see Fig. 2.10) as well as the estimated signal and noise (Fig. 2.12) do not strongly depend on the block size. In contrast to Elberling and Don, we therefore conclude that the main error in estimating the residual noise is due to a bias produced by the "desired" signal which is removed by the iteration procedure. The small remaining dependence of the residual noise on the block size may be due to the small number of degrees of freedom although a second iteration step seems to further diminish the dependence of the residual noise on the block size (data not shown).

With iterative averaging, the noise in a single sweep can therefore be estimated more accurately than without iteration. Hence, it is not necessary to form a block-weighted average in order to improve the estimate of the noise from a block of epochs.

For ABRs the approximation of eq. (2.4) is very good since the SNR of single epochs is very low. For evoked and event related potentials generated in cortical areas the above approximation is worse since the SNR of the single epochs is generally higher than in ABR recordings. It is therefore to be expected that more pronounced differences between iterated and non-iterated averages will be seen in these cases. Iterative averaging should produce better signal and noise estimates. Additionally, due to the smaller number of sweeps required for late evoked potentials the computational cost for the single-sweep-based methods is at least an order of magnitude smaller than for ABRs.

Taking into account the following three aspects – elimination of the arbitrariness of the artifact threshold, small residual noise (cf. Fig. 2.9, Tab. 2.1) and good estimation properties

(Fig. 2.11) – weighted averaging in iteration appears to be the most favorable averaging method.

2.5 Summary and conclusions

- For accurate estimation of the signal and the residual noise, the methods based on single sweeps are shown to be superior to those based on averages.

- Weighted averaging avoids the arbitrariness of the choice of the artifact threshold.

- The effect of underestimation of signal and noise of weighted averaging can be overcome by the use of the iteration procedure.

- For a given number of recorded sweeps, iterative weighted averaging provides the most reliable estimates of the signal and the residual noise, while iterated sorted averaging appears to be the second best method.

Chapter 3

Auditory brain stem responses evoked by lateralized clicks[1]

Abstract

The dependence of binaurally evoked auditory brain stem responses and the binaural difference potential on simultaneously presented interaural time- and level differences is investigated in order to assess the representation of stimulus lateralization in the brain stem. Auditory brain stem responses to binaural click stimuli with all combinations of three interaural time- and three interaural level differences were recorded from 12 subjects and 4 channels. The latency of Jewett wave V is shortest for zero interaural time difference and longest for the trading stimuli. The amplitude of wave V is largest for centrally perceived stimuli, i.e., the diotic and trading stimuli, and smallest for the most laterally perceived stimuli. The latency of the most prominent peak of the binaural difference potential DN1 mainly depends on the interaural time difference. The amplitude of the components of the binaural difference potential, DP1-DN1, depends similarly on stimulus conditions as wave V amplitude in the case of the binaural stimuli: smallest amplitudes are found for the most lateral stimuli and largest amplitudes for central stimuli. The results demonstrate that interaural level- and time differences are not processed independently. This supports the hypothesis that directional information in humans is already extracted and represented at the level of the brain stem.

---

[1] A modified version of this chapter was published in Hear. Res. 163 (1-2), 12–26: Riedel and Kollmeier (2002a): "Auditory brain stem responses evoked by lateralized clicks: Is lateralization extracted in the human brain stem ?"

3.1   Introduction

From psychoacoustical studies it is known that the most important cues for directional hearing are the interaural time difference (ITD) and the interaural level difference (ILD) (Feddersen et al., 1957; Mills, 1958, 1960; Colburn and Durlach, 1978). Presenting a stimulus with an ITD and an ILD that point towards the same direction ('synergistic presentation') leads to a further lateralized perception than presenting the same respective ITD or ILD alone. In contrast, the 'antagonistic presentation', when ILD and ITD point to opposite lateral positions, results in more central percepts. The relation between an ILD and the ITD that causes the same lateralization is commonly expressed in terms of the time-intensity-trading ratio (TIR). It can be measured either by subjectively matching the respective lateralization produced by ILD and ITD or by trading ILD against ITD resulting in a centered image. TIRs were measured for a variety of stimuli and levels (e.g., Durlach and Colburn, 1978; Nordby et al., 1982; Furst et al., 1985; McPherson and Starr, 1995; Damaschke et al., 2000).

From neuroanatomical studies it is known that left and right auditory afferent fibers first intersect in the superior olive (SO) in the brain stem (Nieuwenhuys et al., 1988). Neurophysiological studies in animals demonstrate binaural processing in cells of the SO and subsequent stations of the auditory brain stem, namely the nuclei of the lateral lemnisci (NLL) and the inferior colliculus (IC) (Semple and Aitkin, 1979; Achor and Starr, 1980; Caird and Klinke, 1983; Caird et al., 1985; Yin and Chan, 1990; Popper and Fay, 1992; Joseph and Hyson, 1993; Gummer and Zenner, 1996; van Adel et al., 1999).

In evoked response studies binaural processing or binaural interaction is assessed in terms of the binaural difference potential (BD). It is defined as the difference between the potential obtained with binaural stimulation and the sum of the potentials obtained with monaural stimulation, symbolically BD = B – (L + R). Any significant deviation from BD=0 is understood as hint to some non-linearity, i.e., a functional coupling of left and right signals. The BD was analyzed in a number of studies, often as a function of ILD and ITD (Dobie and Berlin, 1979; Dobie and Norton, 1980; Ainslie and Boston, 1980; Levine, 1981; Wrege and Starr, 1981; Gerull and Mrowinski, 1984; Kelly-Ballweber and Dobie, 1984; Furst et al., 1985; Ito et al., 1988; McPherson et al., 1989; Jones and van der Poel, 1990; McPherson and Starr, 1995; Jiang, 1996; Cone-Wesson et al., 1997; Brantberg et al., 1999a,b). Using 2000 sweeps per stimulus condition Gerull and Mrowinski (1984) contended the null hypothesis BD = 0, i.e., that the binaural response B can be perfectly predicted by the sum of the monaural responses L + R, supporting the view of two indepen-

dent channels. Ainslie and Boston (1980), also using 2000 sweeps, found a non-vanishing BD but explained it by acoustic crosstalk (ACT). A thorough investigation of possible artifacts resulting in an artificial BD was performed by Levine (1981). He described two possible sources of artifacts: acoustic crosstalk and the middle ear reflex (MER). These have to be ruled out before associating a measured BD with neural processes.

Furst et al. (1985) suggested that the first major peak in the BD, DN1 or β in their nomenclature, is a physiological correlate of the categorial percept of binaural fusion. DN1 was present up to ITDs of 1 ms with a relatively constant amplitude, but undetectable for ITDs longer than 1.2 ms. With increasing ILD, DN1 amplitude decreased gradually and became undetectable for ILDs greater than 30 dB. Brantberg et al. (1999a) studied DN1 as a function of the ITD and found approximately constant amplitudes for ITDs up to 1 ms. In contrast, McPherson and Starr (1995) reported that the DN1 component gradually decreased with increasing ILD and ITD, and became undetectable for ILD > 16 dB and ITD > 1.6 ms. They stated an inverse correlation between DN1 amplitude and the psychophysical lateralization (introduced by either ILD or ITD).

The major problem of measuring the BD is its poor signal-to-noise-ratio (SNR). The BD is about an order of magnitude smaller than binaural responses, and its residual noise is about twice as large as that of a directly measured potential due to the calculation of sums and differences. Sweep numbers of about 2000, which are generally sufficient in auditory brain stem response (ABR) recordings, may result in a poor SNR of the BD. Hence, for a reliable detection of BD components high quality recordings and a precise judgement of the SNR are required. In the present study, this was accomplished by storing all unfiltered single sweeps to disk, offline linear phase filtering and iterative, weighted averaging. The residual noise was calculated as the standard error of the mean over all filtered single sweeps, or more precisely, as the rms-value of the time dependent standard error $\sigma(t)$ for every channel, stimulus condition and subject (Riedel et al., 2001). This is in contrast to other studies, where the residual noise was estimated from the difference of two averages recorded, (e.g., Wrege and Starr, 1981) or from the average in the prestimulus interval, (e.g., Furst et al., 1985).

In all studies known to the authors, only the effect of a single parameter (either ILD or ITD) on the BD was studied without relating it to the psychophysical effect of lateralization which arises from a combination of both paramters. The aim of the present study therefore is to analyze the dependence of ABRs and BDs on psychophysical lateralization if not only a single cue, ILD or ITD, is presented, but also for synergistic and trading stimulus configurations. The question is whether stimuli with similar lateralization evoke

similar responses. If so, this would imply that a representation of the laterality of a stimulus would already exist at brain stem level, and that ILD and ITD were not processed independently.

3.2 Methods

3.2.1 Subjects

Twelve subjects from the staff of the University of Oldenburg (3 females, 9 males) between the ages of 25 and 36 participated voluntarily in this study. They had no history of audiological or neurological problems and were classified as normal hearing by routine audiometry. The audiometric loss was less than 10 dB for frequencies below 4 kHz and less than 15 dB for the higher frequencies.

3.2.2 Stimuli

Rarefaction click stimuli were produced by applying rectangular voltage pulses of 100 μs duration to Etymotic Research ER-2 insert earphones. The time interval between the onsets of two successive stimuli was chosen to vary randomly and equally distributed between 62 and 72 ms, yielding an average stimulation rate of approximately 15 Hz. A 700 ms segment of the click train comprising 11 clicks was used to determine the thresholds in quiet. They were measured three times by all subjects with a 3-alternative forced-choice method in conjunction with a 2-down–1-up algorithm for both ears and averaged over runs, subjects and ears. The threshold level – referred to as 0 dB normal hearing level (nHL) – corresponds to 39 dB peak equivalent sound pressure level (peSPL)[2]. The standard deviation of the individual thresholds from the averaged threshold was 3 dB. For reference also the thresholds for single clicks were determined. On average, single click thresholds were 5 dB higher than click train thresholds.

15 stimulus conditions were tested, 9 binaural and 6 monaural. The monaural clicks were presented at the levels 53, 59 and 65 dB nHL and are denoted as L–m, L0m, L+m and R–m, R0m, R+m for monaural left and right stimulation, respectively. The binaural stimuli, see Fig. 3.1, were the nine possible combinations of 3 ITDs (–0.4, 0 and 0.4 ms) and 3 ILDs (–12, 0 and 12 dB).

---

[2]A sinusoid of frequency 1 kHz with the same peak-to-peak-amplitude showed 39 dB SPL in a Brüel & Kjær (B&K) amplifier type 2610. The calibration was performed using a half inch microphone (B&K 4134) with an artificial ear and a preamplifier (B&K 2669).

The binaural stimuli are named as follows: the first letter refers to the perceived lateralization of the stimuli: 'L' for left, 'C' for center and 'R' for right. The second and third characters ('–', '0' and '+') are used to specify the ILD and ITD, respectively. For example, for diotic stimulation to both ears a click at 59 dB nHL was presented simultaneously. This stimulus C00 is found in the center of the diagram (ILD=ITD=0). The stimulus R+0 in the middle of the top row has zero ITD, but is lateralized to the right due to its positive ILD. On the other hand, the stimulus R0+ at the right of the middle row has zero ILD, but is also lateralized to the right due to its positive ITD. The arrows in Fig. 3.1 point into the approximate direction of the lateralization of the stimuli. Both an ITD of 0.4 ms and an ILD of 12 dB cause a strong, but not extreme lateralization of about 70° (Furst et al., 1985; McPherson and Starr, 1995). A stronger, almost complete lateralization is produced by the synergistic stimuli L–– and R++ whose ILDs and ITDs point into the same direction. In contrast, the stimuli C+– and C–+ refer to the antagonistic situation: ILD and ITD act in opposite direction resulting in a centered image. In the lower left corner of each subplot in Fig. 3.1, the respective binaural stimulus is depicted schematically. In the ITD-ILD-plane lines of equal lateralizaton are the diagonal dotted lines. Identical stimuli were used for all subjects for better comparability of the results. Comparatively large values for the ITD and the ILD were used to obtain as large differences in the evoked potentials for the different stimulus conditions as possible without leaving the physiological range.

3.2.3 Electrodes

For the ABR recordings Ag/AgCl-electrodes were used. The 4 active channels were placed at the left (A1) and right (A2) mastoid and the parieto-occipital positions PO9 and PO10 according to the extended 10-20-system (Jasper, 1957; Sharbrough et al., 1991). The common reference electrode was placed at the vertex (Cz), the ground electrode at the forehead (Fpz). Electrode impedances were measured at a test signal frequency of 30 Hz and brought well below 5 kΩ, common values were 2-3 kΩ. Since DC recordings were performed, the criteria for a good contact between electrodes and skin were both a low impedance and a vanishing voltage drift seen in the raw EEG signal.

3.2.4 Recordings

During the ABR recordings, subjects lay in a sound insulated and electrically shielded room. They were instructed to relax and lie as comfortably as possible. ABRs were recorded with a DC-coupled differential amplifier (Synamps 5803). Inside the shielded

room the EEG was preamplified by a factor 150, further amplified by the main amplifier by a factor 33 resulting in a total amplification of 74 dB. The voltage resolution was approximately 16.8 nV/bit. The sweeps were filtered by an analog antialiasing-lowpass with a corner frequency of 2 kHz, digitized with 10 kHz samplingrate and 16 bit resolution, and stored to hard disk.

During the recording the artifact rejection level was set to $\pm 500\,\mu$V, since filtering, artifact analysis and averaging was done offline. The clipping level of the DA-converters was $\pm 550\,\mu$V. The recording interval comprised 400 samples in the time interval from -15 to 25 ms relative to stimulus onset. For the dichotic stimuli the leading click was defined as stimulus onset.

10.000 single sweeps for all of the 15 stimuli were recorded in the following manner: The 5 stimuli belonging to a certain ILD (a row in Fig. 3.1, e.g., L–m, C+–, R+0, R++ and R+m) were presented in random order on a sweep-by-sweep basis. One run consisting of 12500 stimuli, 2500 of each type, lasted approximately 14 minutes. After each run the impedances were checked, and adjusted if necessary. Afterwards the runs belonging to the two other ILDs were recorded. Totally, 12 runs (4 repetitions of the 3 runs) were carried out for every subject. The duration of the recording session was about 3 hours without preparation and pauses.

No contralateral masking was used during the monaural presentation. This has been used in other studies to avoid acoustic crosstalk (ACT) (Jones and van der Poel, 1990; McPherson and Starr, 1995; Brantberg et al., 1999a). Instead, following Levine (1981) and Ito et al. (1988), left, right and binaural stimuli were presented randomly without contralateral masking. In combination with the use of insert earphones and moderate presentation levels of $\leq$65 dB nHL, this avoids contribuations of ACT and the middle ear reflex to the binaural difference potential.

Before averaging, the single sweeps were filtered with a linear phase FIR bandpass with 200 taps and corner frequencies 100 and 1500 Hz (Granzow et al., 2001). An iterated weighted average of the filtered sweeps was computed for all subjects and stimulus conditions. The residual noise of the averages was computed as the standard error σ across sweeps, see chap. 2.

3.2.5   Binaural interaction

The binaural interaction was computed in terms of the binaural difference potential BD = B – (L + R). This was done channel-wise and sample by sample. The six monaural stimuli

were chosen to allow for the computation of the BD for all 9 binaural stimulus conditions. For the stimuli with an ITD, the monaural response of the lagging ear was digitally delayed by the ITD before computing the BD. For the stimuli with an ILD, the monaural stimuli with the corresponding level were used, e.g., $BD_{C+-} = C+- - (L-m + R'+m)$ with R'+m being the delayed version of R+m. All 9 BDs were computed from stimulus triplets which had been recorded quasi-simultaneously, i.e., during the same measurement run. This avoids artifacts in the BD components due to long term changes of the recording conditions or subject's state. The residual noise of the BD was estimated as the square root of the summed variances of the three measurements, e.g., $\sigma_{BD_{C00}} = (\sigma_{C00}^2 + \sigma_{L0m}^2 + \sigma_{R0m}^2)^{1/2}$, assuming that C00, L0m and R0m are statistically independent.

3.2.6 Peak identification

To increase the accuracy of amplitude and latency measurements, data were interpolated by a factor of 10, i.e., they were upsampled to convert the sampling rate from 10 to $100\,\mathrm{kHz}$. This was accomplished by zero-padding in the spectral domain which in the time domain corresponds to a convolution with a sinc-function. Since the original analog signal was band-limited to frequencies below $2\,\mathrm{kHz}$, a perfect interpolation was possible.

Peaks in the interpolated signal were identified by a sign change in its derivative. For baseline-to-peak measurements peaks with voltages V_{bp} smaller than 2σ were not regarded as significant and hence were discarded. In ABR measurements amplitude histograms of single sweeps show very good approximations of Gaussian distributions. The residual noise can therefore be interpreted as the standard error of the Gaussian measurement error σ, see chap. 2. With a confidence of 95% the true evoked potential is then in the interval $[V - 2\sigma \ V + 2\sigma]$ with V being the measured potential. For peak-to-peak-measurements peaks with voltages V_{pp} greater than $\sqrt{2} \cdot 2\sigma$ were accepted. The additional factor of $\sqrt{2}$ is due to the fact that the variances of both peaks in the pair add up. Latency errors were estimated from the amplitude errors and the curvature of the peaks according to Hoth (1986).

Automatic labelling of the peaks was consistent for wave V in the monaural and binaural stimulus conditions. Here, latencies and amplitudes of wave V were determined for all 12 subjects and all 15 stimulus conditions. Amplitudes were measured baseline-to-peak for two reasons: first, the 5-ms-baseline containing 50 samples is well defined, second, peak-to-peak-measurements V - VI' would yield erroneous amplitudes for the subjects exhibiting muscular artifacts: in three out of 12 subjects muscle artifacts with latencies from 8-12 ms in channels A1 and A2 interfere with wave VI' at a latency of about $8\,\mathrm{ms}$.

In the case of the binaural difference potential, the components were labelled manually due to to their larger variability. The convention introduced by Ito et al. (1988) (see their Fig. 1) was adopted. The first main component of the BD is the negative wave DN1 preceeded by a smaller positive wave labelled DP1. DN1 corresponds to the β-wave described by Levine (1981). BD amplitudes were measured peak-to-peak from DP1 to DN1 because the baseline of the BD shows larger fluctuations than for the monaural and binaural responses. Since the latencies of DP1 and DN1 do not deviate more than 0.5 ms from the latency of the binaural wave V there is no interference with muscular artifacts for these BD components. Latencies of the larger DN1 component were analyzed.

3.3   Results

3.3.1   Monaural and binaural potentials

Fig. 3.2 shows the recordings for all stimulus conditions, all channels and one subject. The stimuli are arranged as in Fig. 3.1. The binaural responses exhibit considerably larger amplitudes than the monaural responses. The highest values of wave V amplitude, denoted as A_V, are reached for the diotic (C00) and for the antagonistic stimuli (C+− and C−+). These stimuli are perceived more or less in the center of the head. With growing lateralization, either due to ILD or ITD, A_V decreases about the same amount. A further amplitude reduction is observed for the more lateralized synergistic stimuli.

Fig. 3.3 shows the dependence of wave V amplitude A_V on the stimulus parameters. In the left column, A_V, including error bars, is plotted two-dimensionally. The ILD is coded by different gray-scale values. In the right column, the same data are replotted three-dimensionally to illustrate A_V as a function of ILD and ITD, i.e., over the plane spanned by these parameters.

In the upper row, data for a single channel (PO10) and a single subject (cr) are plotted. As in the previous figure, the highest amplitudes are observed for the centrally perceived stimuli. This can be seen as a 'ridge' along the central diagonal in the three-dimensional graph. A_V decreases with increasing lateralization. Antagonistic stimuli have significantly higher amplitudes than synergistic stimuli, there is no overlap of the small intraindividual standard errors. The second row shows the mean over the four channels for the same subject. The third row depicts the mean over 12 subjects for channel PO10 and the bottom row the mean over channels and subjects. Apparently, interindividual standard deviations (3rd and 4th row) are larger than intraindividual standard errors (1st and 2nd row). This

is due to the large variance of ABRs across subjects. However, the differences obtained for different stimulus conditions are similar for all subjects. To reveal those differences, a signed Wilcoxon rank test was performed for all pairs of binaural stimuli. Tab. 3.1 (upper right triangle) summarizes the results for the amplitudes from the bottom row (mean over channels). Stimuli are rearranged in three groups according to their lateralization: (i) three central stimuli: C00, C–+, and C+–, (ii) four stimuli either lateralized by ILD or ITD: L–0, L0–, R+0, R0+, for convenience called the lateralized stimuli, and (iii) two synergistic stimuli: L–– and R++. Three significance levels were tested: significant differences for $\alpha = 0.05$, 0.01 and 0.001 are marked by '*', '**' and '***', respectively, 'ns' means not significant. There are no significant amplitude differences within the central group. With the exception of L–0 all lateralized and synergistic stimuli have significantly different (smaller) wave V amplitudes than the central stimuli. There are also significant differences between the groups of lateralized and synergistic stimuli. With one exception (L–– / R+0) amplitudes of the synergistic stimuli are significantly smaller than amplitudes of the lateralized stimuli.

	C00	C–+	C+–	L–0	L0–	R+0	R0+	L––	R++
C00	—	ns	ns	ns	**	***	***	**	***
C–+	***	—	ns	ns	**	**	***	**	***
C+–	***	ns	—	ns	***	***	**	**	***
L–0	*	***	***	—	*	*	**	***	***
L0–	***	**	***	***	—	ns	ns	*	***
R+0	*	***	***	*	***	—	ns	ns	***
R0+	***	***	**	***	ns	***	—	*	***
L––	*	**	***	ns	**	**	ns	—	*
R++	ns	***	***	ns	***	ns	***	*	—

Tab. 3.1: Differences between amplitudes (upper right triangle) and latencies (lower left triangle) of wave V for all pairs of binaural stimulus conditions as revealed by signed Wilcoxon rank tests across all subjects. Average data over channels (lower right panel in Fig. 3.4) were used for analysis. Three significance levels were tested: $\alpha < 0.05$ (*), $\alpha < 0.01$ (**) and $\alpha < 0.001$ (***), 'ns' stands for not significant.

Due to the high curvature of the peaks, individual latencies of wave V can be determined with high accuracy. Intraindividual standard latency errors for all subjects and stimulus

conditions vary between 0.016 and 0.069 ms, mean 0.026 ms. The interindividual standard errors of wave V latencies are about an order of magnitude larger and vary between 0.19 and 0.29 ms, mean 0.23 ms. Fig. 3.4 shows the mean amplitude of wave V (A_V) as a function of the mean latency of wave V (t_V), for all channels as well as for the mean over channels. In addition, the lower left triangle in Tab. 3.1 shows the results of the signed Wilcoxon rank test for all pairs of binaural stimulus conditions to reveal the latency differences. Latencies of the antagonistic stimuli are significantly longer than those for all other binaural stimuli. The shortest latencies are observed for the synergistic stimuli and those with ITD = 0 ms. The monaural stimuli exhibit the usual inverse relation between latency and amplitude: with increasing level A_V increases while t_V decreases. For the binaural stimuli with non-vanishing ITD this relation is inverted: The synergistic stimuli L–– and R++ show shortest latencies and smallest amplitudes. With decreasing lateralization both amplitude A_V and latency t_V increase.

3.3.2   Binaural difference potentials

Fig. 3.5 illustrates the computation of the BD from the binaural and monaural responses for one subject, one channel and diotic stimulation. The error bars show ± 3 standard errors corresponding to a 99.7%-confidence-interval for Gaussian measurement errors. Filled triangles indicate extrema whose peak-to-peak voltages V_{pp} exceed $\sqrt{2}\cdot 3\sigma$. Open triangles denote extrema with $V_{pp} > \sqrt{2}\cdot 2\sigma$ that are significant only at a 95%-level. This convention is maintained for all following figures.

Standard errors of monaural and binaural responses are of comparable size since they were averaged by the same number of sweeps (10000). The standard errors of the composed responses $\sigma_\mathrm{L0m+R0m}$ and $\sigma_\mathrm{BD_{C00}}$ are higher by a factor of about $\sqrt{2}$ and $\sqrt{3}$, respectively. This reflects the addition of the variances when adding or subtracting responses. For waves V and VI the binaural response has a slightly shorter latency than the monaural responses and their sum. The BD components DP1 and DN1 are associated with the rising and falling slope of wave V, respectively. Analogously, but not as clear as for wave V, the BD components DP2 and DN2 can be associated with the rising and falling slope of wave VI, respectively.

Fig. 3.6 compares the binaural difference potential for all subjects in the diotic stimulus condition. Interindividual differences are mainly due to differences in residual noise level. Subjects with high residual noise level, e.g., ib and channel A2 from mk, do not show systematic BD components. Muscular artifacts in the mastoidal channels are reduced by

the difference operation, but are still observable for subjects jd,kt and ow. However, the BD waves DP1 and DN1 are not affected by these artifacts. Subjects with low noise level, e.g., cr, dj, hr and kw, show a systematic BD with clear components DP1 and DN1. A DP2-DN2 complex is less reliably found, it exists for subject jo, for other subjects only in some channels. DP2 is better identifiable than DN2. Typical is the V-shaped DP1-DN1-DP2 complex.

In contrast to the monaural and binaural responses peak-to-peak measurements of amplitudes are preferable for BDs due to the low SNR and poorly defined baseline. A pair of consecutive BD components is considered as significant if its peak-to-peak-value V_{pp} exceeds $\sqrt{2} \cdot 2\sigma_{\mathrm{BD}}$ for the channel and subject considered, i.e., if its SNR is $\geq 6\,\mathrm{dB}$.

Tab. 3.2 summarizes amplitudes, residual noise and SNR for wave V and BD wave DP1-DN1 for all subjects. Data are mean values over the four channels for diotic stimulation. Due to the quasisimultaneous measurement of the monaural and binaural responses the residual noise σ is nearly constant for all stimulus conditions. Data are sorted according to the SNR of the BD component DP1-DN1 which varies between 7.4 and 15.2 dB. Higher wave V SNRs do not necessarily entail higher SNRs of BD wave DP1-DN1. The standard error of DP1-DN1 is roughly a factor $2.4 \approx \sqrt{6}$ higher than the standard error of wave V. This can be explained by the following consideration: the addition and subtraction of responses to yield the BD contributes a factor $\sqrt{3}$, another factor $\sqrt{2}$ is produced by the peak-to-peak-measurement in the case of the BD compared to the baseline-to-peak-measurement for wave V.

Fig. 3.7 shows the dependence of average BD amplitudes $A_{\mathrm{DP1-DN1}}$ on stimulus parameters. In the left column, the BD amplitude, DP1-DN1, including error bars, is plotted two-dimensionally. The ILD is coded by different gray-scale values. In the right column the same data are replotted three-dimensionally to illustrate $A_{\mathrm{DP1-DN1}}$ as a function over the parameter plane spanned by ILD and ITD.

In the upper row, data from subject dj, channel A2, are presented. The same systematic dependence of the BD amplitude on stimulus conditions as for ABR wave V (see Fig. 3.3) is found: central stimuli (C00, C+− and C−+) exhibit the highest amplitudes. Except for $A_{\mathrm{R0+}} > A_{\mathrm{C-+}}$, all stimuli lateralized only by ILD or ITD (L–0, L0– and R+0) show smaller amplitudes. A further amplitude reduction of DP1-DN1 is seen for the synergistic stimuli L–− and R++. However, the BDs exhibit larger intraindividual standard errors than the binaural responses. In the second row, the mean over channels for the same subject is shown. The above mentioned exception disappeared and the systematic dependency of

subject	ABR wave V			BD wave DP1-DN1		
	$A[\mu V]$	$\sigma[nV]$	SNR[dB]	$A[\mu V]$	$\sigma[nV]$	SNR[dB]
dj	0.69	15.4	33.2	0.21	37.7	15.2
ow	0.48	15.2	30.4	0.14	37.1	13.7
jo	0.65	20.8	29.9	0.24	50.9	13.4
hr	0.55	12.9	33.4	0.12	31.7	12.1
cr	0.71	15.4	25.6	0.15	37.7	11.7
mk	0.42	25.5	32.8	0.16	62.0	11.7
kw	0.62	13.5	33.5	0.12	33.1	11.5
jd	0.72	23.6	30.7	0.16	57.0	10.7
hk	0.33	12.9	28.5	0.10	31.6	10.5
kt	0.51	22.5	26.9	0.19	54.7	10.2
rh	0.47	16.9	29.1	0.10	41.5	7.4
ib	0.61	25.6	27.7	—	62.7	—
mean	0.56	18.4	30.1	0.15	44.8	11.8

Tab. 3.2: Amplitude A, residual noise σ and SNR of ABR wave V and BD wave DP1-DN1 for all subjects. Mean over channels, diotic stimulation.

BD amplitude and lateralization is clearly visible in the 'ridge' for the central stimuli in the three-dimensional plot. In the third and fourth row, average data over subjects for channels A2 and the mean over channels, respectively, are drawn. As in the upper two rows increasing amplitudes are found with decreasing lateralization. Interindividual standard deviations of the BDs are in the same order of magnitude as intraindividual standard errors. To evaluate the significance of the differences in BD amplitude, signed Wilcoxon rank tests were performed for all pairs of binaural stimulus conditions. In the upper right triangle of Tab. 3.3 the test results are shown for the mean data over channels (bottom row in Fig. 3.7).

Due to the smaller SNR of the BD in comparison to the binaural responses there are fewer significant differences as in the binaural case. Compared to the diotic stimulus condition (C00) BD amplitudes of the stimuli lateralized by the ITD (L0– and R0+) are significantly smaller, BD amplitudes of the stimuli with ILD (L–0 and R+0) do not show a significant amplitude decrement. Except for the pair C00 and L–– , BD amplitudes of the central

stimuli are significantly higher than for the synergistic stimuli.

	C00	C−+	C+−	L−0	L0−	R+0	R0+	L−−	R++
C00	—	ns	ns	ns	*	ns	***	ns	***
C−+	***	—	ns	ns	ns	ns	*	*	**
C+−	***	ns	—	ns	ns	ns	ns	*	***
L−0	ns	***	***	—	ns	ns	ns	ns	**
L0−	***	ns	ns	***	—	ns	ns	ns	*
R+0	ns	ns	**	ns	*	—	ns	ns	**
R0+	***	ns	ns	***	*	*	—	ns	*
L−−	**	ns	ns	***	ns	**	ns	—	ns
R++	***	ns	ns	**	*	***	ns	ns	—

Tab. 3.3: Differences between BD amplitudes DP1-DN1 (upper right triangle) and latencies of DN1 (lower left triangle) for all pairs of binaural stimulus conditions as revealed by signed Wilcoxon rank tests across all subjects. Average data over channels (lower right panel in Fig. 3.8) were used for analysis. Three significance levels were tested: $\alpha < 0.05$ (*), $\alpha < 0.01$ (**) and $\alpha < 0.001$ (***), 'ns' stands for not significant.

Fig. 3.8 shows the mean BD amplitude, $A_{\mathrm{DP1-DN1}}$, as a function of the mean latency of BD wave DN1 (t_{DN1}) for all channels as well as for the mean over channels, in a similar fashion as Fig. 3.4. The mean intraindividual standard latency error of BD wave DN1 averaged over all channels, subjects and stimulus conditions is 0.08 ms, the mean interindividual standard deviation of DN1 latency amounts to 0.29 ms. The lower left triangle in Tab. 3.3 gives the results of signed Wilcoxon rank tests for all pairs of binaural stimulus conditions to reveal significant latency differences. The stimuli with ITD = 0 ms (C00, L−0 and R+0) result in significantly shorter latencies than stimuli with non-vanishing ITD. In contrast to the amplitudes, the latencies of the BD wave DN1 do not clearly depend on the lateralization of the stimuli, but rather on ITD and partially on ILD. t_{DN1} is mainly determined by the ITD and is, on average, 0.21 ms $\approx$ ITD/2 longer for stimuli with ITD = ± 0.4 ms than for stimuli with ITD = 0 ms.
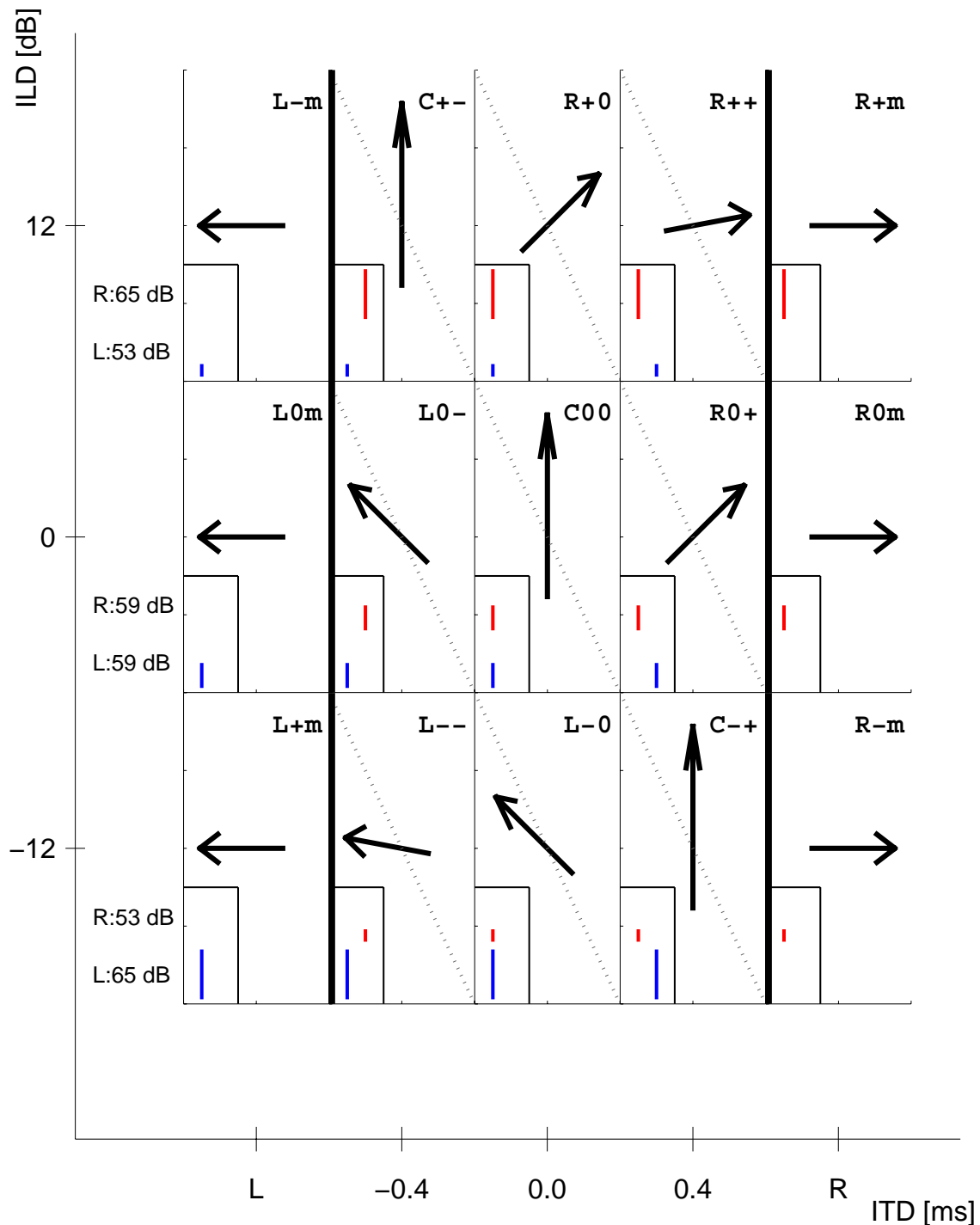
Fig. 3.1: Naming convention and lateralization of the stimuli: Centrally perceived stimuli are marked with a 'C'. Stimuli lateralized to the left and right side are marked with 'L' and 'R', respectively. For the monaural stimuli the second character indicates the level. For the binaural stimuli the second and third character denote the sign of ILD and ITD, respectively. Arrows point into the approximate direction of perceived lateralization. In the lower left corner of each subplot, the corresponding binaural stimulus is depicted. The dotted lines connect stimuli eliciting similar lateralization.
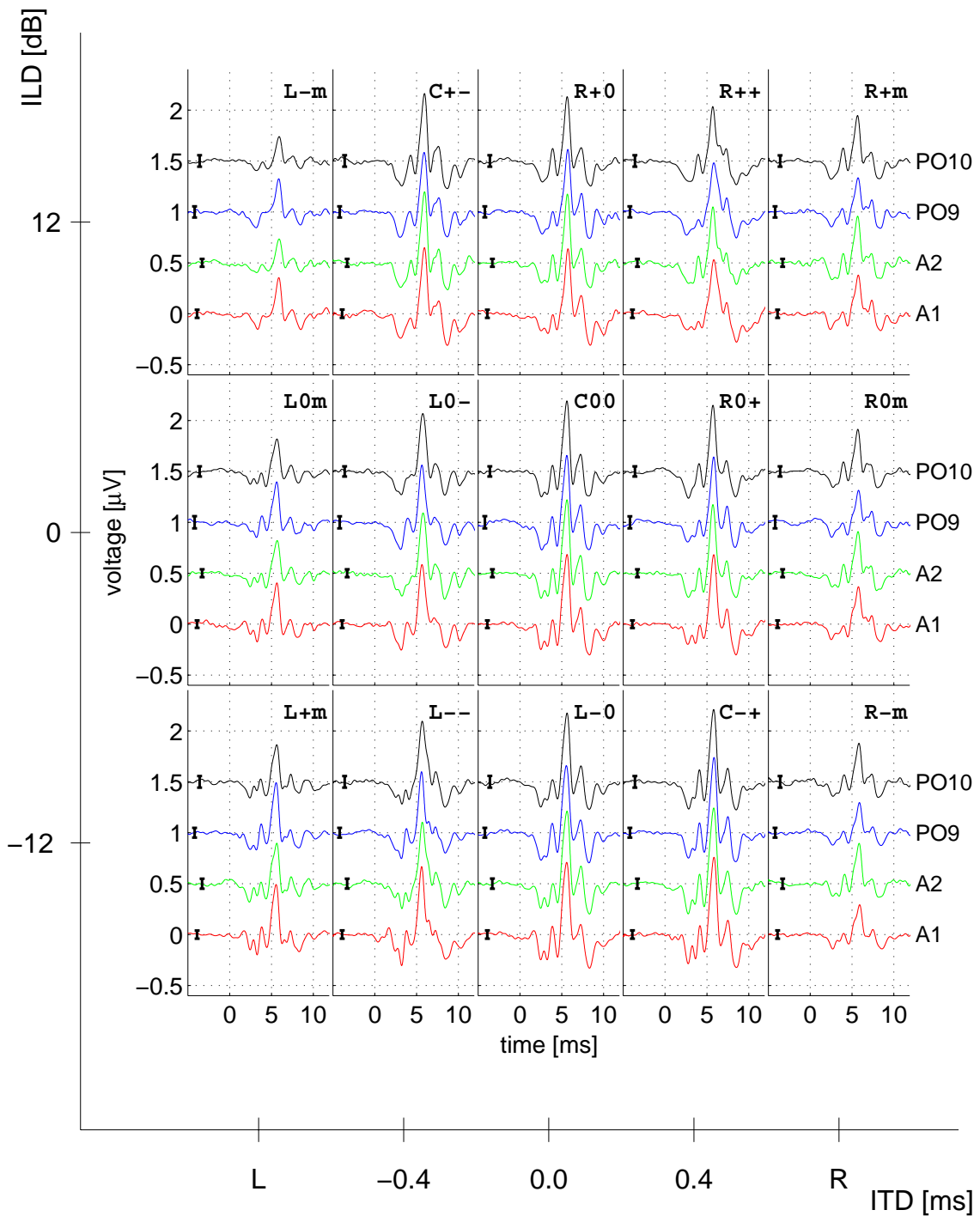
Fig. 3.2: ABRs for the 15 stimulus conditions, 9 binaural and 6 monaural, for one subject (dj). Stimuli are arranged as in Fig. 3.1. Plot offset between the channels is 0.5 μV. Error bars indicate ±3 standard errors of the mean (±3σ).
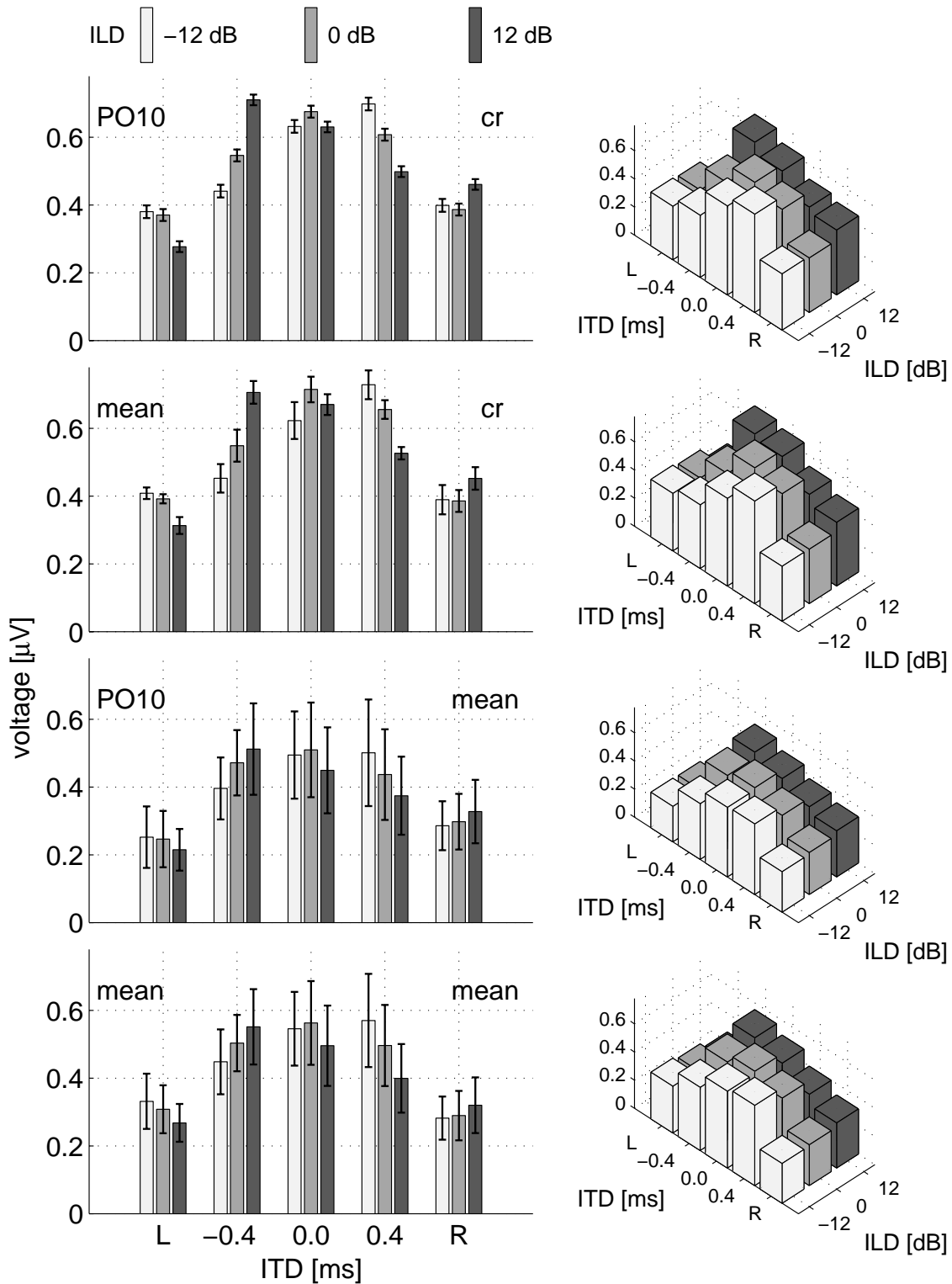
Fig. 3.3: *Amplitudes of wave V as a function of ILD and ITD.* **Left column:** *two-dimensional representation. Brightness of the bars codes the ILD: bright –12 dB, gray 0 dB and dark 12 dB.* **Right column:** *corresponding three-dimensional representation.* **Top row:** *Data for channel PO10 and subject cr, error bars indicate the intraindividual standard error σ.* **Second row:** *Data for mean over channels and subject cr.* **Third row:** *Data for channel PO10 and mean over subjects, error bars show interindividual standard deviations.* **Bottom row:** *Data for mean over channels and subjects.*
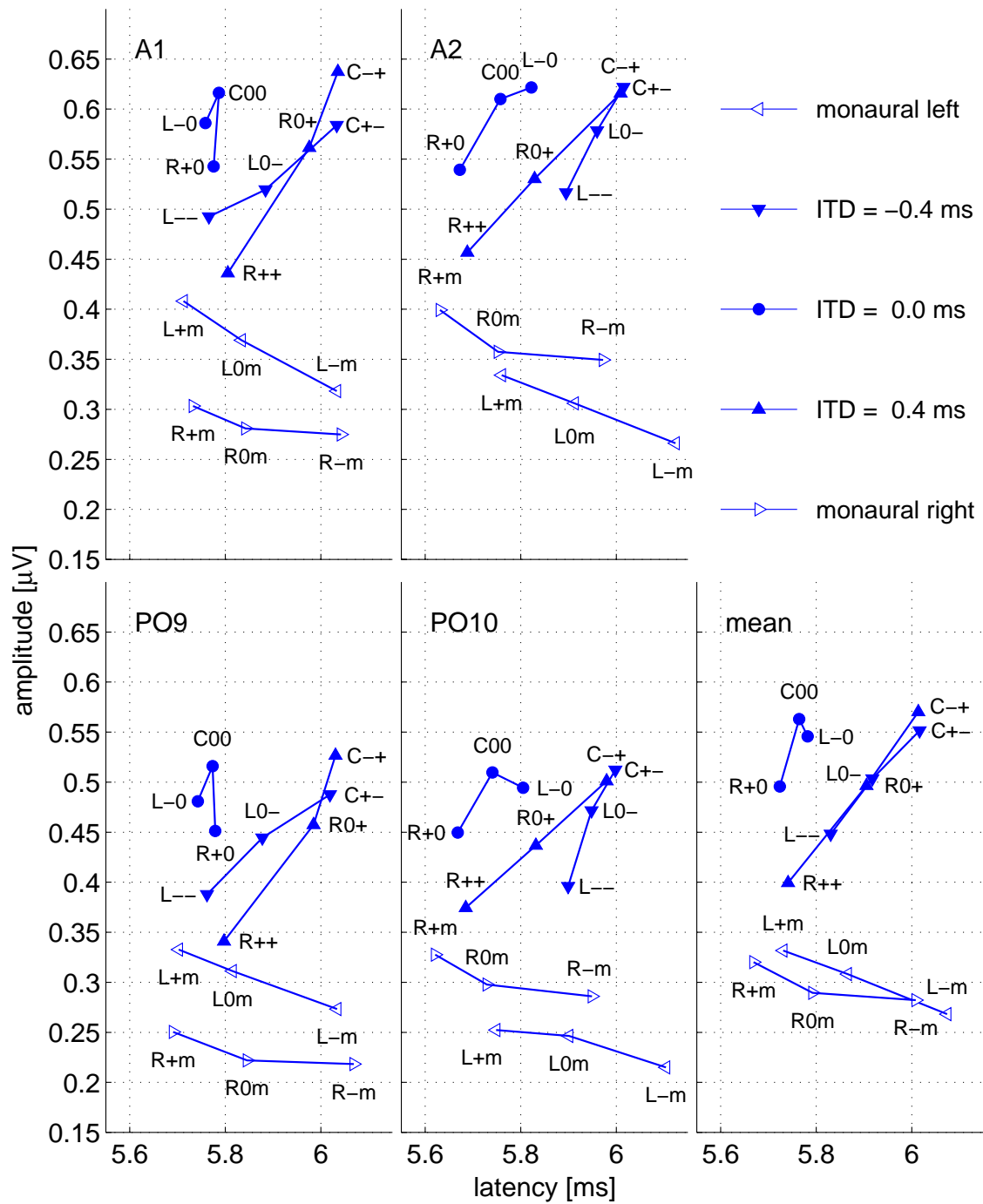
Fig. 3.4: Amplitude of wave V as a function of wave V latency, mean over subjects. The subplots are for different channels and mean over channels. Data belonging to the same ITD are connected with lines, stimulus names are according to Fig. 3.1.
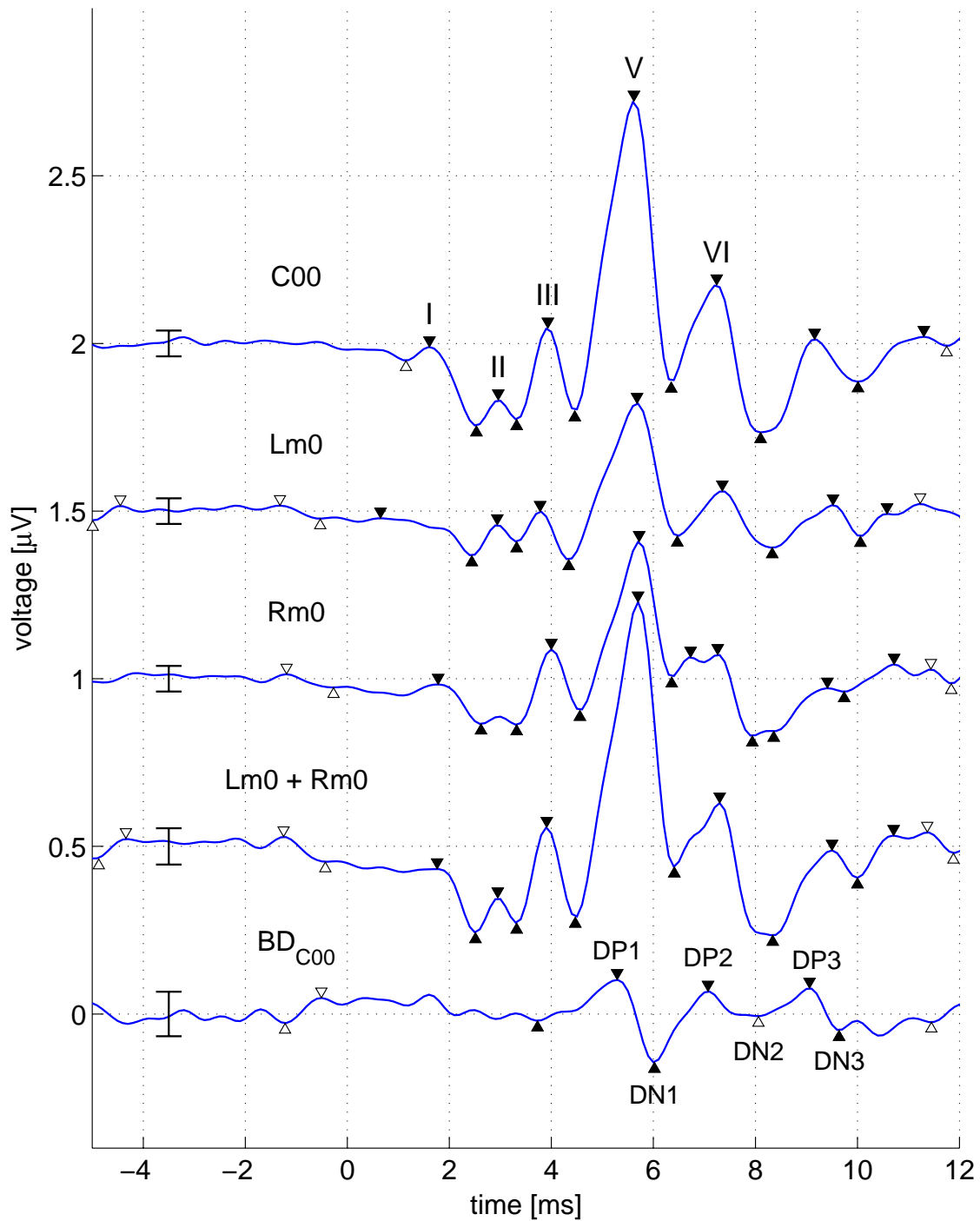
Fig. 3.5: *Derivation of the binaural difference potential for subject dj to diotic stimulation, channel A2.* **Top row:** *response to the binaural stimulus with Jewett peaks,* **second row:** *response to the monaural left stimulus,* **third row:** *response to the monaural right stimulus,* **fourth row:** *sum of the monaural responses,* **Bottom row:** *binaural difference potential with nomenclature of peaks. The voltage offset between adjacent curves is $0.5\,\mu V$. Error bars indicate ± 3 standard errors of the mean ($\pm 3\sigma$). Open triangles: $V_{pp} \geq \sqrt{2} \cdot 2\sigma$, filled triangles: $V_{pp} \geq \sqrt{2} \cdot 3\sigma$.*
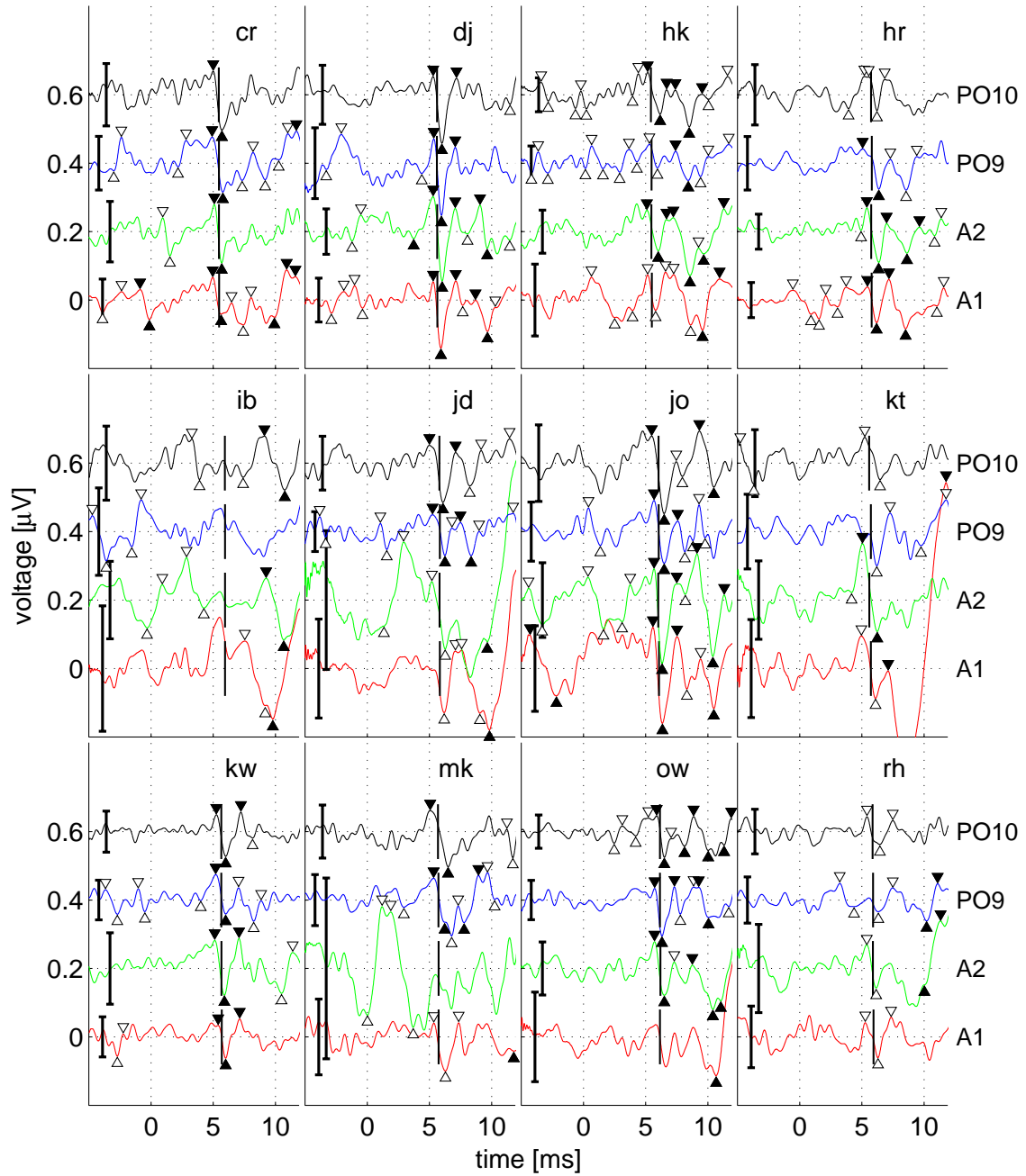
Fig. 3.6: BDs for 12 subjects and 4 channels. The voltage offset between adjacent curves is $0.2\,\mu V$. Error bars indicate ± 3 standard errors of the mean. Vertical lines show the latency of wave V in the corresponding binaural potential. Open triangles: $V_{pp} \geq \sqrt{2}\cdot 2\sigma$, filled triangles: $V_{pp} \geq \sqrt{2}\cdot 3\sigma$.
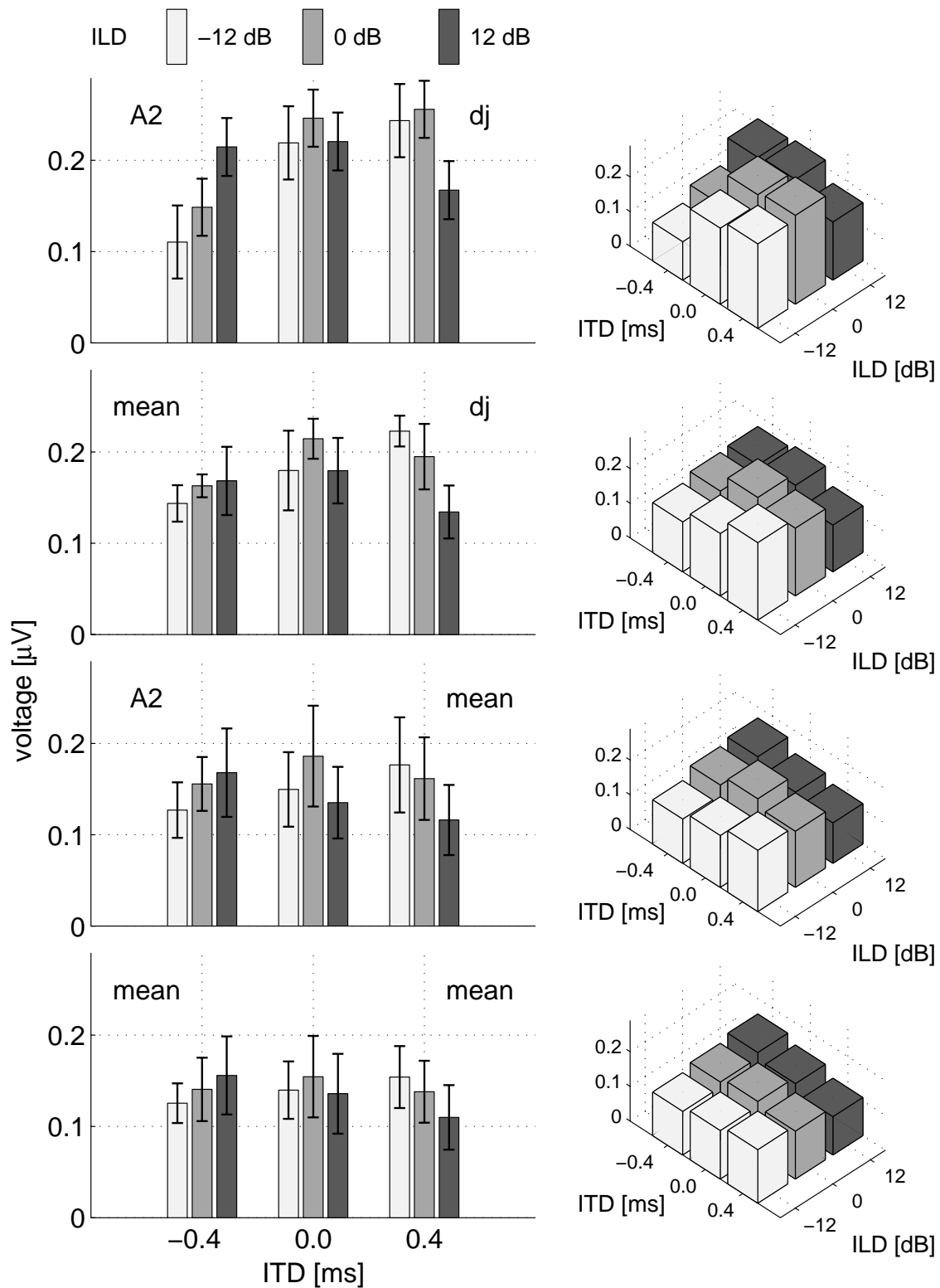
Fig. 3.7: Amplitudes of BD wave DP1-DN1 as a function of ILD and ITD. **Left column:** two-dimensional representation. Brightness of the bars codes the ILD: bright –12 dB, gray 0 dB and dark 12 dB. **Right column:** corresponding three-dimensional representation. **Top row:** Data for channel A2 and subject dj, error bars indicate the intraindividual standard error σ. **Second row:** Data for mean over channels and subject dj. **Third row:** Data for channel A2 and mean over subjects, error bars show interindividual standard deviations. **Bottom row:** Data for mean over channels and subjects.
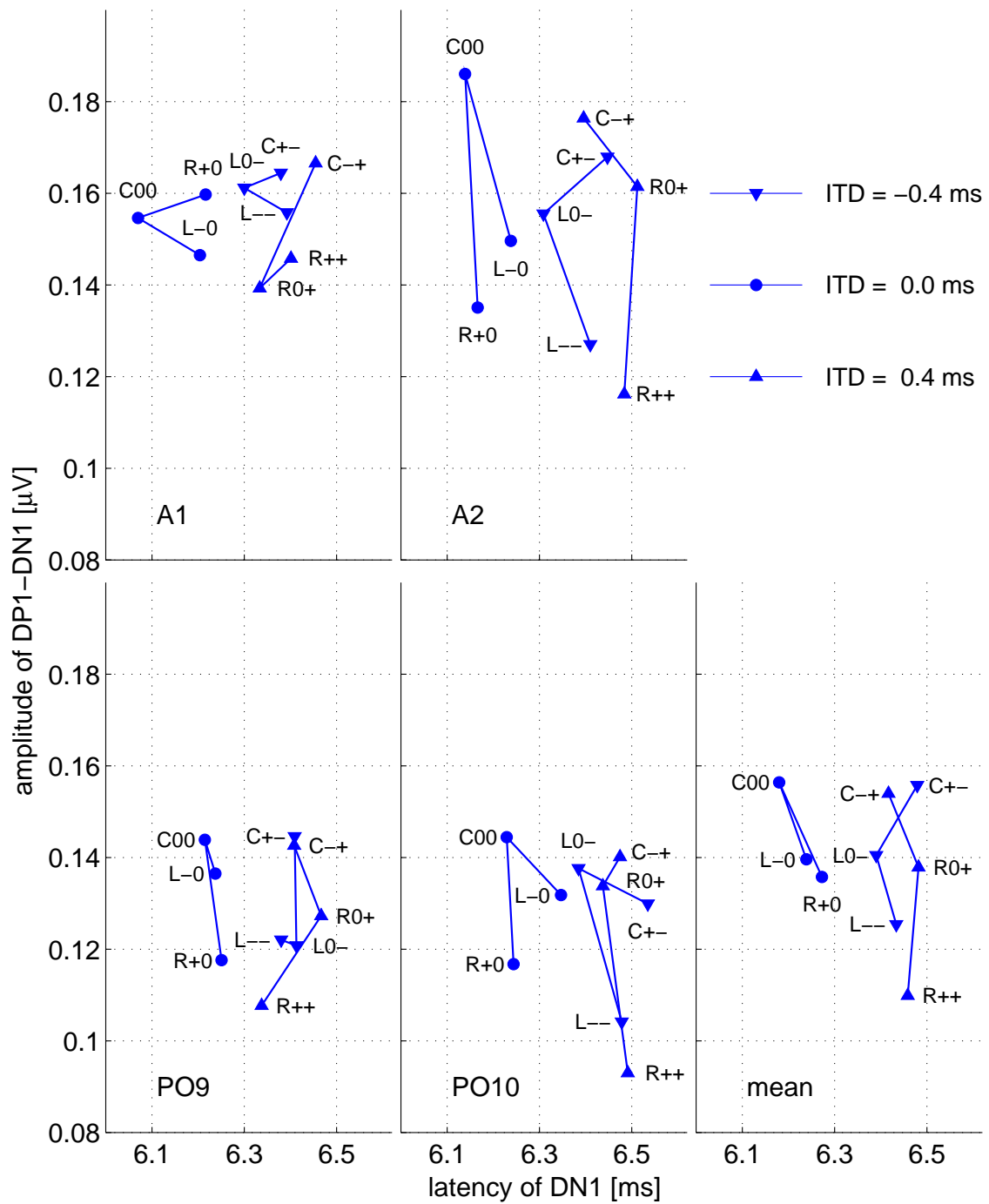
Fig. 3.8: BD amplitude DP1-DN1 as a function of DN1 latency, mean over subjects. The subplots are for different channels and mean over channels. Data belonging to the same ITD are connected with lines, stimulus names are according to Fig. 3.1.

3.4   Discussion

High quality recordings of ABRs and BDs for 9 combinations of ILD and ITD were per-
formed. Single epochs were recorded and analyzed offline allowing for an improved filtering
and averaging technique as well as for an estimation of the residual noise on single sweep
basis (Granzow et al., 2001; Riedel et al., 2001). An objective SNR criterion was applied to
assess the significance of the peaks. Artifact sources were ruled out by use of randomized
stimulation, moderate stimulation levels and the use of insert earphones (Levine, 1981).

A systematic relation between binaural wave V amplitude and stimulus lateralization was
clearly demonstrated for all subjects. ABRs exhibit the largest amplitude for centered
stimuli. With increasing lateralization the amplitude decays gradually. Wave V latency
does not correlate with stimulus lateralization. It is shortest for ITD = 0 and the syner-
gistic stimuli and longest for the trading stimuli. Intermediate values are seen for stimuli
with ITD or ILD only.

The BD component DP1-DN1 was significant in 11 out of 12 subjects in at least 3 of the 4
channels measured. BD amplitudes are maximal for diotic stimulation and decrease when
ILD or ITD are introduced. However, the amplitude difference between diotic responses
and responses to stimuli with ILD $= \pm 12\,$dB fail to reach significance. This finding is in
agreement with the study by Furst et al. (1985) who found only a marginal decrease of
the BD for ILD $= \pm 12\,$dB, but in contrast to the study by McPherson and Starr (1995).
They found a large difference in the BD for ILD $= 0\,$dB and ILD $= 12\,$dB, respectively,
since they did not observe any BD for ILD $> 8\,$dB.

Our data show a significant decrease of the BD at ITD $= \pm\,0.4\,$ms. This is in accordance
with the studies by Furst et al. (1985) and McPherson and Starr (1995). Other studies
did not observe any change of the BD amplitude for ITD $= \pm\,0.4\,$ms: Jones and van der
Poel (1990), Brantberg et al. (1999b) found a constant BD amplitude up to ITD $= 1\,$ms.
However, the number of sweeps recorded per stimulus condition used in those studies
ranged from 4000 – 6400, and in none of them a randomized stimulation paradigm was
used. With the higher number of sweeps (10.000) used in the present study, SNRs of the
BD between about 2 and 6, or equivalently, 7 and 15 dB, were obtained. Therefore it is
possible that the differences between the studies mentioned above result from too low SNR
or from interindividual variation.

The main issue of this study was to investigate the BD in lateralization conditions rather
than in pure ILD and ITD conditions. The BD amplitudes show a similar dependence on
the stimulus parameters as observed for ABR wave V amplitude: The BD amplitude is

smallest for the synergistic stimuli, largest for the diotic and trading stimuli. The significant differences between antagonistic and synergistic responses allow the conclusion that ILD and ITD are not processed independently in the brain stem: Let us assume that the functions $A_L(ILD)$ and $A_T(ITD)$ describe the dependencies of the BD amplitude on ILD and ITD, respectively. They should be symmetric functions, i.e., $A_L(ILD) = A_L(-ILD)$ and $A_T(ITD) = A_T(-ITD)$, since the sign of the interaural disparities should not lead to different BDs. Assuming independent processing of ILD and ITD, the BD amplitude as a function of ILD and ITD would separate into two factors and could be rewritten as $A(ILD, ITD) = A_L(ILD) \cdot A_T(ITD) = A_L(ILD) \cdot A_T(-ITD) = A(ILD, -ITD)$. Thus, synergistic and antagonistic stimuli should reveal the same responses. This clearly contradicts the experimental results. It therefore must be assumed that ILD and ITD are not processed independently in the brain stem. The experiments support the hypothesis that the lateralization angle γ is represented by the BD ($\gamma = 0$ for centrally perceived stimuli). A possible relation of the BD and γ could be: $A(ILD, ITD) = \gamma(TIR * ILD + ITD)$ with TIR being the time-intensity-trading-ratio, i.e., the ILD required to compensate for the lateralization of a given ITD.

Physiological recordings in animals showed that the superior olive (SO) is the first stage of binaural interaction (e.g., Yin and Chan, 1988, 1990; Irvine, 1992). The cells in the SO were classified by the type of input they receive from the cochlear nuclei (CN) and the medial nucleus of the trapezoid body (MNTB), (Goldberg and Brown, 1969). Excitatory-excitatory (EE)-cells in the medial SO receive excitatory input from both sides, whereas inhibitory-excitatory (IE)-cells in the lateral SO receive contralateral inhibitory input via the MNTB and ipsilateral excitatory input. EE-cells are thought to code ITDs for low frequencies performing a running cross-correlation (Jeffress, 1948) whereas IE-cells are believed to code ILDs for high frequencies, thus forming the physiological basis of the duplex theory of sound localization (Rayleigh, 1907). Since in BD studies the binaural response is always found to be smaller than the sum of the monaural responses, at first glance one would claim for binaural interaction of the IE-type. However, as pointed out by Gaumond and Psaltikidou (1991), the reduction of the binaural response could also emerge from the EE-type of interaction since EE-cells could be driven to saturation by monaural stimulation. Therefore, in EEG studies it is difficult to distinguish if the BD originates from inhibition or saturation (or both). However, there is some reasoning based on models as well as experimental data that could allow for the separation of inhibition and saturation effects.

It is generally believed that the ABR to clicks mainly reflects the response to high frequen-

cies. Therefore it can be presumed that the BD measured in response to clicks represents to a larger portion the activity of the IE-cells processing ILDs than of EE-cells processing ITDs.

Gaumond and Psaltikidou (1991) analyzed the capability of two rather simple models, one of the IE-type, the other of the EE-type, to explain the striking constancy of the amplitude ratio $A_{\mathrm{BD}}/A_{\mathrm{V}}$ independent from input level (Levine, 1981). Whereas the IE-model naturally explains the constant amplitude ratio within the EE-model $A_{\mathrm{BD}}/A_{\mathrm{V}}$ generally varies with the input level due to the compressive non-linearity. Therefore, the constant amplitude ratio $A_{\mathrm{BD}}/A_{\mathrm{V}}$ can be understood as a hint that the BD mainly represents binaural interaction of the IE-type.

Ungan et al. (1997) analyzed EE- and IE-models to explain the increase in latency of the BD component DN1 with increasing ITD. With a cross-correlation model of the EE-type using delay lines (Jeffress, 1948) the latency increase should be ITD/2, close to the value found in this study. However, Ungan et al. measured BDs in cat with very fine spacing of the ITD and found DN1 latency increases larger than ITD/2. Proposing a model of the IE-type, they could better explain the experimental data.

In evoked-potential studies, it was shown by means of a spatio–temporal dipole model (Scherg and von Cramon, 1985; Scherg, 1991) that the active structures at the latency of wave V are the SO and the lateral lemniscus (LL). This view is also supported by leason studies from Melcher and Kiang (1996). Therefore, it must be assumed that the analysis and coding of directional information first takes place in the SO and the LL as the neural generators of the peaks in the BD. Our results indicate that a combined evaluation of interaural time and intensity cues already takes place at these stations in the human auditory pathway.

The interaural parameters used in this study are quite extreme, even though in the physiological range. In future studies a finer resolution and smaller values for the ITD and the ILD are desirable. However, although significant BD peaks were shown, a further improvement in SNR seems to be highly desirable to distinguish more clearly between stimulus conditions. Of course, the number of recorded sweeps per stimulus condition could be increased in order to achieve a higher SNR. However, since measurement time grows with the square of the SNR, this is not a practical solution. Another possibility to achieve a further noise reduction is a source analysis via a dipole fit from multi-channel measurements. Data from this approach is presented in the next chapter.

Chapter 4

Dipole source analysis of auditory brain stem responses evoked by lateralized clicks[1]

Abstract

The objective of this chapter is to elucidate the relation between psychophysical lateralization and the neural generators of the corresponding auditory evoked potentials. Multichannel EEG-recordings of auditory brain stem responses from 12 subjects to binaural click stimuli with different interaural time- and level differences are analyzed. Data are modeled by equivalent current dipoles representing the generating sources in the brain. A generalized maximum-likelihood method is used to solve the inverse problem by including the noise covariance matrix of the data. Confidence regions for the estimated dipole parameters are derived. The fit quality is assessed by computing the goodness-of-fit as the outcome of a χ^2-test. This measure appears to be advantageous over the usually employed residual variance. At the latency of Jewett wave V, a systematic variation of the moment of a rotating dipole with the lateralization of the stimulus is found. Dipole trajectories to stimuli with similar lateralization are similar. A sign reversal of the interaural differences results in a mirrored trajectory. Centrally perceived stimuli correspond to dipoles with the largest vertical components. With increasing lateralization, the vertical component of the moment decreases while the horizontal components increase. An alternative, physiologi-

---

[1]A modified version of this chapter has been submitted for publication to Z. Med. Phys.: Riedel and Kollmeier (2002): "Dipole source analysis of auditory brain stem responses evoked by lateralized clicks"

cally motivated model using two constrained, hemispherically symmetric fixed dipoles is also capable of explaining the data. However, compared to the rotating dipole, this model exhibits a worse goodness-of-fit. Finally, a comparison of the forward models allows the conclusion that the differentiation between these models is virtually not restricted by poor spatial sampling, but limited by the residual noise of the measurement. The similarity between rotating dipole trajectories induced by the same lateralization show that interaural time- and level differences are not processed independently. The data support the notion that directional information is already extracted and represented at the level of the brain stem.

4.1   Introduction

Hearing with two ears allows to accurately localize sounds in space. The perceived spatial position of an acoustic stimulus predominantly depends on its interaural time difference (ITD) and its interaural level difference (ILD) (Rayleigh, 1907; Feddersen et al., 1957; Mills, 1958, 1960; Jeffress and McFadden, 1971; Durlach and Colburn, 1978; Colburn and Durlach, 1978; Yost, 1981; Yost and Gourevitch, 1987). In order not to blur the interaural information, the binaural system compares the inputs from both ears at a relatively early stage in the neural pathway. The first intersection of left and right auditory nerve fibers occurs at the superior olive (SO) within the brain stem (Nieuwenhuys et al., 1988). Binaurally sensitive cells also reside in the subsequent brain stem nuclei, i.e., the lateral lemniscus (NLL) and the inferior colliculus (IC) (Goldberg and Brown, 1968, 1969; Caird and Klinke, 1983; Caird et al., 1985; Yin and Chan, 1988, 1990; Irvine, 1992; Melcher and Kiang, 1996; Gummer and Zenner, 1996; van Adel et al., 1999).

Evoked potentials have been widely used to investigate directional hearing in humans. The majority of the studies dealing with the dependence of auditory brain stem responses (ABRs) on ITD and ILD focused on the analysis of the waveforms of single EEG channels (Levine, 1981; Wrege and Starr, 1981; Gerull and Mrowinski, 1984; Furst et al., 1985; Ito et al., 1988; Jones and van der Poel, 1990; McPherson and Starr, 1995; Jiang, 1996; Cone-Wesson et al., 1997; Brantberg et al., 1999a,b; Riedel and Kollmeier, 2002a). The so-called three-channel Lissajous' trajecory represents an attempt to infer the activity of the brain stem from ABR recordings. It is measured using three bipolar pairs of electrodes with preferably perpendicular orientation. The three time signals are then visualized as a trajectory in the three-dimensional voltage space (Pratt et al., 1983; Jewett, 1987; Pratt et al., 1990). Three-channel Lissajous' trajecories have been used to analyze binaural

processing (Polyakov and Pratt, 1994, 1996; Pratt et al., 1997; Polyakov and Pratt, 1998; Pratt et al., 1998). The benefit of the method is that it avoids the problem of the reference electrode by using a bipolar montage. However, it relies on the assumption of a single dipole located exactly in the center of the head, and volume conduction effects are not taken into account.

A more sophisticated approach to localize active neural tissue in the brain is dipole source analysis from multi-channel EEG measurements (Kavanagh et al., 1978; Scherg, 1984, 1990, 1991; Mosher et al., 1992). Scherg and von Cramon (1985) proposed a model of six fixed current dipoles to describe the five waves of the monaurally evoked ABR. This model not only used a fixed location and orientation for each dipole, but also heavily constrained the time course of the dipole moment magnitude. The active structures at the latency of the largest deflection, wave V, were identified as the superior olive and the lateral lemniscus.

The present study focuses on the source analysis of both the binaurally evoked ABR and the binaural difference potential (BD), i.e., the difference between binaural and summed monaural responses. In order to analyze the influence of lateralization, nine different binaural stimulus conditions, the combinations of three ITDs and three ILDs are used. Since the same lateralization can be generated by different combinations of the interaural differences, the variation of both cues allows to draw conclusions about the representation of the stimulus laterality in the brain stem.

4.2 Methods I – Data acquisition

4.2.1 Subjects

Twelve subjects (three females) from the staff of the University of Oldenburg without any history of audiological or neurological problems participated voluntarily in this study. They were aged between 25 and 36 years and were classified as normal hearing by routine audiometry. The audiometric loss was less than 10 dB for frequencies below 4 kHz and less than 15 dB for the higher frequencies.

4.2.2 Stimuli

Rarefaction click stimuli were produced by applying rectangular voltage pulses of $100\,\mu$s duration to Etymotic Research ER-2 insert earphones. The time interval between the

onsets of two successive stimuli was chosen to vary randomly and equally distributed between 62 and 72 ms, yielding an average stimulation rate of approximately 15 Hz.

15 stimulus conditions were tested, 9 binaural and 6 monaural. The monaural clicks were presented at the levels 53, 59 and 65 dB nHL and are denoted as L–m, L0m, L+m and R–m, R0m, R+m for monaural left and right stimulation, respectively. The binaural stimuli were the nine possible combinations of 3 ITDs (–0.4, 0 and 0.4 ms) and 3 ILDs (–12, 0 and 12 dB), as indicated in Fig. 3.1.

The binaural stimuli are named as follows: the first letter refers to the perceived lateralization of the stimuli: 'L' for left, 'C' for center and 'R' for right. The second and third characters ('–', '0' and '+') are used to specify the ILD and ITD, respectively. For example, for diotic stimulation, to both ears a click at 59 dB nHL was presented simultaneously. This stimulus C00 is found in the center of the diagram (ILD=ITD=0). The stimulus R+0 in the middle of the top row has zero ITD, but is lateralized to the right due to its positive ILD. On the other hand, the stimulus R0+ at the right of the middle row has zero ILD, but is also lateralized to the right due to its positive ITD. The arrows in Fig. 3.1 point into the approximate direction of the lateralization of the stimuli. Both an ITD of 0.4 ms and an ILD of 12 dB cause a strong, but not extreme lateralization of about 70° (Furst et al., 1985; McPherson and Starr, 1995). A stronger, almost complete lateralization is produced by the synergistic stimuli L–– and R++ whose ILDs and ITDs point into the same direction. In contrast, the stimuli C+– and C–+ refer to the antagonistic situation: ILD and ITD act in opposite direction resulting in a centered image. In the lower left corner of each subplot in Fig. 3.1, the respective binaural stimulus is depicted schematically. In the ITD-ILD-plane lines of equal lateralizaton are the diagonal dotted lines. Identical stimuli were used for all subjects for better comparability of the results. Comparatively large values for the ITD and the ILD were used to obtain as large differences in the evoked potentials for the different stimulus conditions as possible without leaving the physiological range.

4.2.3   Electrode configuration

Ag/AgCl-electrodes were attached to a flexible cap worn by the subjects. ABRs were recorded from 32 sites according to the extended 10-20-system (Jasper, 1957; Sharbrough et al., 1991). The common reference electrode was placed at the vertex (Cz), the ground electrode at the forehead (Fpz). Electrode impedances were measured at a test signal frequency of 30 Hz and brought below 5 kΩ. Since DC recordings were performed, the criteria for a good contact between electrodes and skin were both a low impedance and a vanishing

voltage drift seen in the raw EEG signal. The electrode positions in the three-dimensional space were determined using a measuring instrument exploiting different ultrasonic signal propagation delays at different sensors (CMS30P by Zebris Medizintechnik). Electrode positions were measured before and after each recording session.

4.2.4 Recordings

During the ABR recordings, subjects lay in a sound insulated and electrically shielded room. They were instructed to relax and lie as comfortably as possible. ABRs were recorded with a DC-coupled differential amplifier (Synamps 5803). Inside the shielded room the EEG was preamplified by a factor 150, further amplified by the main amplifier by a factor 33 resulting in a total amplification of 74 dB. The voltage resolution was approximately 16.8 nV/bit. The sweeps were filtered by an analog antialiasing-lowpass with a corner frequency of 2 kHz, digitized with 10 kHz samplingrate and 16 bit resolution, and stored to hard disk. During the recording the artifact rejection level was set to ± 500 μV, since filtering, artifact analysis and averaging was done offline. The clipping level of the DA-converters was $\pm 550 \mu$V. The recording interval comprised 400 samples in the time interval from -15 to 25 ms relative to stimulus onset. For the dichotic stimuli the leading click was defined as stimulus onset. 10.000 single sweeps for all of the 15 stimuli were recorded in the following manner: The 5 stimuli belonging to a certain ILD (a row in Fig. 3.1, e.g., L–m, C+–, R+0, R++ and R+m) were presented in random order on a sweep-by-sweep basis. One run consisting of 12500 stimuli, 2500 of each type, lasted approximately 14 minutes. After each run the impedances were checked, and adjusted if necessary. Afterwards the runs belonging to the two other ILDs were recorded. Totally, 12 runs (4 repetitions of the 3 runs) were carried out for every subject. The duration of the recording session was about 3 hours excluding preparation and pauses.

No contralateral masking was used during the monaural presentation. This has been used in other studies to avoid acoustic crosstalk (ACT) (Jones and van der Poel, 1990; McPherson and Starr, 1995; Brantberg et al., 1999a). Instead, following Levine (1981) and Ito et al. (1988), left, right and binaural stimuli were presented randomly without contralateral masking. In combination with the use of insert earphones and moderate presentation levels of $\leq$65 dB nHL, this avoids contributions of ACT and the middle ear reflex to the binaural difference potential. Before averaging, the single sweeps were filtered with a linear phase FIR bandpass with 200 taps and corner of frequencies 100 and 1500 Hz (Granzow et al., 2001). An iterated weighted average of the filtered sweeps was computed

for all subjects and stimulus conditions, as described in chap. 2.

4.2.5   Binaural interaction

The binaural interaction was computed in terms of the binaural difference potential BD = B − (L + R). This was done channel-wise and sample by sample. The six monaural stimuli were chosen to allow for the computation of the BD for all 9 binaural stimulus conditions. For the stimuli with an ITD, the monaural response of the lagging ear was digitally delayed by the ITD before computing the BD. For the stimuli with an ILD, the monaural stimuli with the corresponding level were used, e.g., $BD_{C+−} = C+− − (L–m + R'+m)$ with R'+m being the delayed version of R+m. All 9 BDs were computed from stimulus triplets which had been recorded quasi-simultaneously, i.e., during the same measurement run. This avoids artifacts in the BD components due to long term changes of the recording conditions or subject's state. The residual noise of the BD was estimated as the square root of the summed variances of the three measurements, e.g., $\sigma_{BD_{C00}} = (\sigma^2_{C00} + \sigma^2_{L0m} + \sigma^2_{R0m})^{1/2}$, assuming that C00, L0m and R0m are statistically independent.

4.3   Methods II – Dipole source analysis

Modelling of multi-channel EEG data is generally done by minimizing a cost function that describes the difference between measured and modeled EEG. Two ingredients are required to calculate a model EEG: a source model and a head model, together they constitute the so-called forward model. Information about the noise in the measured data is needed in three steps of the modeling procedure: first, it enters the cost function, second, it is needed to estimate the uncertainties of the fit parameters and third, it is required to assess the goodness-of-fit. The programs performing the source analysis were written in MATLAB. In the following the elements of dipole source analysis are explained in more detail.

4.3.1   The forward model

Equivalent current dipoles (ECDs) are used as source model. ECDs are focal sources representing the center of a small brain region in which many cells are synchronously activated (Scherg, 1990, 1991). A homogeneous sphere characterized by a constant volume conductivity $g = 0.0033\ \Omega^{-1}cm^{-1}$ (Cuffin and Cohen, 1979) was chosen as head model. The forward model is mathematically described by the so-called leadfield matrix **F** which only depends on parameters of the head model, here g, the positions $\vec{r}_d$ of the D dipoles

and the positions $\vec{r}_c'$ of the C channels, i.e., electrodes. $\mathbf{F}$ does not depend on the dipole moments $\vec{m}_d$. The linearity of the EEG in the moment parameters allows to express the modeled potentials, i.e., the forward model $\vec{V}_{\mathrm{f}}$ (a Cx1 vector), as the scalar product of the leadfield matrix and the dipole moment vector $\vec{m}$. In the case of D dipoles, the moment vectors $\vec{m}_d$ have to be concatenated to the $3D$x1 vector $\vec{m}$.

$$\vec{V}_{\mathrm{f}} = \mathbf{F}(g, \vec{r}_1', \ldots, \vec{r}_C', \vec{r}_1, \ldots, \vec{r}_D)\,\vec{m}\,. \tag{4.1}$$

$\mathbf{F}$ describes the potential at electrode positions $\vec{r}_c'$ evoked by D current dipoles at positions $\vec{r}_d$ with unit dipole moment. $\mathbf{F}$ is a CxL matrix with C being the number of channels and L the number of linear parameters, e.g., $L = 3D$ for D rotating dipoles. The leadfield matrix of the homogeneous sphere model is given by analytic formulae which have the advantages of stable and relatively quick computation (Brody et al., 1973).

In the current work three source models are compared in their capability to explain the data:

1. **A moving dipole:** for each time sample it has $P = 6$ fit parameters: $N = 3$ nonlinear parameters for the dipole location $\vec{r}$, and $L = 3$ linear parameters for the dipole moment $\vec{m}$.

2. **A rotating dipole:** The location is fitted, but constrained to be constant for the time interval under consideration. For each time sample the three moment parameters are adjusted. The total number of parameters of a rotating dipole is $P = 3 + 3T$ with $N = 3$, $L = 3T$ and T the number of samples in the time interval.

3. **A pair of constrained fixed dipoles:** The location and the orientation of a fixed dipole is fitted, but constrained to be constant for the time interval under consideration. For each time sample only the moment magnitude is fitted. The pair of fixed dipoles is constrained to reflect hemispheric symmetry: the x-coordinates (left-right) and the azimuths are mirrored, i.e., have opposite signs. The y- and z-coordinates and the elevation must be identical for both dipoles. The model has $N = 5$ and $L = 2T$ parameters, i.e., totally $P = 5 + 2T$ parameters. Additionally, stronger constraints are employed: the dipole location is assumed to be known from a-priori physiological information, and the dipole orientation is required to be the same for all stimulus conditions. This model with stronger constraints has $N = 2$ and $L = 2T$ parameters, i.e., totally $P = 2 + 2T$ parameters.

4.3.2   The noise model

Measured data is supposed to be contaminated with additive multivariate Gaussian noise. It has the density function

$$f(\vec{n}) = \frac{1}{(2\pi)^{C/2} \, |\mathbf{S}|^{1/2}} \, e^{-\frac{1}{2}\vec{n}^T \mathbf{S}^{-1} \vec{n}} \tag{4.2}$$

with $|\mathbf{S}|$ being the determinant of the CxC noise covariance matrix. The difference between data and forward model, i.e., the residual $\vec{e} = \vec{V} - \vec{V}_{\mathrm{f}}$ is assumed to be a sample of the Cx1 noise vector $\vec{n}$ (John et al., 1987). The elements of the time-averaged noise covariance matrix $\mathbf{S}$ are computed from the J single sweeps of a measurement according to

$$\sigma_{cd} = \frac{1}{T}\sum_{t=1}^{T} \frac{1}{J(J-1)} \sum_{j=1}^{J} \Big(V_{cj}(t) - \mu_c(t)\Big)\Big(V_{dj}(t) - \mu_d(t)\Big) \tag{4.3}$$

with c and d being the indices of two channels. $\mu_c(t) = 1/J \sum_{j=1}^{J} V_{cj}(t)$ represents the ensemble average, i.e., the evoked potential of channel c at time t. The diagonal elements of $\mathbf{S}$ are the variances, i.e., the squared standard errors of the measurement. Similar to the noise variances, the elements of the noise covariance matrix do not vary much over time, see Fig. 2.2. Therefore, the time dependence of $\mathbf{S}$ was eliminated by averaging over time. An analogous formula holds for the weighted noise covariance matrix.

4.3.3   The cost function

The cost function is a functional that maps the residual $\vec{e} = \vec{V} - \vec{V}_{\mathrm{f}}$, weighted by the elements of the noise covariance matrix σ_{cd}, onto a scalar error value E. The assumption of multivariate Gaussian noise leads to the cost function of generalized maximum-likelihood estimation E_S that incorporates the complete noise covariance matrix $\mathbf{S}$

$$E_S = \vec{e}^T \, \mathbf{S}^{-1} \, \vec{e} = \sum_{c=1}^{C}\sum_{d=1}^{C} e_c \, [\mathbf{S}^{-1}]_{cd} \, e_d \,. \tag{4.4}$$

If only the noise variances are known, i.e., $\sigma_{cd} = \sigma_c^2 \, \delta_{cd}$, the cost function reduces to the one of ordinary maximum likelihood estimation.

$$E_V = \sum_{c=1}^{C} (e_c/\sigma_c)^2 \,. \tag{4.5}$$

If no information of the measurement errors is available, i.e., $\mathbf{S} = \mathbf{1}$, the error functional simplifies to the least-squares cost function

$$E_1 = \vec{e}^T \vec{e} = \sum_{c=1}^{C} e_c^2 . \qquad (4.6)$$

4.3.4 The inverse problem

The estimation of source parameters from the voltage distribution measured at the head surface is called the 'inverse' problem. In the present work, the inverse problem is tackled by minimizing the cost function E_S by means of the simplex algorithm by Nelder and Mead (1965). The search of the minimum is performed in the space spanned by the N nonlinear parameters, i.e., the dipole locations (and the dipole orientations in the case of fixed dipoles). Within the nonlinear fitting routine, the Lx1 vector of the optimal moment parameters is directly determined by

$$\vec{m} = (\mathbf{F}^T \mathbf{S}^{-1} \mathbf{F})^{-1} \mathbf{F}^T \mathbf{S}^{-1} \vec{V} . \qquad (4.7)$$

The resulting forward model that minimizes the cost function is termed 'the fit' $\vec{V}_{\text{fit}}$. Since the time-averaged noise covariance matrix is considered, its inverse $\mathbf{S}^{-1}$ entering the cost function, has to be calculated only once before the minimization procedure.

Although it is common practice to transform data into average reference, due to the inclusion of the complete noise covariance matrix it is necessary to treat the inverse problem in recording reference. The transformation of $\mathbf{S}$ into average reference leads to a matrix $\mathbf{S}'$ which does not have full rank, i.e., is not invertible. The transformation formulae and the proof of the singularity of $\mathbf{S}'$ are given in appendix B.

4.3.5 Estimation of the parameter uncertainties

The LxL covariance matrix $\mathbf{S}_L$ of the linear moment parameters is determined by error propagation

$$\mathbf{S}_L = (\mathbf{F}^T \mathbf{S}^{-1} \mathbf{F})^{-1} . \qquad (4.8)$$

The variances of the linear parameters are found on the diagonal of $\mathbf{S}_L$.

The NxN covariance matrix $\mathbf{S}_N$ of the nonlinear location parameters are estimated by calculating the curvatures in a small region around the minimum.

$$\mathbf{S}_N = \left(\vec{\nabla}_{\vec{r}} \vec{V}_{\text{fit}}^T \; \mathbf{S}^{-1} \; \vec{\nabla}_{\vec{r}} \vec{V}_{\text{fit}} \right)^{-1} . \tag{4.9}$$

$\vec{\nabla}_{\vec{r}} \vec{V}_{\text{fit}}$ is the CxN matrix of the partial derivatives of the forward model with respect to the nonlinear parameters.

4.3.6   Evaluation of the goodness-of-fit

The fit quality is assessed by the comparison of the residual, i.e., the difference between data and forward model, with the noise in the measurement. If the residual is below or in the range of the measurement errors, the fit can be regarded as of good quality. On the other hand, if the deviation between model and data drastically exceeds the noise, the fit cannot be considered as satisfactory. In mathematical terms, the fit quality is assessed by means of a chi-square test.

Under the assumption of Gaussian measurement errors the quantity e_c/σ_c is normally distributed with mean zero and variance one. Consequently, the cost function E_V (eq. 4.5) is a sum of squared standard normal distributions and is therefore χ^2-distributed.

In the present work, the complete noise covariance matrix $\mathbf{S}$ resulting in the more general cost function E_S (eq. 4.4) is used. However, E_S remains invariant under a unitary transform $\mathbf{U}$, i.e., $\mathbf{U}^{-1} = \mathbf{U}^T$, into the eigen-system of $\mathbf{S}^{-1}$. With $\widehat{\mathbf{S}^{-1}} = \mathbf{U}\mathbf{S}^{-1}\mathbf{U}^T$ and $\widehat{\vec{e}} = \mathbf{U}\vec{e}$ it follows

$$\widehat{E_S} = \widehat{\vec{e}^T} \, \widehat{\mathbf{S}^{-1}} \, \widehat{\vec{e}} = \vec{e}^T \, \mathbf{U}^T \, \mathbf{U} \mathbf{S}^{-1} \, \mathbf{U}^T \, \mathbf{U} \vec{e} = \vec{e}^T \, \mathbf{S}^{-1} \, \vec{e} = E_S . \tag{4.10}$$

Since $\widehat{\mathbf{S}^{-1}}$ is diagonal, $\widehat{E_S}$ has the form of E_V (eq. 4.5). As a consequence, E_S also obeys a χ^2-distribution. E_S^*, denoting E_S at the end of the fitting procedure, is distributed as a χ^2-distribution with $\nu = C - P$ degrees of freedom. The goodness-of-fit (gof) was calculated using the incomplete gamma function Γ_{inc}:

$$\text{gof} = 1 - \Gamma_{\text{inc}}\left(\frac{E_S^*}{2}, \frac{\nu}{2}\right) = 1 - \frac{\displaystyle\int_0^{E_S^*/2} x^{\frac{\nu}{2}-1} e^{-x} \, dx}{\displaystyle\int_0^{\infty} x^{\frac{\nu}{2}-1} e^{-x} \, dx} . \tag{4.11}$$

For vanishing error E_S gof approaches 1. For increasing E_S the numerator approaches the denominator resulting in a decreasing gof. The χ^2-distribution has a mean of ν and a variance of 2ν. Therefore, gof-values above 0.1 are considered as good, above 0.001 as

acceptable, models with gof < 0.001 should be rejected (Press et al., 1992). The goodness-of-fit is determined by the ratio of E_S and ν. For example, at a single time instant the moving dipole has $\nu = 26$ degrees of freedom. If for each channel the deviation between model and data equals the standard error of that channel, E_S amounts to 32 and gof to 0.193 which is rated as a good fit.

For the purpose of comparison, the so-called residual variance, which is widely used in the literature, was also computed:

$$\text{rv} = \frac{\sum_{c=1}^{C}(V_{\text{fit},c} - V_c)^2}{\sum_{c=1}^{C} V_c^2} = \frac{E_1}{P_{\text{global}}} \tag{4.12}$$

The residual noise rv is the least-squares cost function normalized by the global power of the measured signal P_{global}.

4.4 Results

4.4.1 Sample data

Fig. 4.1 shows ABRs of subject jo to the diotic stimulus (ITD = ILD = 0) for all 32 measured channels in recording reference. For each channel three responses are depicted: the response to the diotic stimulus (B) in the top trace, the sum of the monaural responses (L+R) in the middle trace, and the binaural difference potential (BD) in the bottom trace.

The residual noise for each channel and condition is shown by the error bars. They denote the threefold standard error of the mean (± 3 S.E.M.). Channels near the reference electrode (CZ) generally exhibit smaller responses and smaller residual noise compared to channels more distant from CZ. Exceptions are channels FP2 and F4 which show irregular waveforms. However, these channels also have large standard errors. While a least squares fit would incorporate these irregular waveforms into the dipole fit, the maximum likelihood method is not distracted by these irregularities because it takes into account the raised residual noise in these channels. Binaural difference potentials (BDs) are most clearly observed for channels that exhibit high signal energy in the binaural stimulus condition. The main BD component is a trough in the downslope of wave V, it is labelled DN1. The preceding positive peak, named DP1, is usually of smaller amplitude.

In Fig. 4.2 the electrode positions measured before (filled circles) and after (open circles) the recording session for one subject (rh) are shown in three projections onto planes. The electrode labels are according to the extended 10-20-system. The radii of the fitted spheres are 8.62 and 8.61 cm, respectively. The standard deviations of the electrode positions from the spheres amount to 0.70 and 0.74 cm, respectively. The mean rms-deviation between the two data sets is 0.23 cm. Since the differences between measured positions and positions projected onto a sphere are small and because the spatial gradient of the evoked potentials are small due to the deep brain stem sources, cf. Fig. 4.3, the sphere model can be considered as a sufficient approximation here.

Fig. 4.3 shows data in average reference from Fig. 4.1 at a latency of 5.8 ms together with the electrode positions. The lines connect head surface sites with identical potential and are estimated using refined spline interpolation. The upper left panel shows the head model from the top, the upper right panel from the front, the lower left panel from the left side, and the lower right panel from the right side, respectively. Positive voltages are seen at superior electrodes, negative voltages at inferior electrodes. The dipole resulting from a fit only considering data at the current latency shows a predominant upward current flow. Its direction is perpendicular to the equipotential lines.

4.4.2   The moving dipole

In Fig. 4.4 the time course of the location of a moving dipole fitted to the diotic response of subject cr is depicted. The four curves show the x-,y-,z-component and the magnitude of the radius vector in the time interval from 0 to 10 ms, respectively. For all components, the dashed lines mark the 95%-confidence regions. The fitting routine did not allow for radii larger than the head radius of 8.8 cm defined by the fitted sphere of the electrodes. The discontinuities at latencies smaller than 2 ms and around 4.7 ms and 7.3 ms are a consequence of vanishing signal energy in these time intervals, (c.f. the upper left panel of Fig. 4.6) and of fitting each time sample independently. In the time interval of interest here, i.e., around the latency of wave V ($t_V = 5.5$ ms) the dipole is approximately located in the middle of the head (x=y=0), but about 3 cm below the horizontal plane, corresponding to a generator site in the brain stem.

Fig. 4.5 shows the time course of the moment of a moving dipole fitted to the diotic response of subject cr. The fitting routine limited the magnitudes of the dipole moments to 20 nAm. At the same latencies as in Fig. 4.4 discontinuities of the moment due to vanishing signal energy are observed (cf. Fig. 4.6). Around the wave V the vertical

component (m_z) is strongest, m_y directing to the front is about a fifth of m_z while m_x virtually vanishes. The magnitude at the peak of wave V is nearly 8 nAm. The dashed lines denote the 95%-confidence regions for the dipole moment. The result of a mainly vertical current direction is in accordance with Fig. 4.3.

Fig. 4.6 shows the characteristics of signal and noise, and exemplifies the two measures of fit quality applied. As in the previous figure, data are ABRs from diotic stimulation of subject cr. In the upper left panel, the rms-value of the signal averaged over the 32 measurement channels, i.e., $\sqrt{P_{\text{global}}}$ is shown. Sharp minima are seen at a latencies of 4.7 and 7.3 ms, respectively. At these latencies the moving dipole fit yields erroneous results, see Figs. 4.4 and 4.5. Wave V is the most pronounced peak of the ABR, its rms-value of 339 nV is more than twice as large as the rms-values of all other peaks.

The straight line in the upper right panel of Fig. 4.6 at 10.4 nV is the rms-value of the noise level averaged over channels, i.e., the square root of the global noise power. Since the noise fluctuations over time are small they were averaged over time, see Fig. 2.2. The noise level varies very little between stimulus conditions, but considerably between subjects. Rms-values are in the range from 8.9 to 16.5 nV, mean 11.8 nV. The residual $\vec{e}$ is the difference between measured EEG $\vec{V}$ and forward model $\vec{V}_{\text{fit}}$. Its rms-value averaged over channels is also plotted in the upper right panel of Fig. 4.6. It varies substantially over time, at t_V it is maximal and amounts to 24.6 nV. Whereas for illustration only the average over channels is depicted in the upper panels of Fig. 4.6, the fitting routine uses noise and residual individually for each channel, c.f. eq. 4.4.

In the lower left panel of Fig. 4.6, the residual variance rv (eq. 4.12) is shown. It is mainly determined by its denominator: At small signal energies, rv is found to be large and vice versa. As a consequence, despite the maximal residual at t_V of 24.6 nV, rv amounts to only 0.0053 or half a percent.

On the other hand, the bottom right panel of Fig. 4.6 depicts the goodness-of-fit (gof) as the outcome of the χ^2-test. At time instants where the residual is smaller than the noise, gof is near 1. At t_V gof is very low since the residual is more than twice the noise, i.e., it is very unlikely that the deviation between data and model is due to the noise. Although at t_V 99.5% of the variance of the data is explained by the model, the fit cannot be regarded as good because of the low gof.

Fig. 4.7 shows the quality measures of the moving dipole fit averaged over subjects. Around the latency of wave V, the mean residual variance (thin lines) lies between 1% and 10%. The goodness-of-fit values (thick lines) are high for the monaural stimuli and markedly

lower for the binaural stimuli implying that the monaurally evoked potentials are well described by a single dipole while the binaurally evoked potentials are not. This finding cannot be inferred from the residual variance.

4.4.3   The rotating dipole

In Fig. 4.8 the locations of a rotating dipole for all stimulus conditions and subject dj are shown. Data were fitted for a time interval from 1 ms before to 1 ms after the latency of wave V. This 2-ms interval comprises 21 samples and will be used in the subsequent analysis. Note that its beginning and end depend on stimulus condition and subject. As in the case of the moving dipole, see Fig. 4.4, the optimized locations are inferior to the center of the head, about 2.5 cm below the horizontal plane ($r_{\mathrm{rot}} = (0\ 0\ \text{-}2.5)$ cm). The fitted locations from the binaural conditions (filled symbols) form a narrow cluster. Their 95%-confidence regions (ellipses) at t_V have a diameter of about 0.4 cm. The locations of the monaural left and right conditions are found 0.5 cm to the right and left from the midline (x=0), respectively. The 95%-confidence regions of the monaural conditions (open symbols) have roughly twice the size of those from the binaural conditions (0.8 cm). For subject dj the spatial separation of the dipoles fitted to the monaural stimuli is most pronounced, it amounts to approximately 1 cm. In the average over subjects, the mean distance in x-direction between dipoles corresponding to monaural left and right stimuli is only 0.4 cm.

Figure 4.9 shows the dipole moment trajectories of the rotating dipole fitted in the interval from 1 ms before to 1 ms after peak V in the frontal plane. Data are averaged over subjects. The triangle denotes the start of the trajectory at t_V–1 ms. The ellipses drawn at t_V are the 95%-confidence regions for the dipole moment. They do not vary over time because neither the leadfield matrix $\mathbf{F}$ nor the covariance matrix $\mathbf{S}$ are time dependent (cf. eq. 4.8). The trajectories for the central stimuli, i.e., the diotic stimulus C00 and the antagonistic stimuli (C+– and C–+), exhibit the largest dipole moments in the vertical direction (m_z). With growing lateralization, m_z decreases and is smallest for the monaural conditions. The moment trajectories allow to distinguish between left and right conditions. The trajectories for stimuli which are lateralized to the right exhibit clockwise trajectories while counter-clockwise trajectories are found for left-lateralized stimuli. Furthermore, trajectories to conditions with ITD or ILD only look very similar. The trajectories to R+0 and R0+ on the one hand, and to L–0 and L0– on the other hand, bear a strong resemblance to each other. This is striking since they are produced by different physical

stimuli but elicit the same subjective lateralization.

Fig. 4.10 shows the quality measures of the rotating dipole fit. At t_V the residual variance (thin lines) ranges between 1% and 2.5% which is, on average, lower compared to the moving dipole fit. The goodness-of-fit (thick lines) is similar as in the case of the moving dipole fit although the number of free parameters per time sample was halved. This validates the concept of fitting a source with a constant location over a time interval. The gof for the monaural conditions is adequate, while the gof for the binaural conditions is insufficient, especially for the central stimuli.

The attempt to fit a rotating dipole to the binaural difference potential (BD) failed. The signal-to-noise ratio (SNR) was too low to allow for a stable solution. For most conditions and subjects, the dipole location was found at large excentricities, i.e., near the surface of the head model. For every subject and condition, the dipole location was therefore set to the location of the rotating dipole fitted to the corresponding binaural condition. This constraint assumes that the generator sites of the binaural potentials and the BDs are the same. The dipole moments were allowed to vary freely and independently for every time sample. The BD trajectories in the frontal plane, mean over subjects, are depicted in Fig. 4.11. A systematic dependence on stimulus conditions is hardly visible. The 95%-confidence regions drawn at t_V are large compared to the extension of the trajectories. However, for all stimulus conditions, the dipole moments have a predominantly vertical direction. The current direction is upward before t_V and downward after t_V, respectively. This corresponds well to the positive BD peak DP1 before t_V and the negative BD peak DN1 after t_V. Additionally, it can be seen that the vertical extension of the BD trajectories is largest for the central condition and smallest for the synergistic condition. This is in analogy to the BD results in chap. 3.

4.4.4 Two constrained fixed dipoles

An alternative, physiologically motivated model, is the assumption of two fixed dipoles with hemispheric constraints, i.e., mirrored x-component of the location and mirrored azimuth (α). The y- and z-component of the location and the elevation (θ) were forced to be identical for both dipoles. Expressed mathematically, the constraints were

$x_2 = -x_1, y_2 = y_1, x_2 = x_1, \alpha_2 = -\alpha_1$ and $\theta_2 = \theta_1$.

Unfortunately, fitting two fixed dipoles using these constraints leads to solutions with nearly identical locations of the dipoles in the vast majority of the cases, i.e., the fitting algorithm converges to a single dipole solution. To ensure the separation of the two

dipoles, a stronger constraint for the location parameters was introduced. For every subject the locations of the rotating dipole to all 15 stimulus conditions were averaged yielding $\vec{r}_{\text{rot}} = (x_{\text{rot}}, y_{\text{rot}}, z_{\text{rot}})$. The locations of the two fixed dipoles were defined by shifting the x-component of the rotating dipole dx cm to the left and dx cm to the right, respectively: $\vec{r}_{1,\text{fix}} = (x_{\text{rot}} - dx, y_{\text{rot}}, z_{\text{rot}})$, $\vec{r}_{2,\text{fix}} = (x_{\text{rot}} + dx, y_{\text{rot}}, z_{\text{rot}})$.

The superior olive (SO) and the nucleii of the lateral lemniscii (NLL) are considered as the likely generator sites of wave V (Scherg, 1991). According to an anatomical atlas of the brain (Nieuwenhuys et al., 1988), the distance between left and right SO is about 1.6 cm. The nuclei of left and right NLL reside at a distance of roughly 2.2 cm. Therefore, dx = 1 cm was chosen, i.e., the distance between the fixed dipoles was 2 cm. Additionally, it was assumed that the current direction in the neural tissue is the same for all stimulus conditions. Therefore, the orientation constraint was extended by fitting a common dipole orientation for all responses. Only the dipole moment magnitude varied between conditions.

Fig. 4.12 shows the locations (circles) and maximal moments (lines) of the two constrained fixed dipoles in the frontal plane (mean over subjects). The orientation of the dipoles is mainly vertical, they are inclined about 9° in the frontal plane and roughly 5° to the front in the sagittal plane. While the maximal moments for the binaural conditions are similar, for the monaural conditions the dipole contralateral to the side of stimulation clearly exhibits a larger moment.

The time courses of the dipole magnitudes of the two fixed dipoles are presented in Fig. 4.13 for the mean over subjects. The error bars mark the 95%-confidence regions of the moment magnitudes corresponding to ±2 standard errors of the mean. For the monaural stimuli, the maximum of the contralateral dipole is reached 0.4 ms earlier than the maximum of the ipsilateral dipole. This time difference is similar for conditions R0+, R+0, L0– and L–0 and slightly larger for the synergistic stimuli. For the central stimuli the latencies of the maximal dipole moment coincide.

In Fig. 4.14 the quality measures of the constrained fit of the two fixed dipoles is shown. Compared to the fit of the rotating dipole the residual variance is slightly raised, at t_{V} values up to 3% are found. The goodness-of-fit of this physiologically motivated model is worse than for the single rotating dipole. However, the number of free parameters per time sample was only two compared to three in the case of the rotating dipole.

4.4.5 Comparison of the forward models

To gain insight why the locations of two fixed dipoles, constrained only by hemispheric symmetry, did not automatically separate, the forward models of three dipole configurations were compared.

1. Potential $\vec{V}_0$ is generated by a single dipole with typical values of location and moment of a fitted rotating dipole:
   location: $\vec{r}_0 = (0\ 0\ -3)\,\text{cm}$,
   moment: $\vec{m}_0 = (0\ 0\ 8)\,\text{nAm}$.

2. Potential $\vec{V}_1$ is generated by a dipole pair at a distance of $2\,\text{cm}$ with orientations from the fit of the two constrained fixed dipoles:
   locations: $\vec{r}_{1,L} = (-1\ 0\ -3)\,\text{cm}$, $\vec{r}_{1,R} = (+1\ 0\ -3)\,\text{cm}$,
   moments: $\vec{m}_{1,L} = (+0.63\ 0\ 4)\,\text{nAm}$, $\vec{m}_{1,R} = (-0.63\ 0\ 4)\,\text{nAm}$.

3. Potential $\vec{V}_2$ is generated by a dipole pair at a distance of $4\,\text{cm}$ with orientations from the fit of the two constrained fixed dipoles:
   locations: $\vec{r}_{2,L} = (-2\ 0\ -3)\,\text{cm}$, $\vec{r}_{2,R} = (+2\ 0\ -3)\,\text{cm}$,
   moments: $\vec{m}_{2,L} = (+0.63\ 0\ 4)\,\text{nAm}$, $\vec{m}_{2,R} = (-0.63\ 0\ 4)\,\text{nAm}$.

Fig. 4.15 shows the dipole locations, moments and model EEGs in the frontal plane. The presentation can be restricted to two dimensions due to the rotational symmetry ensured by the vanishing y-components of the dipole locations and moments. The circle depicts the spherical head model with a typical head radius of 8.8 cm. The EEG is calculated in average reference and plotted in polar coordinates, at the head radius the potential is zero. At radii larger than the head radius the potential is positive, at smaller radii it is negative. The bars represent the most inferior electrode positions used in the EEG recordings. The potential caused by the single central dipole $\vec{V}_0$ (dash-dotted line) does hardly differ from the one generated by the dipole pair in a distance of $2\,\text{cm}$ $\vec{V}_1$ (solid line). There is a larger difference for the dipole pair with $4\,\text{cm}$ distance $\vec{V}_2$ (dashed line). Unfortunately, the largest differences are found at cervical areas which are experimentally unaccessible.

In principle, dipole separability is limited by finite spatial sampling and the residual noise of the measurement. The forward model was determined with a spacing of $1°$. The angular distance between adjacent channels in a 32-recording is roughly $20°$. However, since the sources considered are deep, the EEG does not exhibit steep gradients indicating that

the poor spatial sampling is not the main reason for the indistinguishability of the source configurations compared.

To illustrate the influence of the measurement noise, in Fig. 4.16 the differences between the EEGs caused by the two-dipole configurations and the single central dipole are depicted. The residual noise is assumed to have a typical average value $\bar{\sigma} = 13\,\text{nV}$ at all channels, $\pm\bar{\sigma}$ is plotted as a gray area around the circle signifying the head radius. The difference between the dipole pair at $2\,\text{cm}$ distance and the central dipole $\vec{V_1} - \vec{V_0}$ is drawn as a solid line. It deviates from the noise area at the unaccessible inferior sites, but also at lateral positions, where the maximal difference is found to be $1.8\ \bar{\sigma}$. $\vec{V_2} - \vec{V_0}$ is negative at the top and clearly positive at lateral sites. This means that a dipole pair at $4\,\text{cm}$ distance should be distinguishable from an equivalent central dipole. A dipole pair at only $2\,\text{cm}$ distance is hardly separable from the single dipole. It might be possible to distinguish between both source configurations for subjects with very low residual noise.
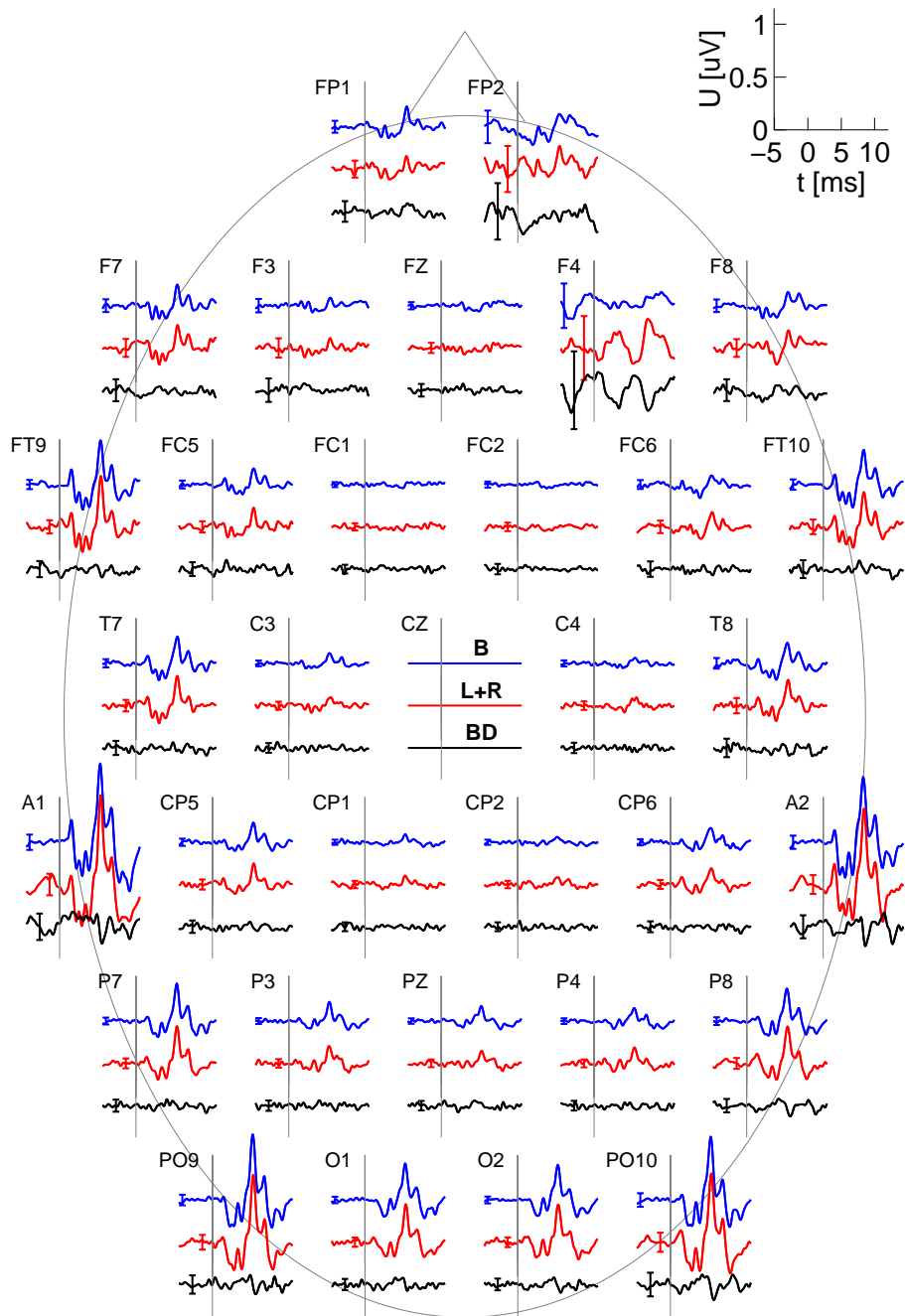
Fig. 4.1: ABRs recorded from 32 channels for subject jo. Data are plotted in recording reference, the reference electrode was CZ. **Top trace:** Binaural response to the diotic click (B). **Middle trace:** Sum of responses to monaural clicks (L+R). **Bottom trace:** Binaural difference potential (BD). The error bars indicate ±3 standard errors of the mean.
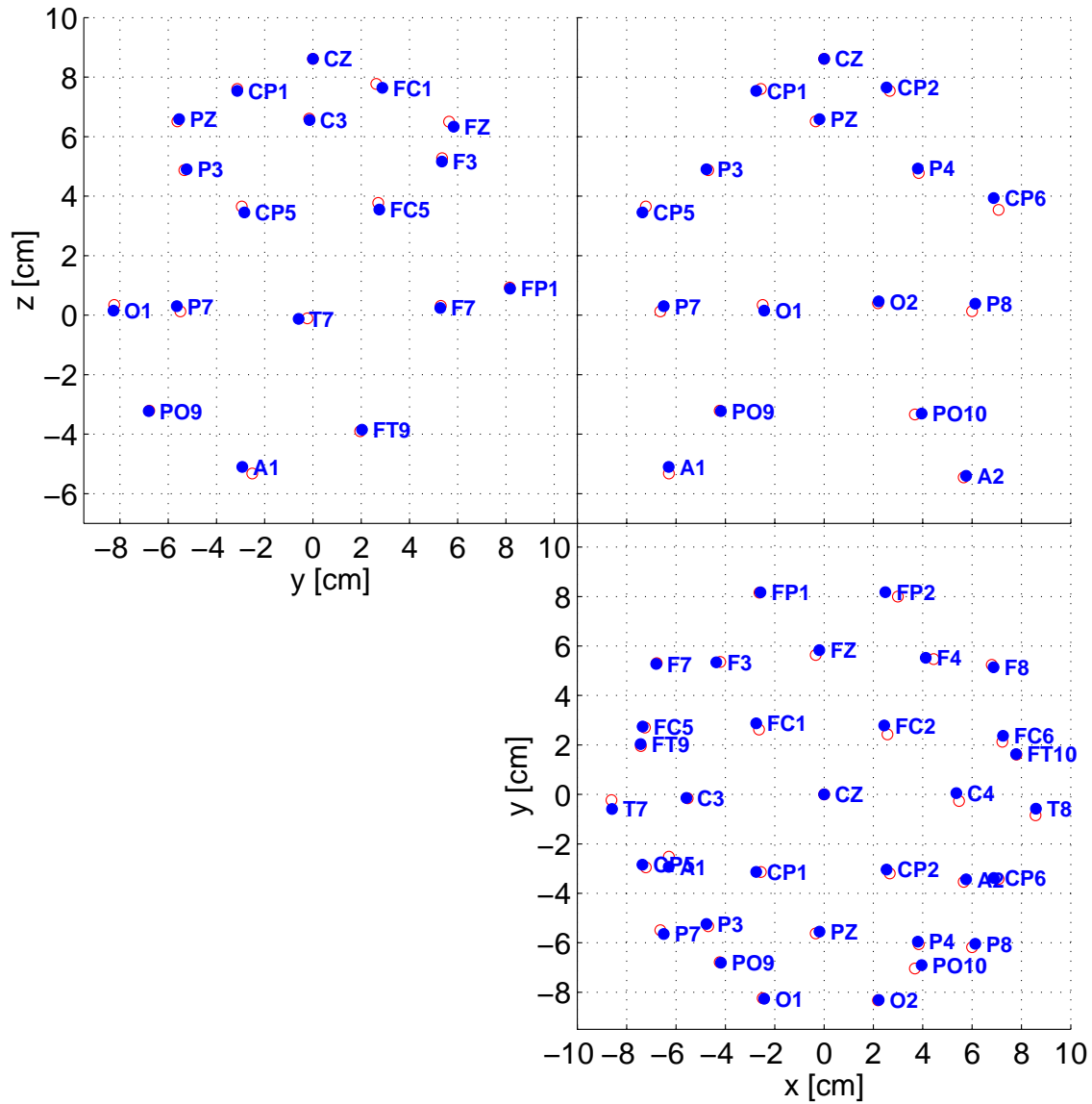
Fig. 4.2: *Projections of the measured electrode positions onto three planes for subject rh.*
x points to the right, y to the front and z to the top. The electrode labels are according
to the extended 10-20-system. **Filled symbols:** *Electrode positions measured before the*
recording session. **Open Symbols:** *Electrode positions measured after the recording*
session.
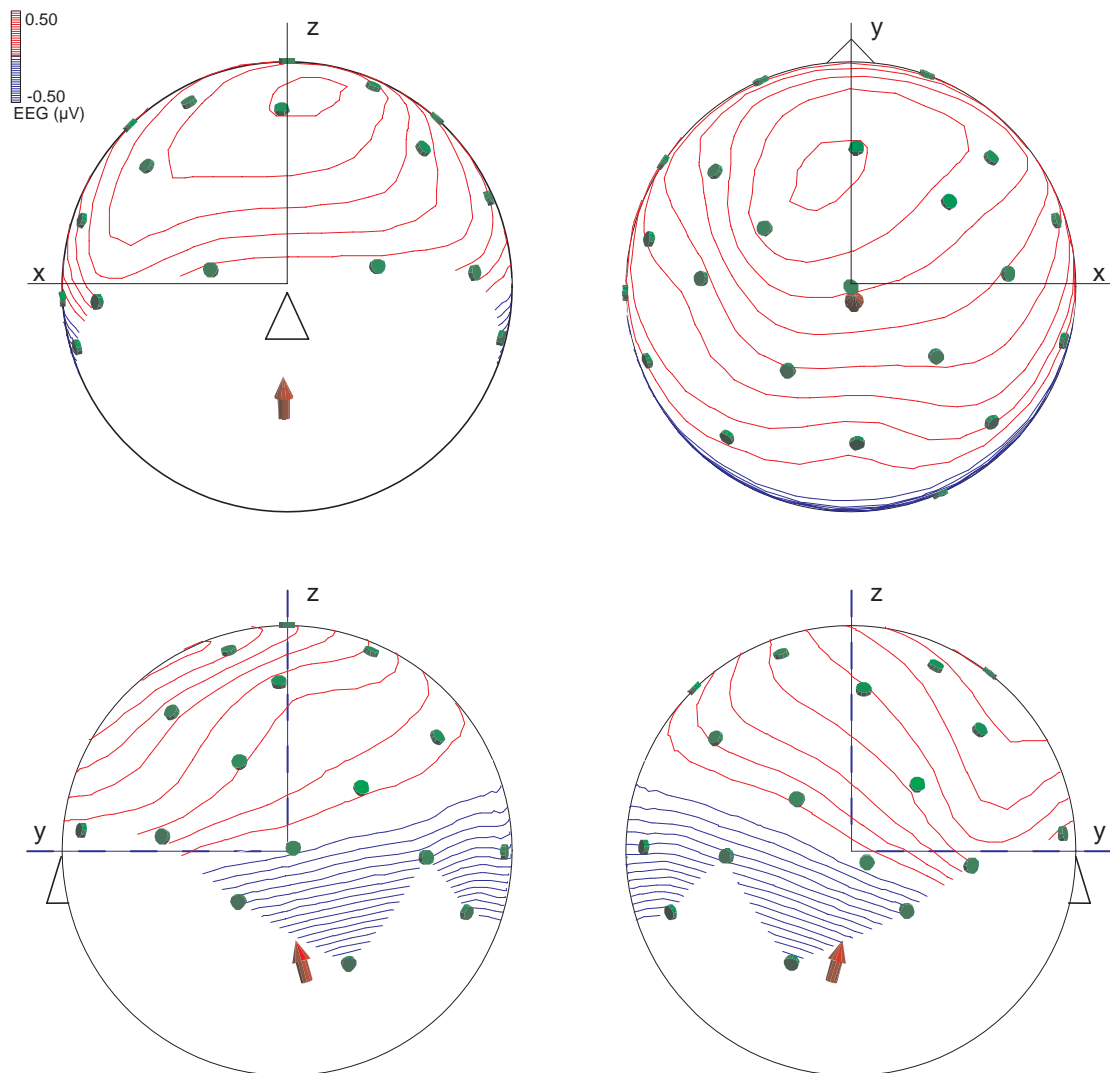
Fig. 4.3: Interpolated equipotential maps in average reference at a latency of 5.8 ms after the onset of the diotic click, subject jo. x points to the right, y to the front and z to the top. The triangles indicate the nose. The black line denotes zero potential. Dark gray lines indicate positive voltage, light gray lines negative voltage. The voltage difference between adjacent equipotential curves is 25 nV. **Top left:** *view from the front,* **top right:** *view from the top,* **bottom left:** *view from the left side,* **bottom right:** *view from the right side. The arrow in each subplot indicates the result of a moving dipole fit, i.e., only data at the latency shown entered the fit routine.*
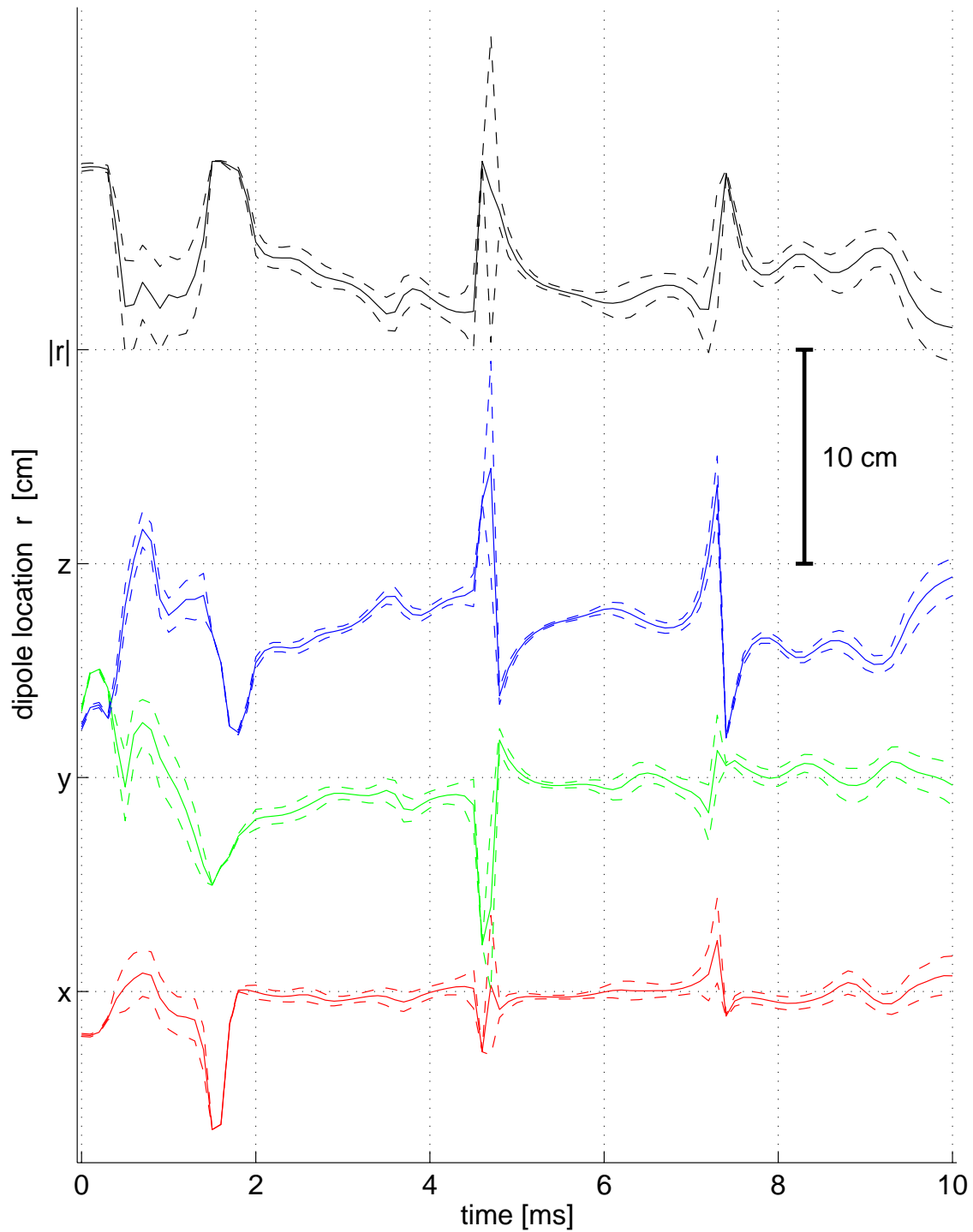
Fig. 4.4: Location of a moving dipole fitted to the diotic response of subject cr. The lower three curves depict the coordinates x, y and z, respectively. The upper curve shows the magnitude $|\vec{r}|$ of the location vector. The dashed lines denote the 95%-confidence regions for the dipole location.
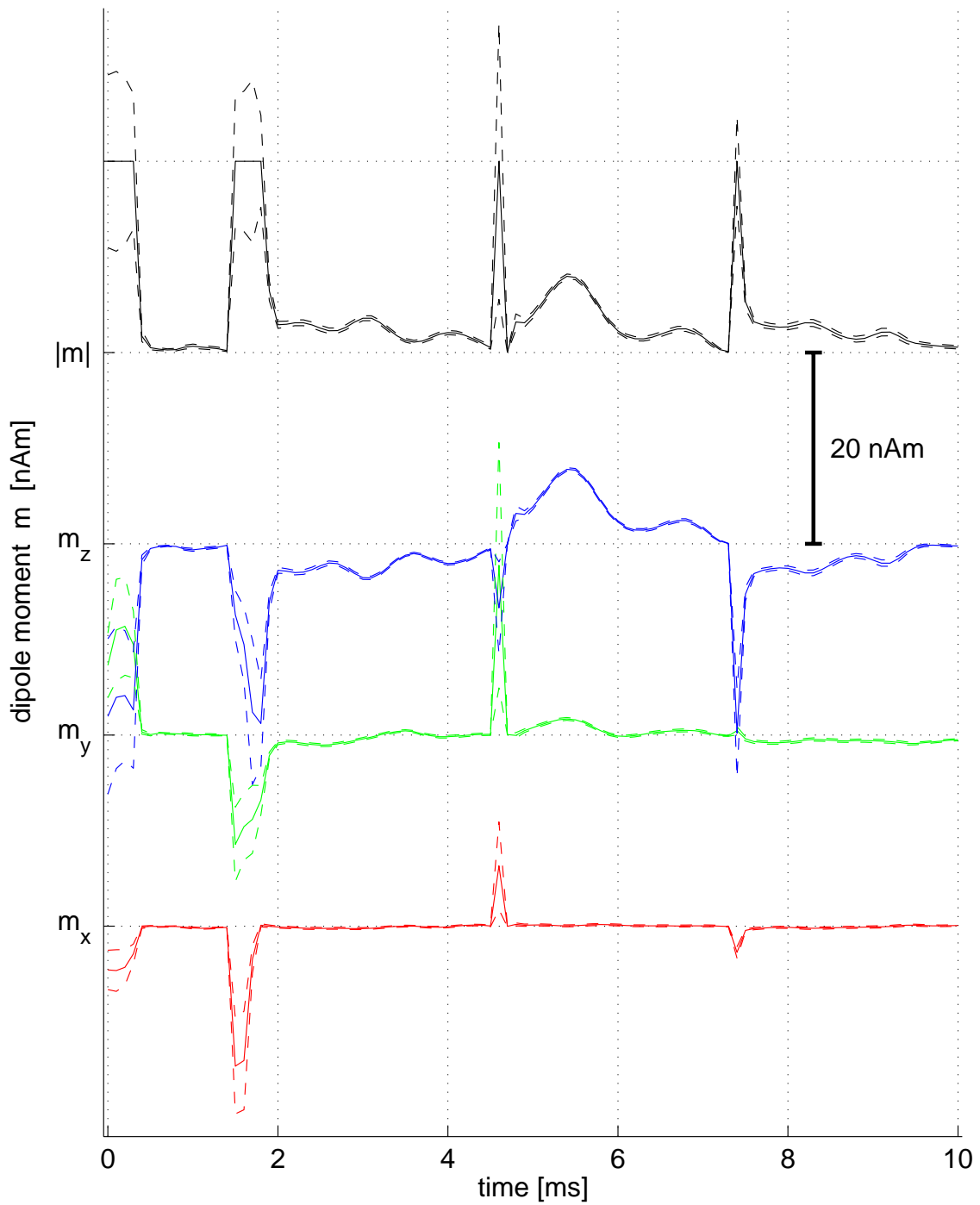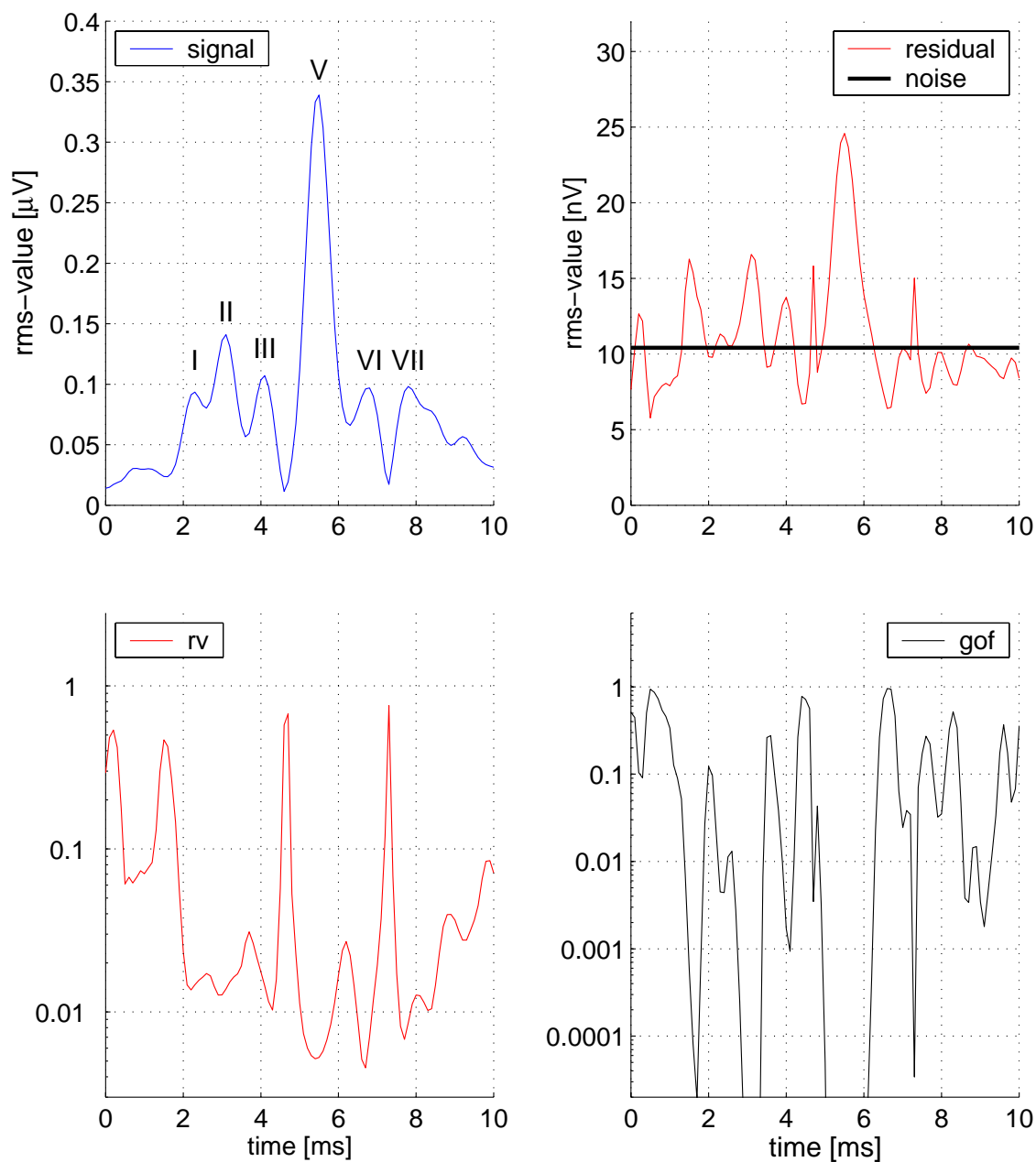
Fig. 4.5: Dipole moment of a moving dipole fitted to the diotic response of subject cr. The lower three curves depict the coordinates of the moment m_x, m_y and m_z, respectively. The upper curve shows the magnitude $|\vec{m}|$ of the dipole moment. The dashed lines denote the 95%-confidence regions for the dipole moment.

Fig. 4.6: *Signal characteristics and measures of the fit quality for the moving dipole fit. Data from subject cr, response to the diotic click.* **Top left:** *Signal rms, roman numerals denote the peaks of the ABR.* **Top right:** *rms of the difference between signal and fit (residual, thin line) compared to the residual noise averaged over time and channels (thick line).* **Bottom left:** *residual variance.* **Bottom right:** *goodness-of-fit.*
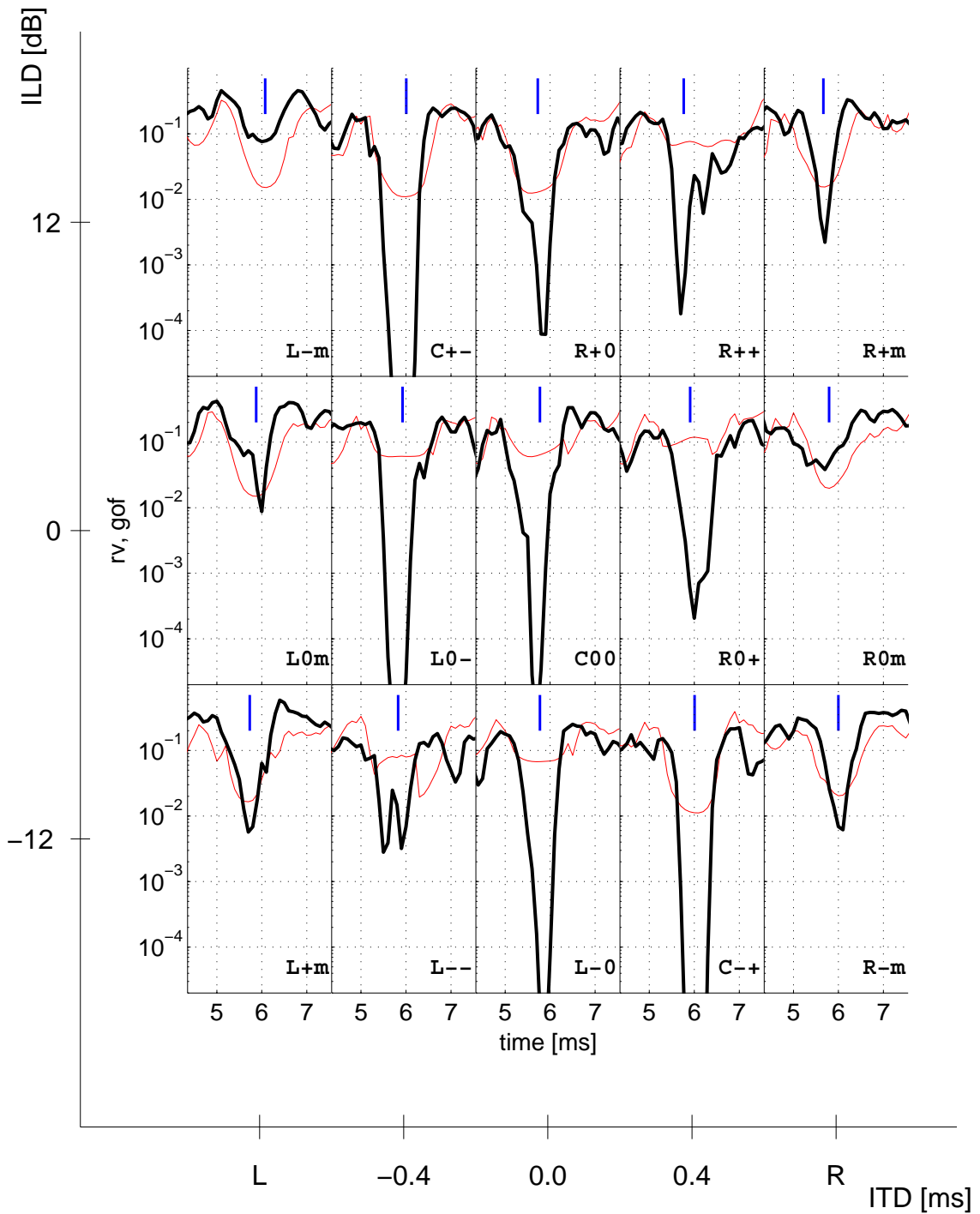
Fig. 4.7: Residual variance (rv, thin line) and goodness-of-fit (gof, thick line) of the moving dipole fit for all stimulus conditions, mean over subjects. The vertical bars mark the mean latency of wave V.
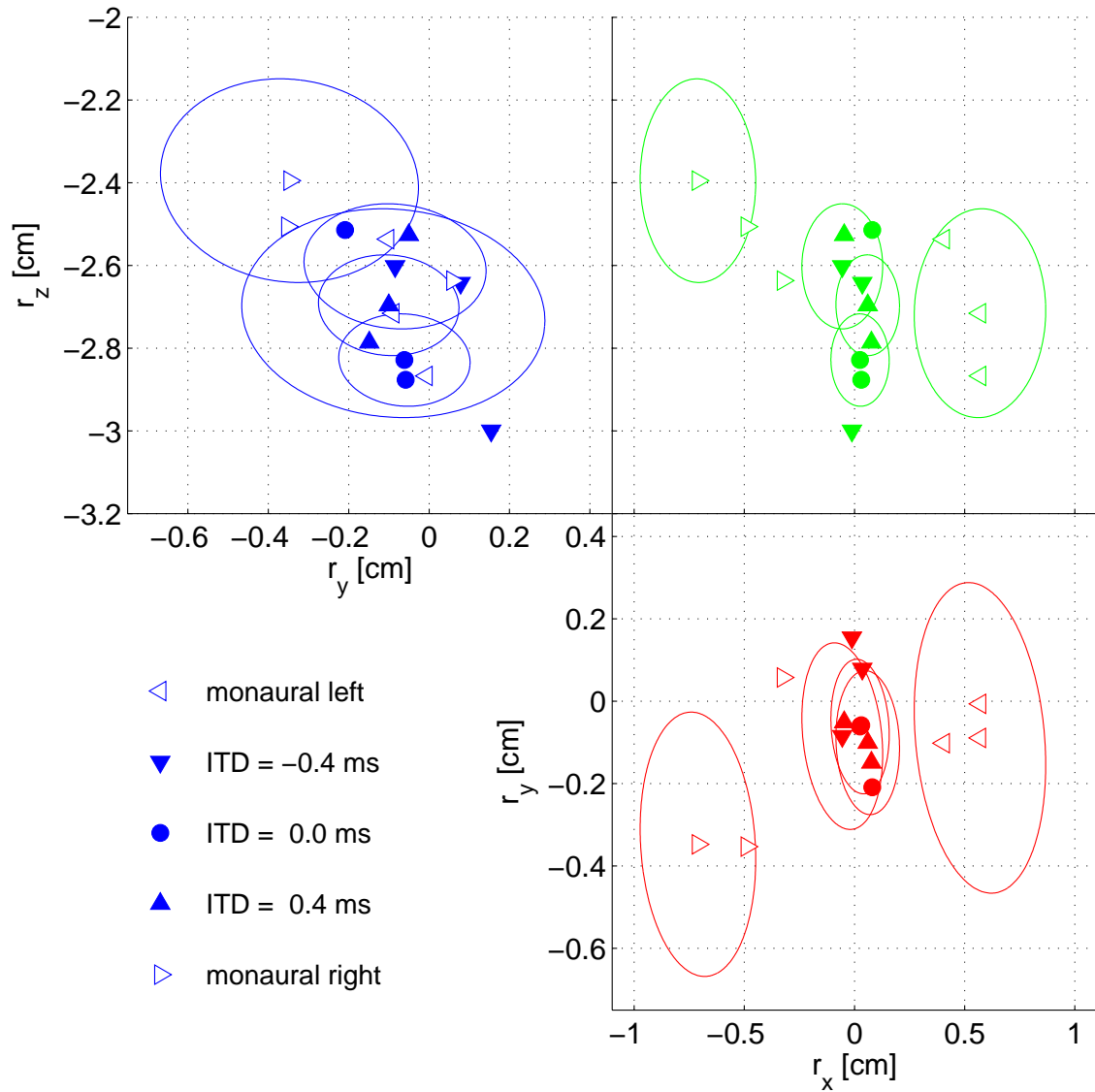
Fig. 4.8: Locations of rotating dipoles fitted to the 15 conditions for subject dj. x points to the right, y to the front and z to the top. The fit interval began 1 ms before and ended 1 ms after peak V. The 95%-confidence regions (ellipses) hold for the latency of wave V (t_V), they are only drawn for the five stimuli with ILD = 0 dB. Filled symbols denote binaural stimuli, open symbols stand for monaural stimuli.
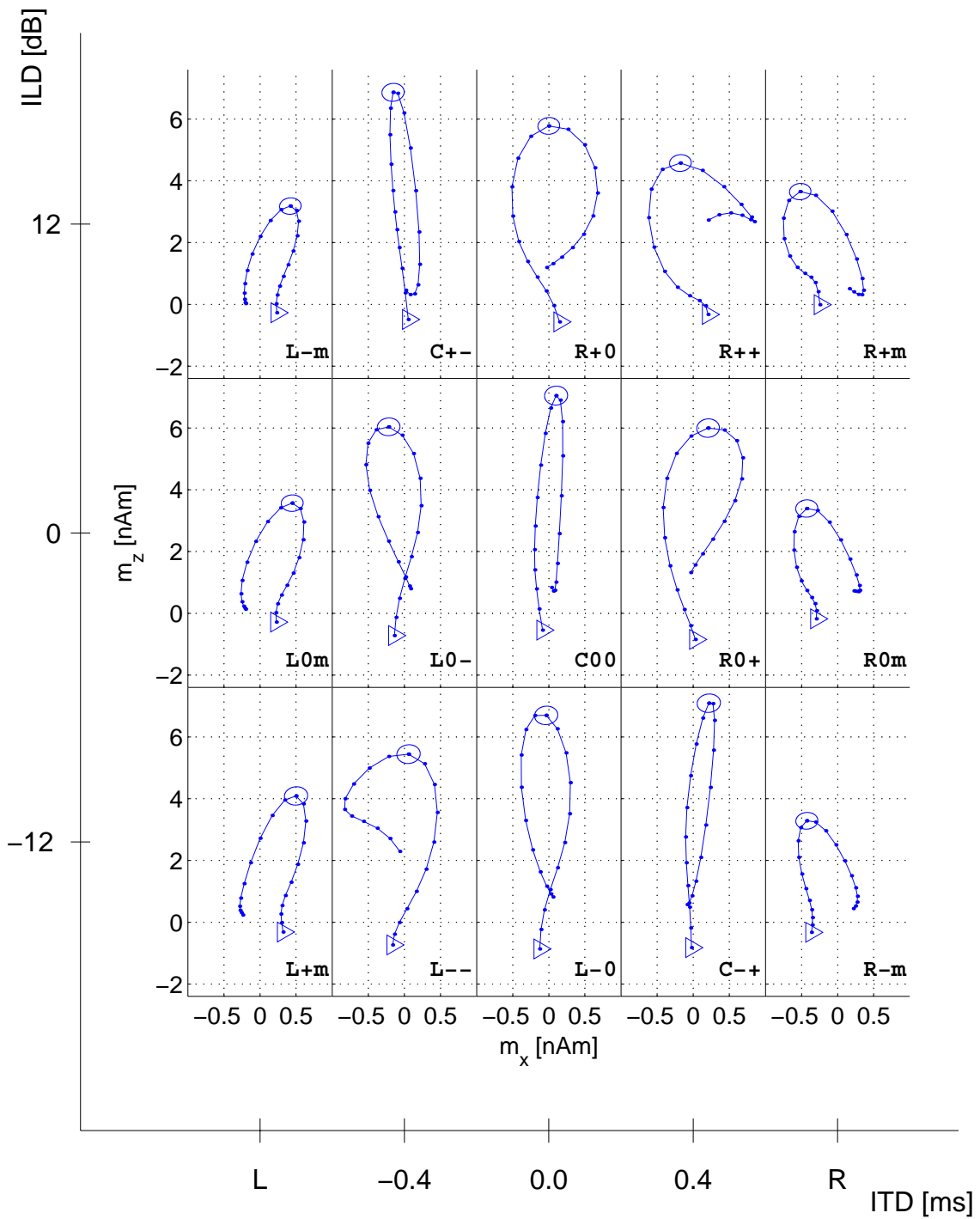
Fig. 4.9: Dipole moment trajectories of a rotating dipole for the 15 stimulus conditions in the frontal plane, mean over subjects. x points to the right and z points to the top. The x-coordinate points to the right, z to the top. The fit interval lasted 2 ms, from 1 ms before (triangles) to 1 ms after peak V. At the latency of wave V, ellipses denoting the 95%-confidence regions for the moment are drawn.
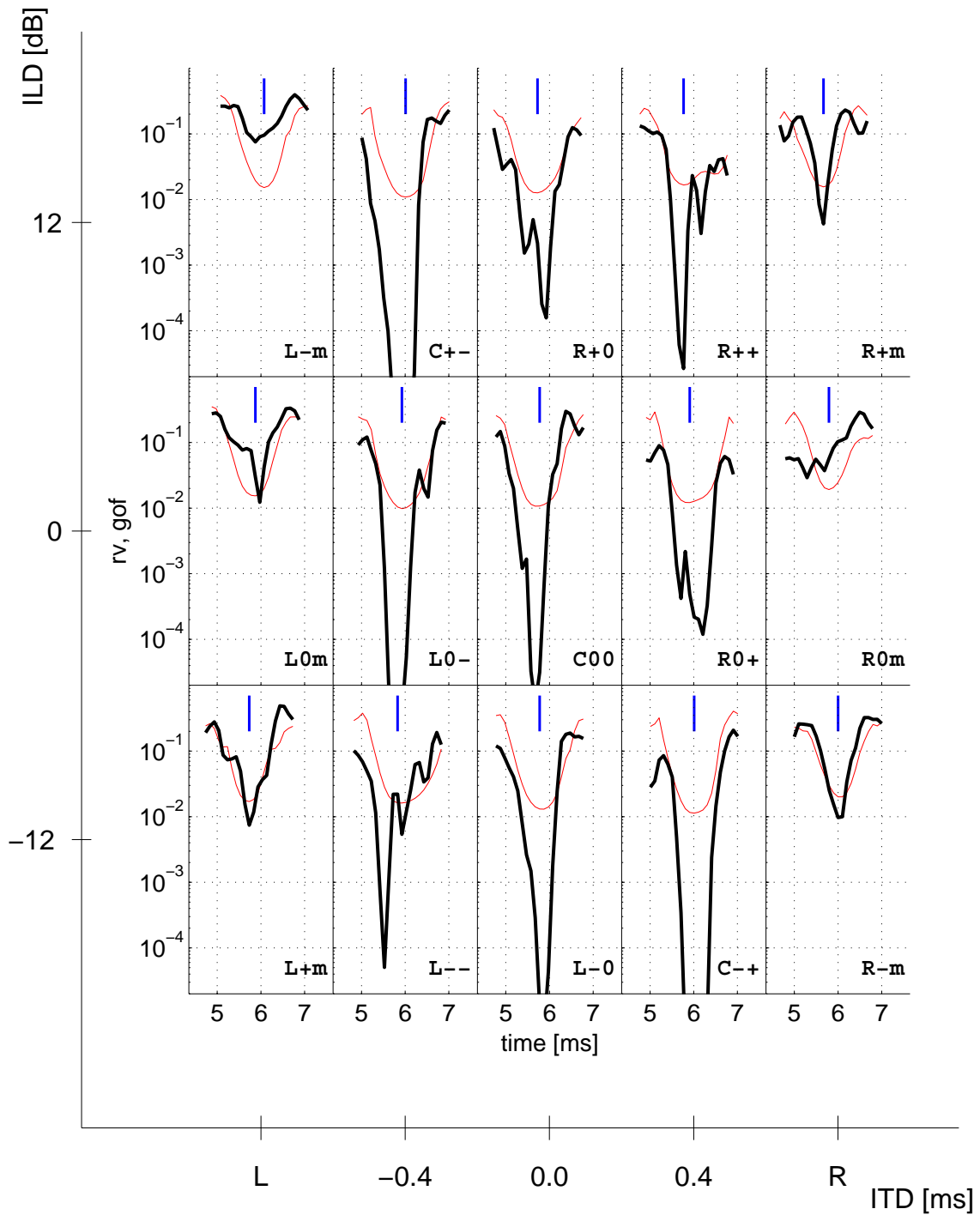
Fig. 4.10: Residual variance (rv, thin lines) and goodness-of-fit (gof, thick lines) of the rotating dipole fit for all stimulus conditions, mean over subjects. The vertical bars mark the mean latency of wave V.
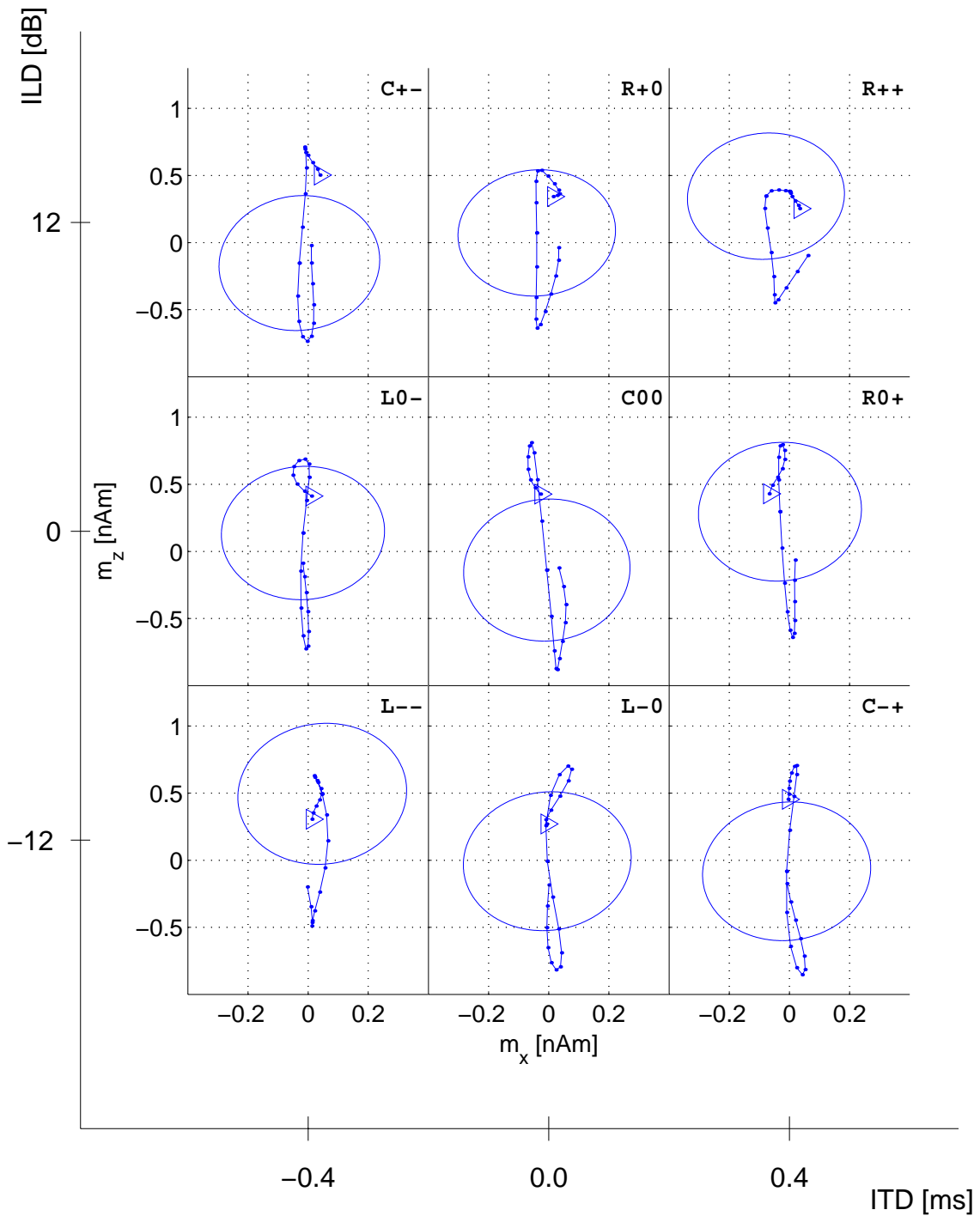
Fig. 4.11: Dipole moment trajectories of a rotating dipole for the 9 BD conditions in the frontal plane, mean over subjects. x points to the right and z points to the top. The fit interval lasted 2 ms, from 1 ms before (triangles) to 1 ms after peak V. At the latency of wave V, ellipses denoting the 95%-confidence regions for the moment are drawn.
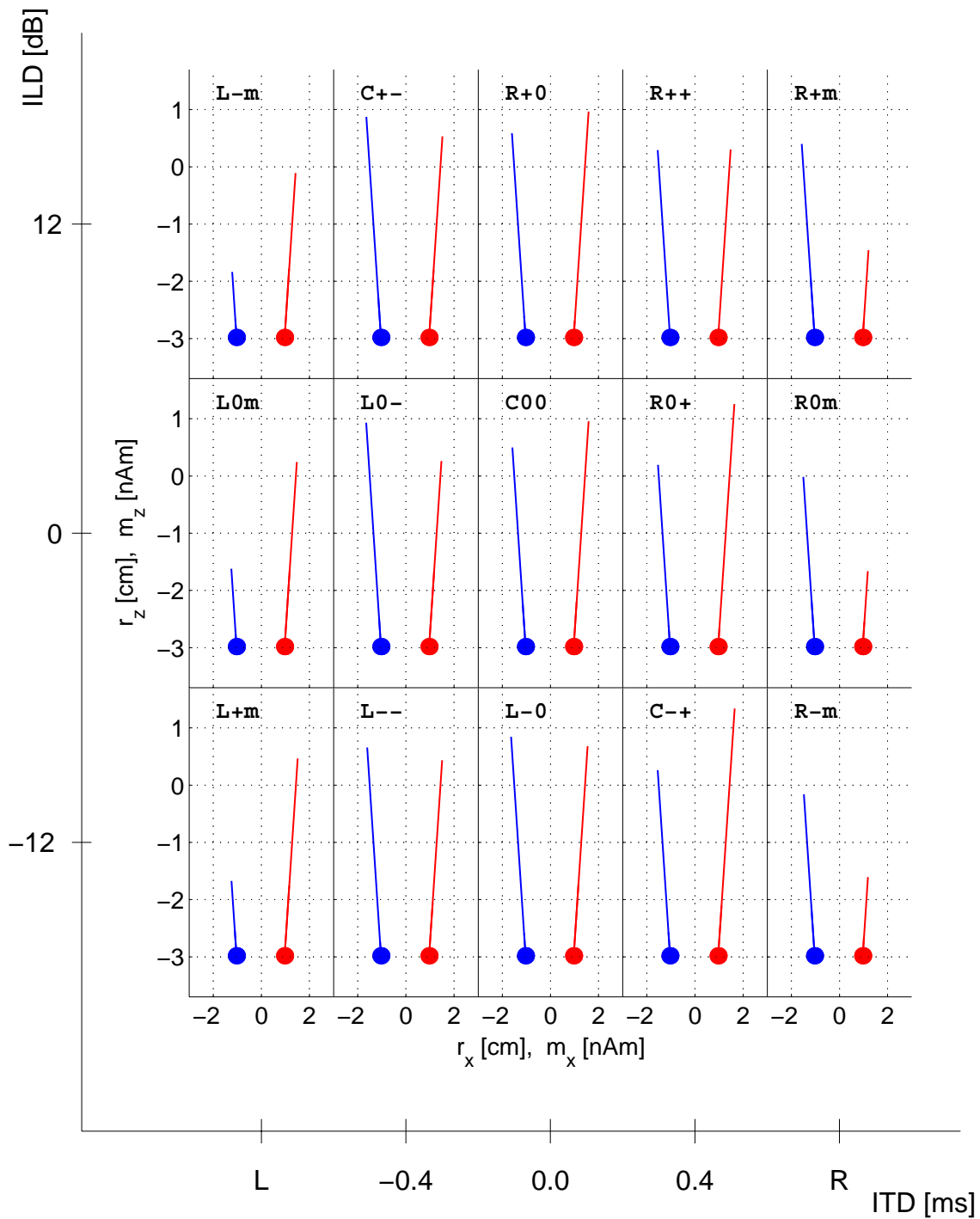
Fig. 4.12: Dipole locations (filled circles), orientations and maximal magnitudes (lines) of two constrained fixed dipoles for all stimulus conditions in the frontal plane, mean over subjects. x points to the right and z points to the top. The dipole locations were set to the mean location of rotating dipole fit. ±1 cm was added to the x-component of the mean location. The orientations of the two dipoles were constrained to be identical for all stimulus conditions and to have the same elevations, but mirrored azimuths.
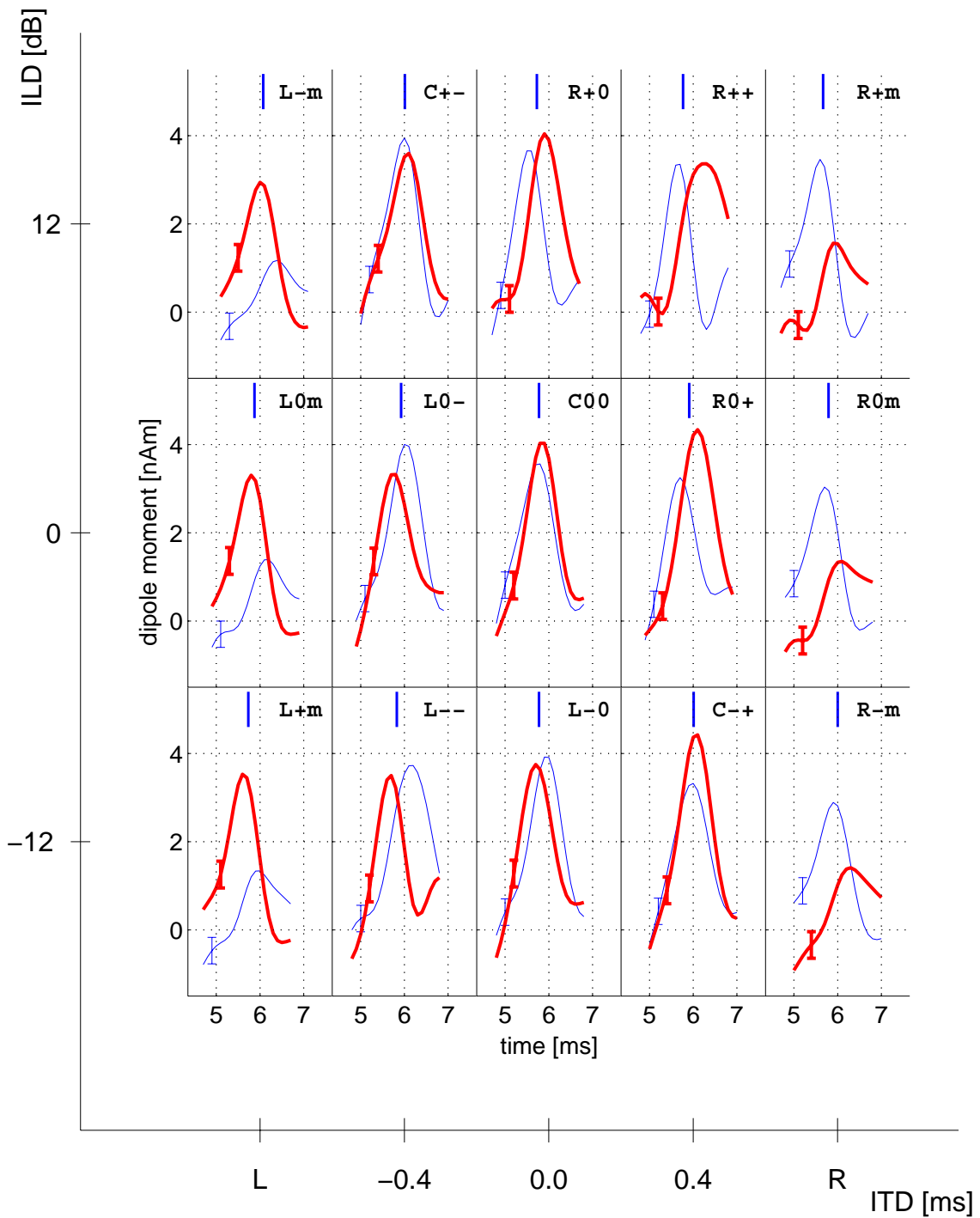
Fig. 4.13: Dipole moment magnitudes of two constrained fixed dipoles for all stimulus conditions in the frontal plane, mean over subjects. **Thin curves:** dipole in the left hemisphere, **thick curves:** dipole in the right hemisphere. The error bars denote the 95%-confidence regions of the moment magnitudes, the vertical bars mark the mean latency of wave V.
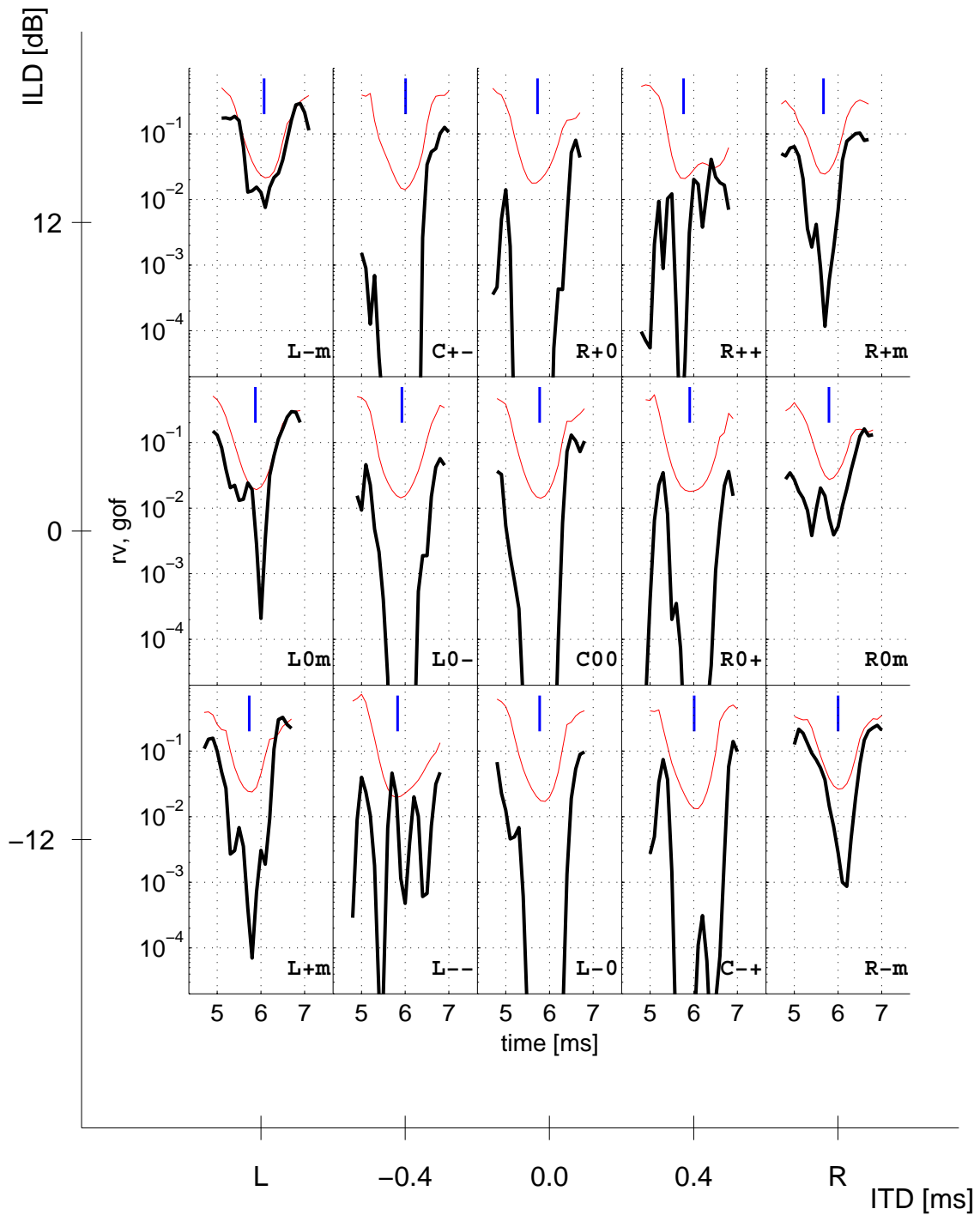
Fig. 4.14: Residual variance (rv, thin lines) and goodness-of-fit (gof, thick lines) of the fit using two constrained fixed dipoles, mean over subjects. The vertical bars mark the mean latency of wave V.
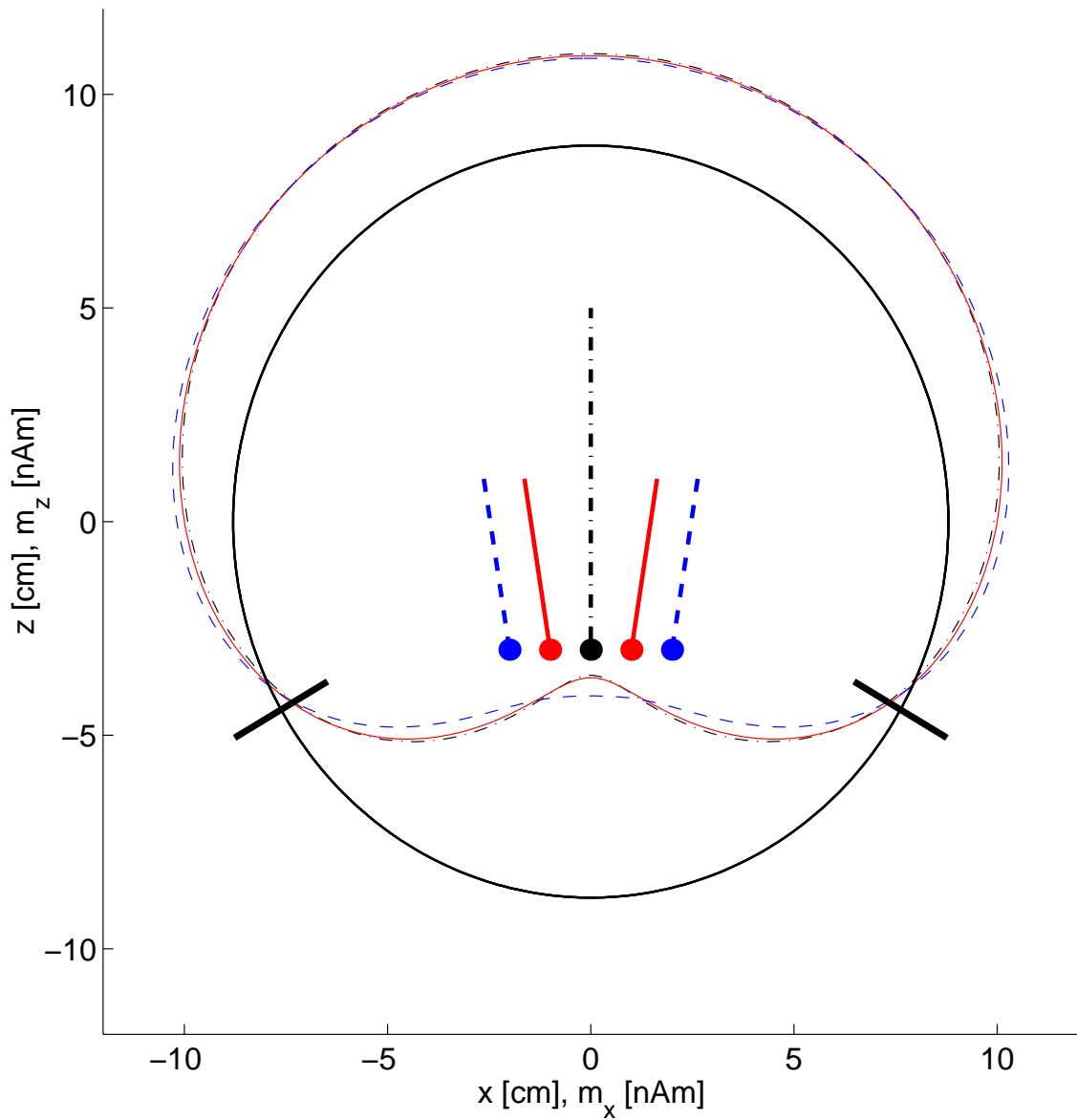
Fig. 4.15: Forward models in average reference of three dipole configurations in the frontal plane: 1. a central vertical dipole (dash-dotted), 2. a pair of inclined dipoles at a distance of 2 cm (solid), 3. a pair of inclined dipoles at a distance of 4 cm (dashed). The head is symbolized by the circle, the bars mark the positions of the most inferior electrodes. x points to the right and z points to the top. The EEG is illustrated as a polar plot: at the head radius, the voltage is zero. Potentials are negative inside the circle and positive outside the circle.
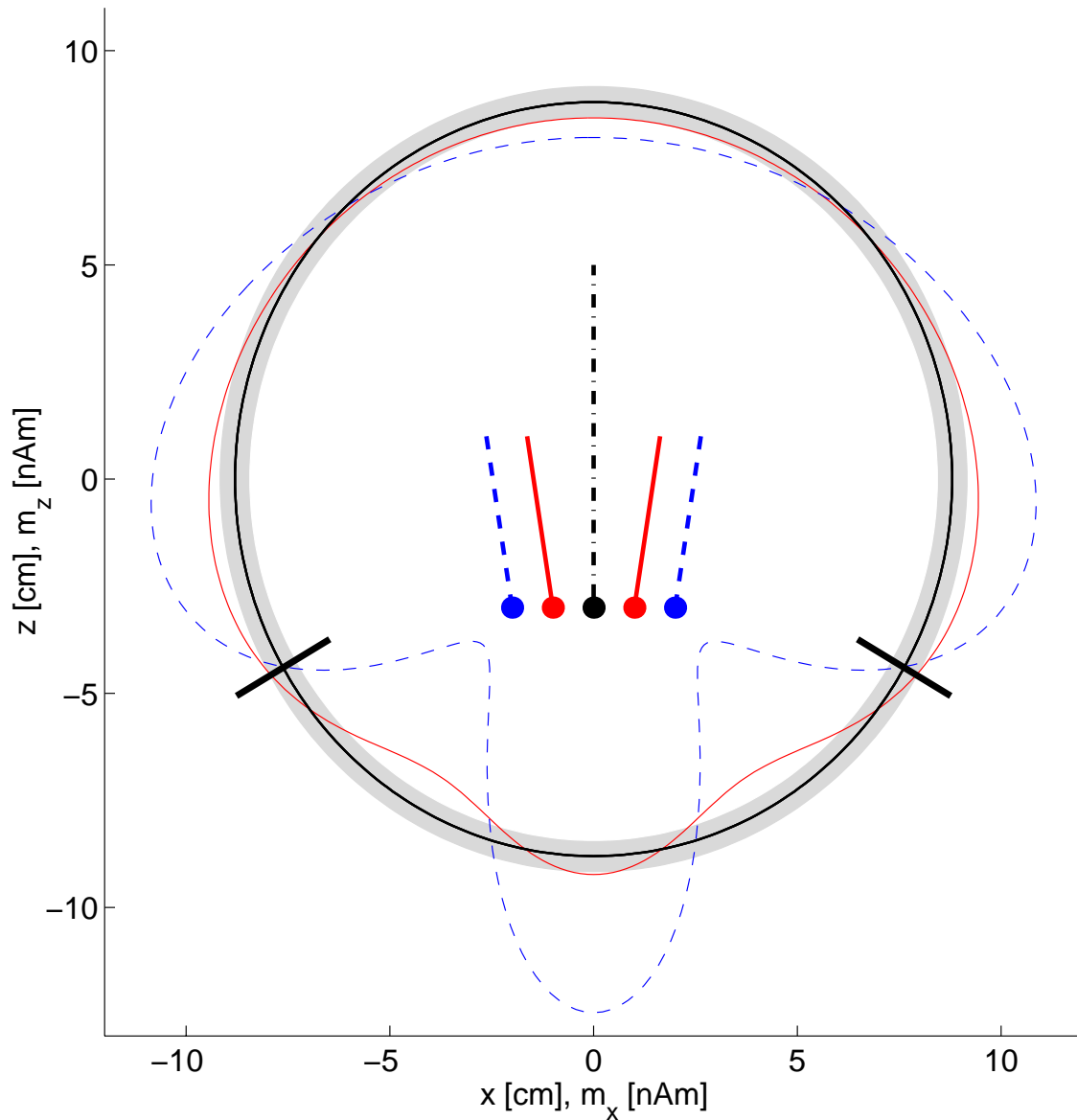
Fig. 4.16: *Differences of forward models in average reference in the frontal plane: 1. difference between a central vertical dipole and a pair of inclined dipoles at a distance of 2 cm (solid), 2. difference between a central vertical dipole and a pair of inclined dipoles at a distance of 4 cm (dashed). The head is symbolized by the circle, the bars mark the positions of the most inferior electrodes. x points to the right and z points to the top. The EEG is illustrated as a polar plot: at the head radius, the voltage is zero. Potentials are negative inside the circle and positive outside the circle. The gray shaded area illustrates an average residual noise of ±13 nV.*

4.5 Discussion

The aim of this work is to analyze the correspondence between psychophyical lateralization and the neural generators of potentials evoked by lateralized stimuli. Since the generators of early auditory evoked potentials are deep, i.e., reside in the brain stem, data exhibit a relatively low SNR (signal-to-noise ratio). Therefore a large number of sweeps per stimulus condition (10000) was collected to ameliorate the signal quality.

4.5.1 Incorporation of the noise covariance matrix

To enhance and accurately determine the quality of the solutions of the inverse problem, information of the measurement errors based on the single sweeps of the EEG recordings was used (see chap. 2). The noise covariance matrix allows (i) a more general formulation of the cost function, (ii) the computation of confidence regions for source parameters by error propagation, and (iii) the evaluation of the fit quality by means of a statistical test. Generalized maximum likelihood estimation was first used by Sekihara et al. (1992) who analyzed the biomagnetic inverse problem. Lütkenhöner (1998a,b) demonstrated the advantage of the incorporation of noise covariance both theoretically and with MEG (magnetoencephalogram) data from late auditory evoked potentials.

4.5.2 Goodness of fit versus residual variance

An important methodological improvement pertains to the judgement of fit quality. In EEG literature the residual variance rv is commonly used to determine the quality of the fit. It has two major drawbacks. First, it introduces a kind of 'quadratic bias'. A residual variance of 1% sounds reasonably small, but is equivalent to a residual standard deviation of 10%. Second, the rv does not compare the residual with the noise in the measurement. Therefore it should be considered as not being appropiate for rating the fit quality. Without noise information it is also impossible to infer confidence regions of source parameters. Commercial software programs like BESA (Brain Electrical Source Analysis) and ASA (Advanced Source Analysis) transform EEG data to average reference before dealing with the inverse problem. In both programs noise information is not considered, i.e., $\mathbf{S}'$ is set to the unit matrix. The fit quality is expressed in terms of rv. Given the sophisticated algorithms which are used in data analysis, e.g., distributed source models or the computation of forward models using realistically shaped head models, it is surprising that the judgement of noise and SNR of data has such a low significance.

4.5.3   The problem of the reference electrode

In MEG studies a reference sensor does not exist, since all sensors record components of the magnetic field (or their spatial derivatives). In EEG studies, however, the electric field is measured by voltage differences between electrodes. Usually, a common reference electrode is used for all other electrodes. A large body of literature has emerged discussing the advantages and disadvantages of the transformation of EEG data from recording reference to average reference. Bertrand et al. (1985) justified the choice of the average reference by the notion that the integral of the potential distribution over a sphere that includes current dipoles is zero. Onofrj et al. (1994) detected focal abnormalities in the event related potential P3 in patients with cortical lesions when they used the average reference instead of a linked earlobe reference. However, other authors describe severe drawbacks of the average reference (Desmedt and Tomberg, 1990; Desmedt et al., 1990b; Tomberg et al., 1990). Desmedt et al. (1990a) pointed out that the average reference creates problems because it dynamically zero-centers all potentials at each latency resulting in so-called 'ghost potentials'. The reason for these spurious effects is the lack of electrodes in the bottom half of the head. However, this criticism mainly concerns the mapping of brain potentials, i.e., the interpolation of measured potentials at the electrode sites to equipotential maps. Often generators of the EEG are inferred visually from these maps without mathematically tackling the inverse problem.

At first glance, the criticism on the average reference becomes irrelevant if the interest is focused on the results of source modeling. Without noise information, i.e., merely using a least-squares approach, all supposable references should work equally well. The rationale is that forward models generate potential distributions with respect to a reference electrode at infinite distance. These potentials are subsequently transformed into the desired reference. However, if the noise covariance matrix is incorporated into the description of the inverse problem, it turns out that a solution in average reference is impossible. The reason is simply that the noise covariance matrix in average reference is singular, see appendix B. One should keep in mind that maximum-likelihood estimation and least-squares fitting differ from generalized maximum likelihood estimation only by simplifications of the noise covariance matrix. This means that the treatment of the inverse problem in average reference is only possible by simplifying the noise covariance matrix. Consequently, source analysis should be carried out in recording reference.

4.5.4   The moving versus the rotating dipole

In the present chapter three source models were compared to explain the multi-channel ABR data. As most simple model without any constraints the moving dipole was adopted. Monaural data are explained with good fit quality while binaural data show small goodness-of-fit values in the latency range of wave V. The second approach was a rotating dipole in a 2-ms time interval centered at the latency of peak V. From a physiological standpoint, a rotating dipole is more satisfactory than a moving dipole. Although at different latencies sources at different brain locations may be active, current sources do not 'move' within the brain. A source at a constant location with time-varying moment better complies with the physiological point of view. With half the number of parameters the rotating dipole achieves smaller residual variances and roughly the same goodness-of-fit as the moving dipole. Additionally, the rotating dipole fit unveils characteristics of the generators. Centrally perceived stimuli cause trajectories of the dipole moment in the frontal plane that mainly extend in the vertical direction. Lateral stimuli generate trajectories with smaller vertical but larger horizontal extension. This corresponds well to the results from chap. 3, taking into account that the single channels A1, A2, PO9 and P010 are orientated predominantly vertical and therefore map the vertical component m_z of the source dipole. The laterality of the stimulus is coded in the sense of rotation of the trajectory. The moment trajectories of the rotating dipole do not code the ITD or ILD alone, but show a striking correlation with the lateralization of the stimuli (see Figs. 3.1 and 4.9), i.e., stimuli with similar lateralization cause similar dipole moment trajectories. This means that ITD and ILD are not processed independently in the brain stem.

Despite the small extension of the brain stem, significantly different dipole locations were detected for monaural and binaural stimuli. For binaural stimuli a centered source with a 95%-confidence region radius as small as 2 mm is found for subjects with high SNR. For monaural stimuli the fitted dipole position is found in the contralateral hemisphere. This is physiologically meaningful because the majority of the auditory fibers projects to contralateral nuclei in the brain stem (Nieuwenhuys et al., 1988). However, the distance between left and right fitted source amounts to maximally 1 cm. Given the anatomical distance of the likely generators of wave V, namely 1.6 cm for the superior olives and 2.2 cm for the nuclei of the lateral lemniscii, the fitted distances appear too small.

Two reasons are conceivable to explain this discrepancy. First, the homogeneous sphere which served as head model may be too simple because it does not model the attenuation effect of the skull. Compared to the brain tissue and the skin, the conductivity of the skull

is about 80 times smaller (Cuffin and Cohen, 1979). Ary et al. (1981) compared the homogeneous sphere with a three-shell head model. They showed that a dipole in the 3-shell head model must have a larger excentricity to generate approximately the same EEG as an identically oriented dipole in a homogeneous sphere. Second, for monaural stimulation the ipsilateral generators will also be activated, albeit weaker. The fitting algorithm has to optimize a single source that must explain two sources of different strengths. It consequently finds a best matching location between both sources, but nearer to the stronger source. From this consideration the extension to a model with two sources evolved naturally.

4.5.5   The rotating versus constrained fixed dipoles

As third source model, a pair of hemispherically symmetric fixed dipoles was chosen. The attribute 'fixed' means that their orientations were fitted but required to be constant during the fitting interval. This constraint is physiologically motivated by the idea that in an activated brain area the direction of the current should remain constant since the orientation of the nerve fibers does not change. However, from a mathematical point of view, each rotating dipole can be considered as a superposition of three perpendicular fixed dipoles at the same location.

For most subjects and stimulus conditions, the two fixed dipoles converged to nearly the same location representing a single rotating dipole that could rotate in only two dimensions. A separation of the two dipoles could only be ensured by constraining the dipoles to have an a-priori known position which was determined from physiological constraints. A distance of 2 cm was chosen to reflect the distance of the nuclei involved in the generation of wave V. For the monaural conditions, the resulting moments show a stronger activation of the contralateral dipole which is physiologically plausible. For the binaural conditions, the differences in the maximal amplitude of the moments have the tendency of being larger at the ipsilateral site. In addition, for the lateralized binaural stimuli, the latency of the maximal dipole moment is larger for the ipsilateral dipole. This undermines the physiological plausibility of the model since a smaller latency was expected for the maximum of the ipsilateral dipole. An explanation could be that the dipole with the smaller latency reflects the activity of the contralateral SO while the other dipole maps activity of the ipsilateral NLL, presuming that the latter is activated later than the contralateral SO. It remains unclear if the ipsilateral SO and the contralateral NLL should be modeled by additional sources. In such a case, the model would have to be augmented to four sources: left SO, right SO, left NLL, right NLL. However, on the basis of the data, i.e., without

physiological constraints, it already appeared to be impossible to separate two sources. Therefore one should concede that a constrained four-dipole model will pose even harder separation problems than the two-dipole model.

4.5.6 Possible improvements

The comparison of the model EEGs of a single central dipole and a pair of dipoles with half the moments and symmetric positions around the central dipole clearly demonstrates the ambiguity of the inverse problem. In this example, the separability is mainly restricted by the residual noise. Since the sources are deep, insufficient spatial sampling seems to play a minor role.

The separation of sources within the brain stem imposes great demands on the signal quality and is a task certainly hard to accomplish. Some improvements are conceivable to approach this task. A three-shell model of the head can be used instead of the homogeneous sphere. It is generally believed that by means of realistically shaped head models derived from MRI (magnetic resonance imaging) scans, the localization accuracy can be augmented. However, recently Cuffin et al. (2001) compared the source localization errors of the three-shell model with a realistically shaped head model. The sources were created by injecting current into implanted depth electrodes of human patients. Virtually no difference occured in the localization errors of both models.

Alternatively, the number of recording channels can be raised to improve spatial sampling. Residual noise can be further reduced by even longer recording sessions. Presumably, evoked potentials to only a few stimulus conditions can be collected. Instead of the click stimulus a rising frequency chirp (Dau et al., 2000) generating larger evoked potentials with a higher SNR can be used. Data from this approach are presented in the subsequent chapter.

4.6 Summary and conclusions

- The incorporation of the noise covariance matrix into the algorithm to solve the inverse problem allows for important methodological improvements.

- The goodness-of-fit, comparing the residual with the noise, is preferable to the residual variance which relies only on the residual and does not account for the noise in the recording.

- The generalized inverse problem has to be treated in recording reference. In the commonly used average reference the noise covariance matrix is singular.

- The rotating dipole well explains the data around the latency of wave V. It is preferable to the moving dipole because it uses only half the number of parameters while producing approximately the same goodness of fit.

- Rotating dipoles corresponding to binaural stimuli are localized in the sagittal plane.

- For monaural stimuli, the locations of the rotating dipoles lie in the hemisphere contralateral to the side of stimulation.

- The moments of the rotating dipoles strongly correlate with the lateralization of the stimuli caused by ITD *and* ILD, but they do not correlate with ITD or ILD alone.

- The physiologically motivated model using two hemispherically symmetric constrained dipoles is also capable of explaining the data.

- The forward models of the rotating and the pair of fixed dipoles are very similar in regions where electrodes can be attached to the head. The larger differences at the inferior, cervical sites can hardly be utilized. Higher signal-to-noise ratios are required to resolve different sources within the brain stem.

Chapter 5

Comparison of binaural auditory brain stem responses and binaural difference potentials evoked by chirps and clicks[1]

Abstract

Rising chirps that compensate for the dispersion of the travelling wave on the basilar membrane evoke larger monaural brain stem responses than clicks (Dau et al., 2000). In order to test if a similar effect applies for the early processing stages of binaural information, monaurally and binaurally evoked auditory brain stem responses were recorded for clicks and chirps for levels from 10 to 60 dB nHL in steps of 10 dB. 10000 sweeps were collected for every stimulus condition from ten normal hearing subjects. Wave V amplitudes are significantly larger for chirps than for clicks for all conditions. The amplitude of the binaural difference potential, DP1-DN1, is significantly larger for chirps at the levels 30 and 40 dB nHL. Both the binaurally evoked potential and the binaural difference potential exhibit steeper growth functions for chirps than for clicks for levels up to 40 dB nHL. For higher stimulation levels the chirp responses saturate approaching the click evoked amplitude. For both stimuli the latency of DP1 is shorter than the latency of the binaural wave

---

[1]A modified version of this chapter was published in Hear. Res. 169 (1-2), 85–96: Riedel and Kollmeier (2002b): "Comparison of binaural auditory brain stem responses and the binaural difference potential evoked by chirps and clicks"

V, which in turn is shorter than the latency of DN1. The amplitude ratio of the binaural difference potential to the binaural response is independent of stimulus level for clicks and chirps.

A possible interpretation is that with click stimulation predominantly binaural interaction from high frequency regions is seen which is compatible with a processing by contralateral inhibitory and ipsilateral excitatory (IE) cells. Contributions from low frequencies are neglegible since the responses from low frequencies are not synchronized for clicks.

The improved synchronisation at lower frequencies using chirp stimuli yields contributions from both, low and high frequency neurons enlarging the amplitudes of the binaural responses as well as the binaural difference potential. Since the constant amplitude ratio of the binaural difference potential to the binaural response makes contralateral and ipsilateral excitatory (EE) interaction improbable, binaural interaction at low frequencies is presumably also of the IE-type.

Another conclusion of this study is that the chirp stimuli employed here are better suited for auditory brain stem responses and binaural difference potentials than click stimuli since they exhibit higher amplitudes and a better signal-to-noise ratio.

5.1   Introduction

The properties of the traveling wave along the basilar membrane are such that the activation maximum for higher frequencies occurs earlier than that for lower frequencies (von Békésy, 1960; Greenwood, 1990). From this dispersion it follows that an acoustic click stimlus is no longer synchronized after passing the inner ear. To compensate for the dispersion on the basilar membrane a chirp with rising instantaneous frequency was developed by Dau et al. (2000). They demonstrated that a rising chirp stimulus evokes a larger response than an equally loud click for monaural stimulation. This effect can be well understood by the enhanced neural synchronisation obtained by the chirp especially for low frequencies, i.e., below 1 kHz.

Binaural interaction in auditory brain stem responses (ABRs) is commonly analyzed in terms of the binaural difference potential (BD), i.e., the difference between the evoked responses to binaural and summed monaural stimulation, symbolically BD = B − (L + R) (Levine, 1981; Furst et al., 1985; Ito et al., 1988; Jones and van der Poel, 1990; Furst et al., 1990; Levine and Davis, 1991; Jiang, 1996; Brantberg et al., 1999a,b; Riedel and Kollmeier, 2002a), see also chap. 3. The BD is thought to reflect the activity of neural units

responding specifically to binaural stimulation. The amplitude of the summed monaural potential is usually found to be larger than the binaural response, i.e., the BD has an inverted polarity compared to the binaural response. At least two mechanisms can be thought to cause this reduction: (i) Contralateral inhibitory and ipsilateral excitatory (IE) cells Goldberg and Brown (1969) in the superior olive (SO) exhibit a reduced response to binaural stimulation. (ii) Contralateral and ipsilateral excitatory (EE) cells are driven (near) to saturation by monaural stimulation and cannot double their response for binaural stimulation. For clicks, the amplitude of the most prominent peak pair DP1-DN1 is about a fifth of the amplitude of wave V for a wide range of stimulus levels (Levine, 1981). Because the noise variance of the BD is about three times the variance of the monaural response, the signal-to-noise ratio (SNR) of the BD is about an order of magnitude smaller than the SNR of the binaural response. Click evoked BD amplitudes barely exceed $0.2\,\mu\mathrm{V}$. Therefore, it is highly desirable to provide methods which augment the SNR of the BD. An increased dynamic range of the latter could expand the possible experimental setups used to study the correlation between spatial stimulation and the corresponding binaural difference potential.

In the present study it is investigated whether a larger binaural potential and a larger BD (with higher SNR) can be obtained with a chirp signal in comparison to the traditionally used clicks. Furthermore, the amplitude ratio of the binaural difference potential to the binaural response is analyzed as a function of stimulus level. This ratio allows to draw conclusions about the cell types involved in the generation of the BD.

5.2 Methods

5.2.1 Subjects

Ten subjects, 2 females and 8 males, aged $25 - 36$ years participated in the experiments. They were either paid or volunteers from the staff of the University of Oldenburg. They were classified as normal hearing by routine audiometry and had no history of audiological or neurological problems. The audiometric loss was less $10\,\mathrm{dB}$ for frequencies below $4\,\mathrm{kHz}$ and less than $15\,\mathrm{dB}$ for the higher frequencies.

5.2.2 Stimuli

Stimuli were generated digitally, downloaded to a DSP32C card in the host computer, and DA converted at a sampling rate of $50\,\mathrm{kHz}$. The click was a sequence of five con-

stant samples, and was converted to a rectangular voltage pulse of 0.1 ms duration. The spectrum of the chirp was approximated to the flat spectrum of the click by attenuating it at the lower frequencies. The flat spectrum chirp with the edge frequencies of 0.1 and 10 kHz had a duration of 10.32 ms (Dau et al., 2000). Acoustic waveforms and spectra as measured with a fast Fourier transfrom (FFT) analyzer (Stanford Research SR780) are shown in Fig.5.1. The time between two subsequent stimulus onsets was chosen to vary randomly and equally distributed between 62 and 72 ms yielding an average stimulation rate of approximately 15 Hz.

A 700 ms segment of the stimuli comprising 11 clicks or chirps was used to determine the thresholds in quiet. They were measured three times by all subjects with a 3-alternative forced-choice method in conjunction with a 2-down–1-up scheme for both ears and averaged over runs, subjects and ears. The threshold level – referred to as 0 dB normal hearing level (nHL) – corresponded to 40.5 dB peak equivalent sound pressure level (peSPL) for the click and 37 dB peSPL for the chirp,[2]. The standard deviation of the individual thresholds from the averaged thresholds was 2.8 dB for the clicks and 3.6 dB for the chirps.

5.2.3 Recordings

Ag/AgCl-electrodes were used for the recordings. Three active channels were placed at the left (A1) and right (A2) mastoid as well as one cm below the inion (Iz). The common reference electrode was placed at the vertex (Cz), the ground electrode at the forehead (Fpz). Electrode labels are according to the 10-20-system (Jasper, 1957).

Electrode impedances were measured at a test signal frequency of 30 Hz and brought well below 5 kΩ, common values were 2-3 kΩ. Since DC recordings were performed, a second criterion for a good contact between electrodes and skin beside low impedance was the voltage drift seen in the raw EEG signal. Electrode contact was improved until any drift vanished.

During the ABR recordings, subjects lay in a sound insulated and electrically shielded room. They were instructed to relax and lie as comfortably as possible. ABRs were recorded with a DC-coupled differential amplifier (Synamps 5803). Inside the shielded room the EEG was preamplified by a factor 150, further amplified by the main amplifier by a factor 33 resulting in a total amplification of 74 dB. The voltage resolution was

---

[2]A sinusoid of frequency 1 kHz with the same peak-to-peak-amplitude as the chirp showed 37 dB SPL in a Brüel & Kjær (B&K) amplifier type 2610. The calibration was performed using a half inch microphone (B&K 4157) with an artificial ear (1.29 cm^3) and a preamplifier (B&K 2669).

approximately $16.8\,\mathrm{nV/bit}$. The sweeps were filtered by an analog antialiasing-lowpass with an edge frequency of $2\,\mathrm{kHz}$, digitized with $10\,\mathrm{kHz}$ samplingrate and 16 bit resolution, and stored to hard disk.

The artifact level was set to $\pm 500\,\mu\mathrm{V}$, since filtering, artifact analysis and averaging was done offline. The clipping level of the DA-converters was $\pm 550\,\mu\mathrm{V}$. The recording interval comprised 500 samples in the time interval from -15 to $35\,\mathrm{ms}$ relative to stimulus onset.

Left, right and binaural stimuli were presented randomly on a sweep-by-sweep basis. One run consisted of 15000 stimuli, 5000 of each type, and lasted approximately 17 minutes. Two runs were performed for both stimuli (clicks and chirps) and 6 levels, 10 to $60\,\mathrm{dB}\,\mathrm{nHL}$ in steps of $10\,\mathrm{dB}$. The recordings were subdivided into 4 sessions with 6 runs each. Every session started with the highest level $60\,\mathrm{dB}\,\mathrm{nHL}$ and successively lowered the level until $10\,\mathrm{dB}\,\mathrm{nHL}$ was reached in the last run.

5.2.4 Data analysis

Before averaging the single sweeps were filtered with a linear phase FIR bandpass filter with 200 taps and the edge frequencies 100 and $1500\,\mathrm{Hz}$ (Granzow et al., 2001). An iterated weighted average of the filtered sweeps was computed for all subjects and stimulus conditions. The residual noise of the averages was computed as the standard error of the single sweeps σ, see chap. 2 (Riedel et al., 2001).

For both stimulus types and all levels, the binaural difference potential was computed channel-wise and sample by sample, symbolically: BD = B – (L + R). All BDs were computed from response triplets which were recorded quasi-simultaneously. This avoids artifacts in the BD components due to long term changes of the recording condition or subject's state. The residual noise of the BD was estimated as the square root of the summed variances of the three measurements: $\sigma_{\mathrm{BD}} = (\sigma_{\mathrm{B}}^2 + \sigma_{\mathrm{L}}^2 + \sigma_{\mathrm{R}}^2)^{1/2}$.

To increase the accuracy of amplitude and latency measurements, data were interpolated by a factor of 10, i.e., they were upsampled to convert the sampling rate from 10 to $100\,\mathrm{kHz}$. This was accomplished by zero-padding in the spectral domain which in the time domain corresponds to a convolution with a sinc-function. Since the original analog signal was band-limited to frequencies below $2\,\mathrm{kHz}$ a near-perfect interpolation was possible.

Peaks in the interpolated signal were identified by a sign change in its derivative. For baseline-to-peak-measurements peaks with voltages V_{bp} smaller than 3σ (99.7% confidence level for Gaussian measurement errors) were not regarded as significant and hence were

discarded. For peak-to-peak-measurements peaks with voltages V_{pp} greater than $\sqrt{2} \cdot 3\sigma$ were accepted. The additional factor of $\sqrt{2}$ is due to the fact that the variances of both peaks in the pair add up. Latency errors were estimated from the amplitude errors and the curvature of the peaks according to Hoth (1986).

Amplitude and latency of wave V was analyzed for each stimulus condition. The amplitude was measured baseline-to-peak because the peak-to-peak measurement from wave V to wave VI' (the negative trough following wave VI) would yield erroneous amplitudes for three subjects exhibiting muscular artifacts in the auricular channels at the latency of wave VI'.

The first main component of the BD is the negative wave DN1 preceeded by a smaller positive wave labelled DP1. The nomenclature introduced by Ito et al. (1988) is adopted here (see their Fig. 1). DN1 corresponds to the β-wave described by Levine (1981). BD amplitudes were measured peak-to-peak from DP1 to DN1. They are not contaminated by any muscle artifacts since the latencies of DP1 and DN1 are close to the latency of wave V. Latencies of the larger component DN1 were analyzed. Signed Wilcoxon rank tests were performed to reveal the significance of amplitude differences between clicks and chirps.

5.3   Results

5.3.1   Binaural potentials

Fig. 5.2 shows binaural potentials for the click (left) and the chirp (right) for levels from 10 to 60 dB nHL for one subject. For all three channels and all levels, the amplitude of wave V (A_V) is larger for the chirp than for the click. Moreover, for this subject the maximal response is found at 40 dB nHL for chirp stimulation. At 10 dB nHL there is still a visible response for the chirp, but not for the click. Relative to stimulus onset, the latency for the chirp response is about 10 ms larger compared to the click. At 60 dB nHL a positive deflection at about 8 ms latency is seen in the chirp response.

In the upper two rows of Fig. 5.3 the amplitude of wave V to diotic clicks and chirps is compared for channel A1 and all subjects. A_V for clicks increases with stimulation level, i.e., exhibits monotonic growth functions. In contrast, A_Vs for chirps are generally non-monotonic functions. They show a steeper increase in response than clicks up to 40 dB nHL. For higher levels there are interindividual differences: for the five subjects in the first row A_V for chirps decreases with increasing level, i.e., the maximal amplitude

is reached at 40 dB nHL. For the subjects in the second row chirp amplitudes level off at 40 dB nHL (kw, ag, jd) or increase even further with level (ow, nw). The small intraindividual standard errors (error bars) show that at 40 dB nHL A_V is larger for the chirp than for the click, for all subjects. In the bottom row grand average data over subjects are shown for the three channels measured as well as the mean over channels. The growth functions are very similar between channels. On the average over subjects, the chirp amplitudes reach their maximum at 40 dB nHL and level off for higher stimulation levels whereas the growth functions for clicks are monotonically increasing. To reveal the differences between amplitudes evoked by clicks and chirps signed Wilcoxon rank tests were performed for all channels and pairs of levels. The results of these tests are summarized in the left half of Tab. 5.1. For levels between 20 and 40 dB nHL wave V amplitudes for chirps are larger than for clicks with high significance ($\alpha < 0.001$). For 50 dB nHL the differences are only significant at the level $\alpha < 0.05$. At 60 dB nHL only the differences for channel A2 and the mean over channels show a significant difference ($\alpha < 0.05$).

	binaural wave V						BD wave DP1-DN1					
level	10	20	30	40	50	60	10	20	30	40	50	60
A1	**	***	***	***	*	ns	ns	ns	*	*	ns	ns
A2	**	***	***	***	*	*	ns	ns	*	ns	ns	ns
IZ	**	***	***	***	*	ns	ns	ns	**	*	ns	*
mean	**	***	***	***	*	*	ns	ns	**	**	ns	ns

Tab. 5.1: *Significance of the differences between amplitudes of wave V and of DP1-DN1 for clicks versus chirps as evinced by signed Wilcoxon rank tests. Stimulus level in dB nHL. Three significance levels were tested: $\alpha < 0.05$ (*), $\alpha < 0.01$ (**) and $\alpha < 0.001$ (***), 'ns' means not significant.*

5.3.2 Binaural difference potentials

Fig. 5.4 shows the derivation of the binaural difference potential (BD) from the binaural and monaural responses at 50 dB nHL for one subject. The error bars show ± 3 standard errors corresponding to a 99.7% confidence interval for Gaussian measurement errors. The binaural responses (B) have slightly shorter latencies and smaller amplitudes than the sum of the monaural responses (L+R). This results in significant peaks in the BD, namely DP1, DN1 and DP2. DP1 and DN1 are associated with wave V of the binaural potential, DP2 with wave VI. In this example, the chirp BD has a larger amplitude as the click BD.

In Fig. 5.5 the dependence of the BD on stimulus type and level is depicted for one subject and all channels. Triangles are drawn for peaks whose peak-to-peak voltage V_{pp} exceeds $\sqrt{2} \cdot 3\sigma$. Only peaks satisfying this criterion are considered significant BD peaks. At the same stimulation level the chirp BDs are larger than the click BDs. Analogous to the chirp evoked binaural potential the growth function of the chirp BD is also steeper than for clicks. In this example the maximal peak-to-peak amplitude DP1-DN1 for chirps is found at 40 dB nHL.

Fig. 5.6 summarizes the peak-to-peak amplitude $A_{DP1-DN1}$ for all subjects and the mean over subjects as a function of stimulus type and level. Amplitudes of chirp and click BDs are marked by filled upward and open downward triangles, respectively. If for a certain stimulus and level $A_{DP1-DN1}$ failed to reach significance, no data are drawn. In analogy to Fig. 5.3 the first two rows show data for single subjects. There is a large variation in BD among subjects. However, similar to the binaural potentials, the chirp evoked BDs are generally larger than click evoked BDs. For one subject (ow) no BD to clicks could be detected at any stimulation level. For 5 subjects, the chirp BD is maximal at 40 dB nHL, for the other subjects chirp BDs level off or increase further. The grand mean data over subjects (see bottom row of Fig. 5.6) show that the dependence of $A_{DP1-DN1}$ on stimulus type and level is similar for all three channels measured: chirp BDs grow faster with level than click BDs and level off at 40 dB nHL.

Signed Wilcoxon rank tests were performed for all channels and pairs of levels to analyze the differences between BDs evoked by chirps and clicks. The right side of Tab. 5.1 shows that, with the exception of channel A2 at 40 dB nHL, chirp BDs are significantly larger than click BDs for 30 and 40 dB nHL. At 50 and 60 dB nHL, except for channel IZ at 60 dB nHL, no significant difference between chirp and click amplitudes is found. However, for many subjects there was no significant BD at some levels and channels either for chirps or for clicks, e.g., for subject ow there were only significant peaks in the BD for chirps. These unpaired data did not enter the above tests. If non-significant and undetectable peaks are considered to have amplitude zero, at 40 dB nHL significant differences result for all channels, and at 50 and 60 dB nHL no significant differences are found. This may still be due to the small number of subjects (N=10), but it shows that, on average, BDs to clicks grow for levels from 40 to 60 dB nHL whereas BDs to chirps saturate at 40 dB nHL. This results in a maximal difference at 40 dB nHL.

In Fig. 5.7 the latencies of wave V (t_V) are compared with the latencies of BD waves DP1 (t_{DP1}) and DN1 (t_{DN1}) as a function of the stimulus level. For both clicks and chirps, the latencies are ordered: $t_{DP1} < t_V < t_{DN1}$. The first two rows of Fig. 5.7 show single

subject data. The error bars denote the intraindividual latency errors estimated from the intraindividual amplitude errors and the peak curvature according to Hoth (1986). In the bottom row mean latencies, averaged over subjects, are shown for the three channels as well as the average over channels. Here, error bars indicate the interindividual standard deviations of the latencies. Averaged over levels, subjects and channels, the latency difference for wave V between chirps and clicks amounts to 9.63 ± 0.14 ms. The mean time differences between DP1 and wave V amount to 0.44 ± 0.12 ms for clicks and 0.48 ± 0.08 ms for chirps. Compared to wave V, the mean delay of BD wave DN1 is 0.39 ± 0.08 ms for clicks and 0.43 ± 0.17 ms for chirps. These data indicate a strong dependence of BD latencies on wave V latencies.

In Fig. 5.8 the amplitude ratio of the BD and the binaural response is shown. For both stimuli and each level averages and standard deviations over 10 subjects and 3 channels are drawn. Except for 10 dB nHL stimulus level the amplitude ratio is nearly constant. Averaged over stimulus levels from 20 to 60 dB nHL $A_{\mathrm{DP1-DN1}}/A_{\mathrm{V}}$ is 0.28 for clicks and 0.27 for chirps. The literature value of one fifth is obtained by measuring wave V peak-to-peak to VI', i.e., to the trough following wave VI: $A_{\mathrm{DP1-DN1}}/A_{\mathrm{V-VI'}}$ amounts to 0.20 for clicks and 0.19 for chirps. No amplitude measures were taken into account for the three subjects showing muscle artifacts in channels A1 and A2. At 10 dB nHL no amplitude ratio for clicks is given because the SNR of the BD was smaller than the threefold standard error for all subjects. For chirps a higher amplitude ratio than for the other stimulus levels is found. However, the measurement of the amplitude ratio is more uncertain at low stimulus levels due to the small amplitudes of both the binaural response and the BD.
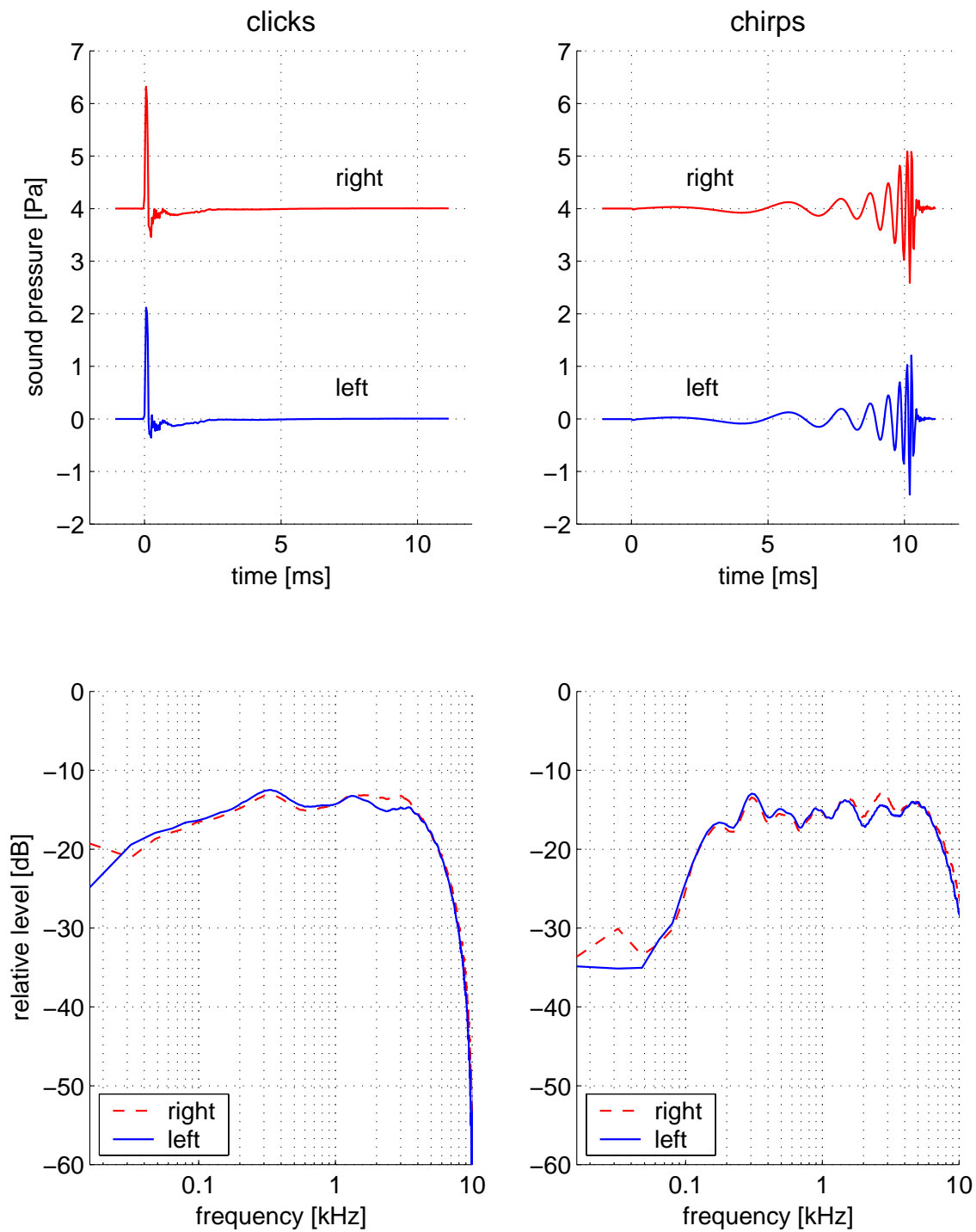
Fig. 5.1: **Top row:** *Acoustic waveforms of the click (left panel) and the chirp (right panel) measured at 60 dB nHL, corresponding to 100.5 dB peSPL for the click and 97 dB peSPL for the chirp. Right stimuli are plotted with an offset of 4 Pa.* **Bottom row:** *Acoustic spectra of the stimuli using 625 FFT bins in steps of 80 Hz.*
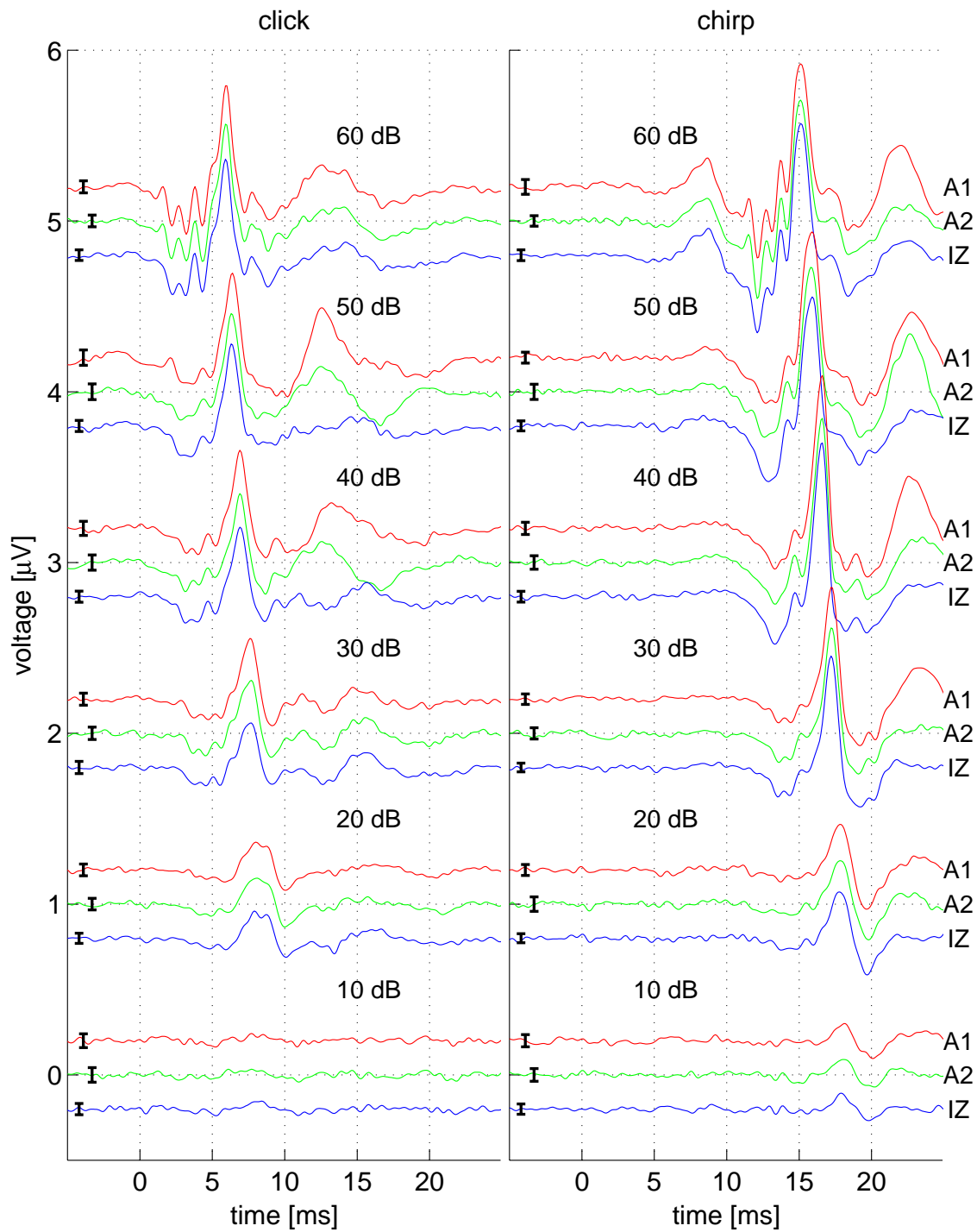
Fig. 5.2: Comparison of ABRs evoked by diotic clicks and chirps for subject rh. Channels A1, A2 and IZ at levels from 10 to 60 dB nHL are shown. Plot offset between levels is 1 μV, between channels 0.2 μV. Error bars indicate ±3 standard errors of the mean (±3σ).

Fig. 5.3: *Amplitudes of wave V as a function of the stimulus level, thick lines are for chirps,
thin lines for clicks.* **Top and middle row:** *Data for single subjects from channel A1
with intraindividual standard error σ.* **Bottom row:** *Data averaged over subjects with
interindividual standard deviations, channels A1, A2, IZ and mean over channels.*
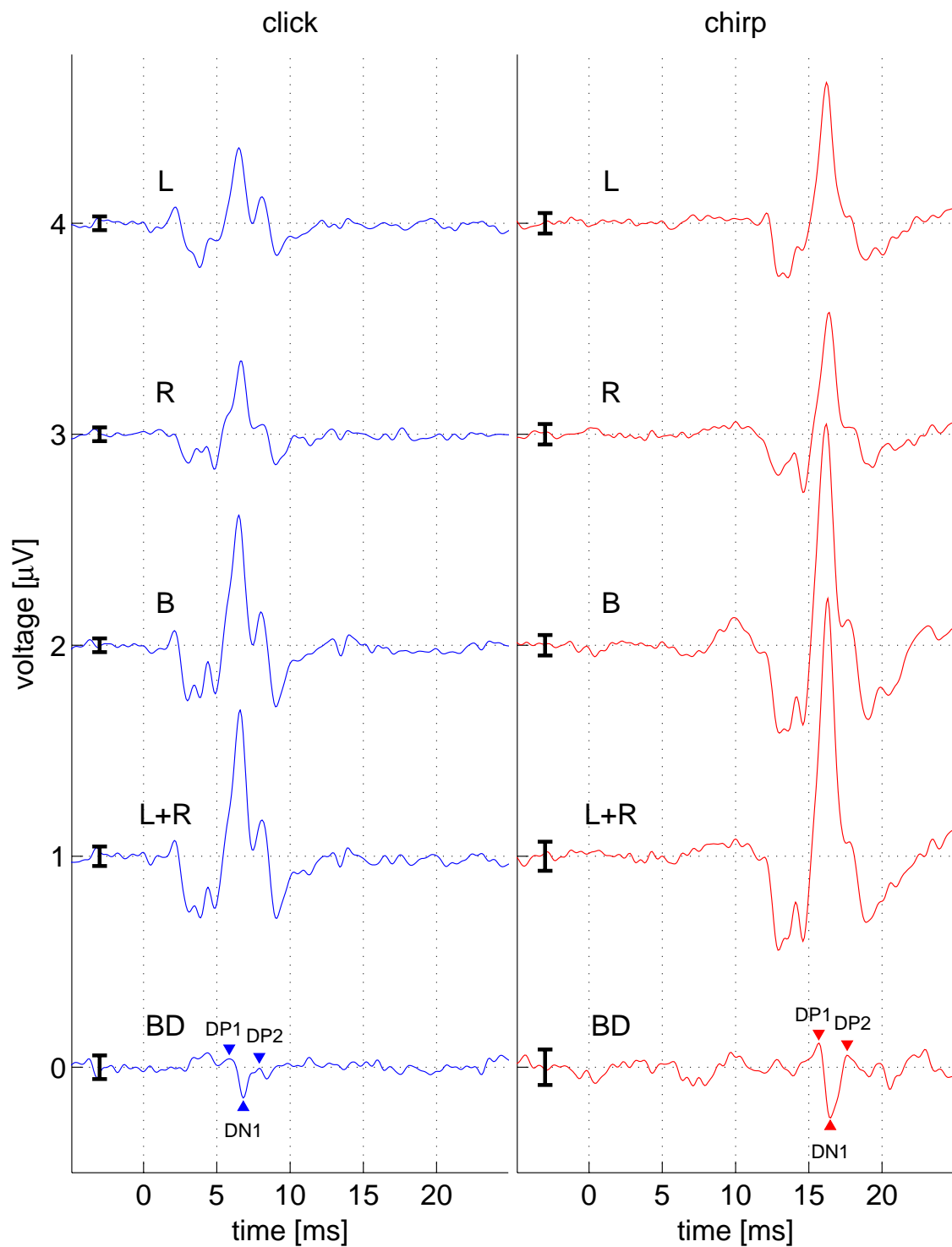
Fig. 5.4: Derivation of the binaural difference potential. Data from subject jo for $50\,dB\,nHL$. **Left column:** clicks, **right column:** chirps. **Top row:** ABR to monaural left stimulation, **second row:** ABR to monaural right stimulation, **third row:** binaural diotic response, **fourth row:** sum of the monaural responses, **bottom row:** binaural difference potential $BD = B - (L + R)$. Error bars indicate ± 3 standard errors ($\pm 3\sigma$). Responses are plotted with an offset of $1\,\mu V$.
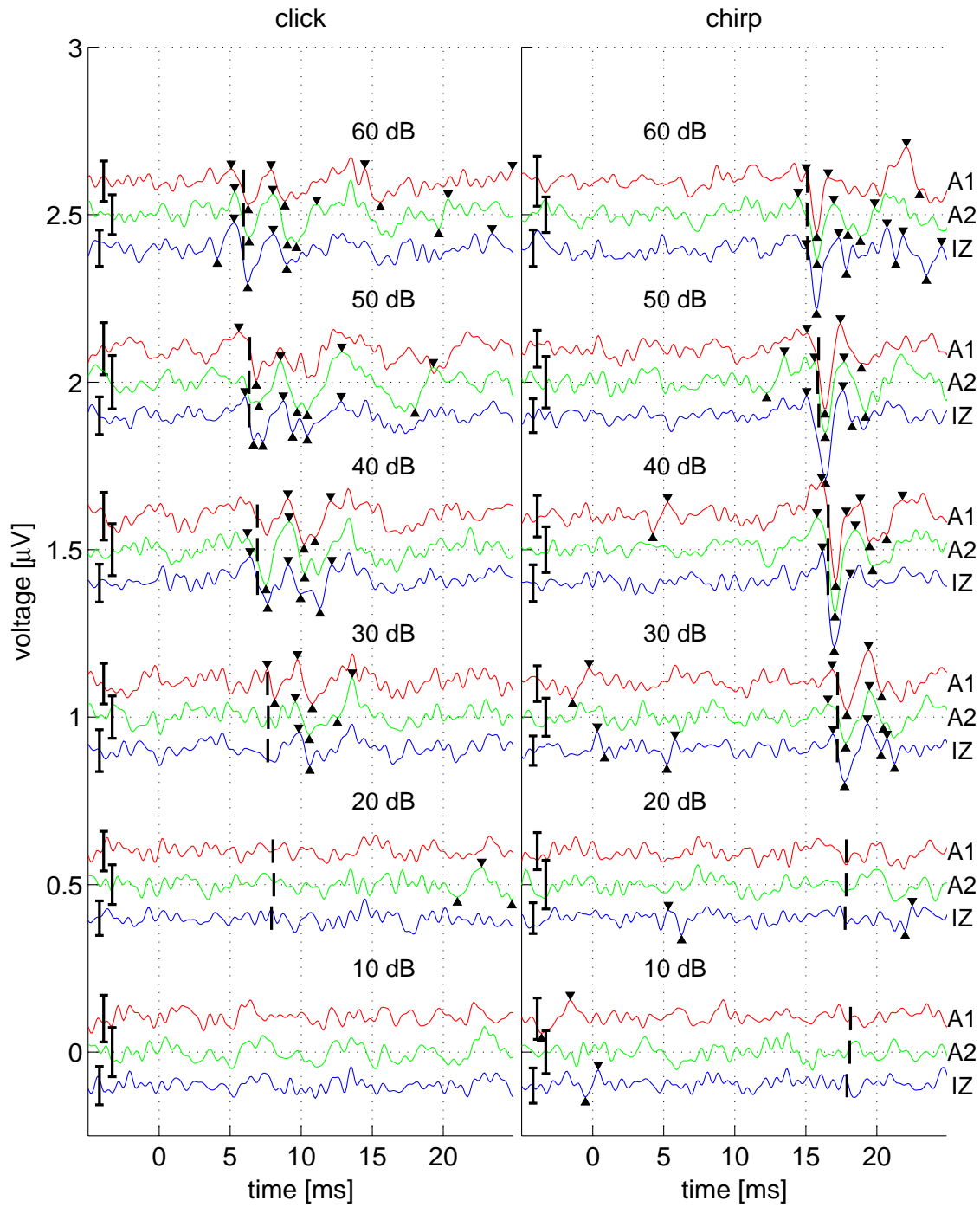
Fig. 5.5: Comparison of the binaural difference potential BD evoked by diotic clicks and chirps for subject rh. Channels A1, A2 and IZ at levels from 10 to $60\,dB\,nHL$ are shown. Plot offset between levels is $1\,\mu V$, between channels $0.2\,\mu V$. Error bars indicate ± 3 standard errors ($\pm 3\sigma$). Triangles indicate peak pairs with peak-to-peak voltages $V_{pp} \geq \sqrt{2}\cdot 3\sigma$. The vertical bars mark the latency of wave V for binaural stimulation.
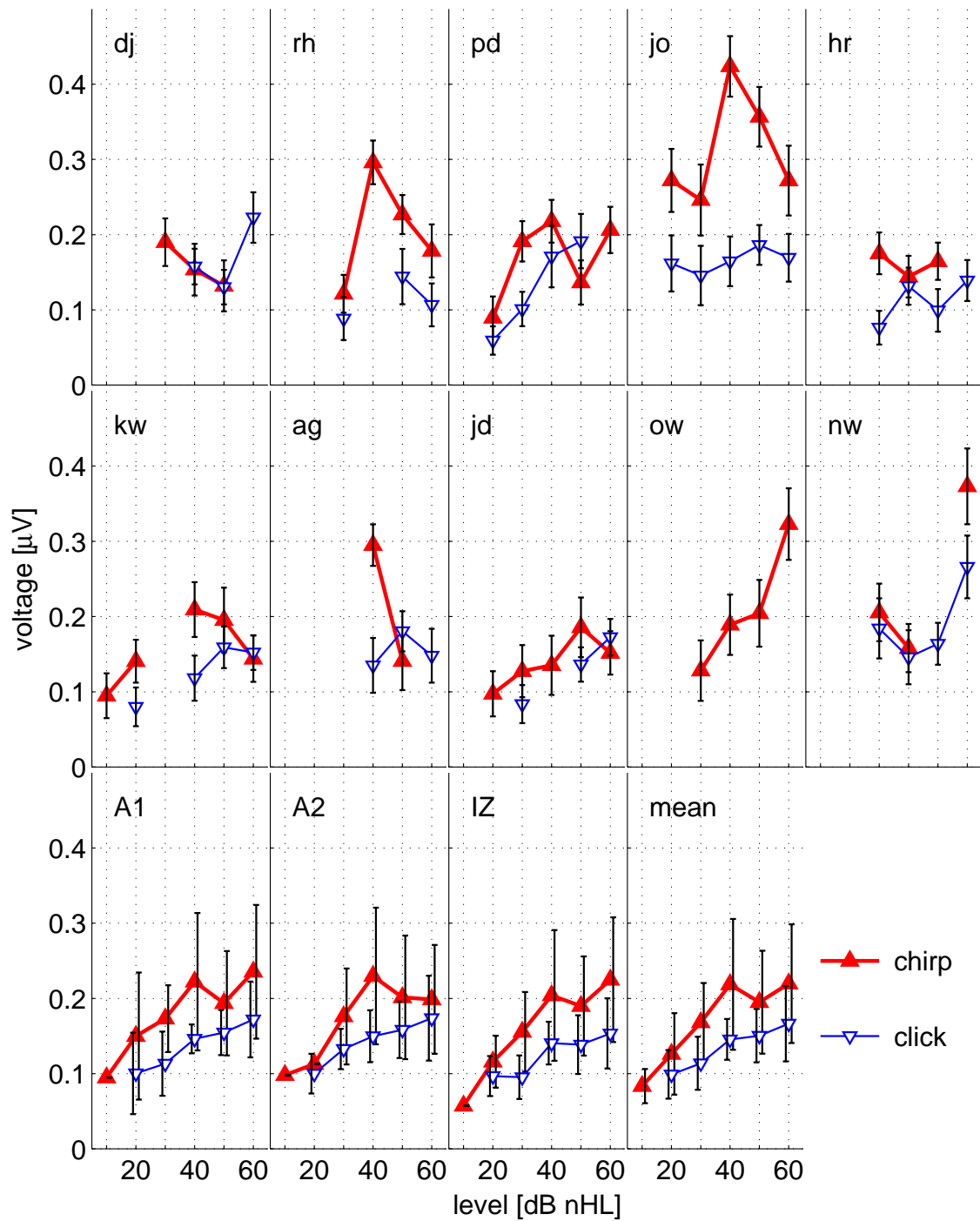
Fig. 5.6: BD amplitude $A_{\mathrm{DP1-DN1}}$ as a function of the stimulus level. Thick lines with filled triangles indicate chirp amplitudes, thin lines with open triangles click amplitudes. **Top and middle row:** Data for single subjects from channel A1 with intraindividual standard errors σ. **Bottom row:** Data averaged over subjects with interindividual standard deviations, channels A1, A2, IZ and mean over channels.
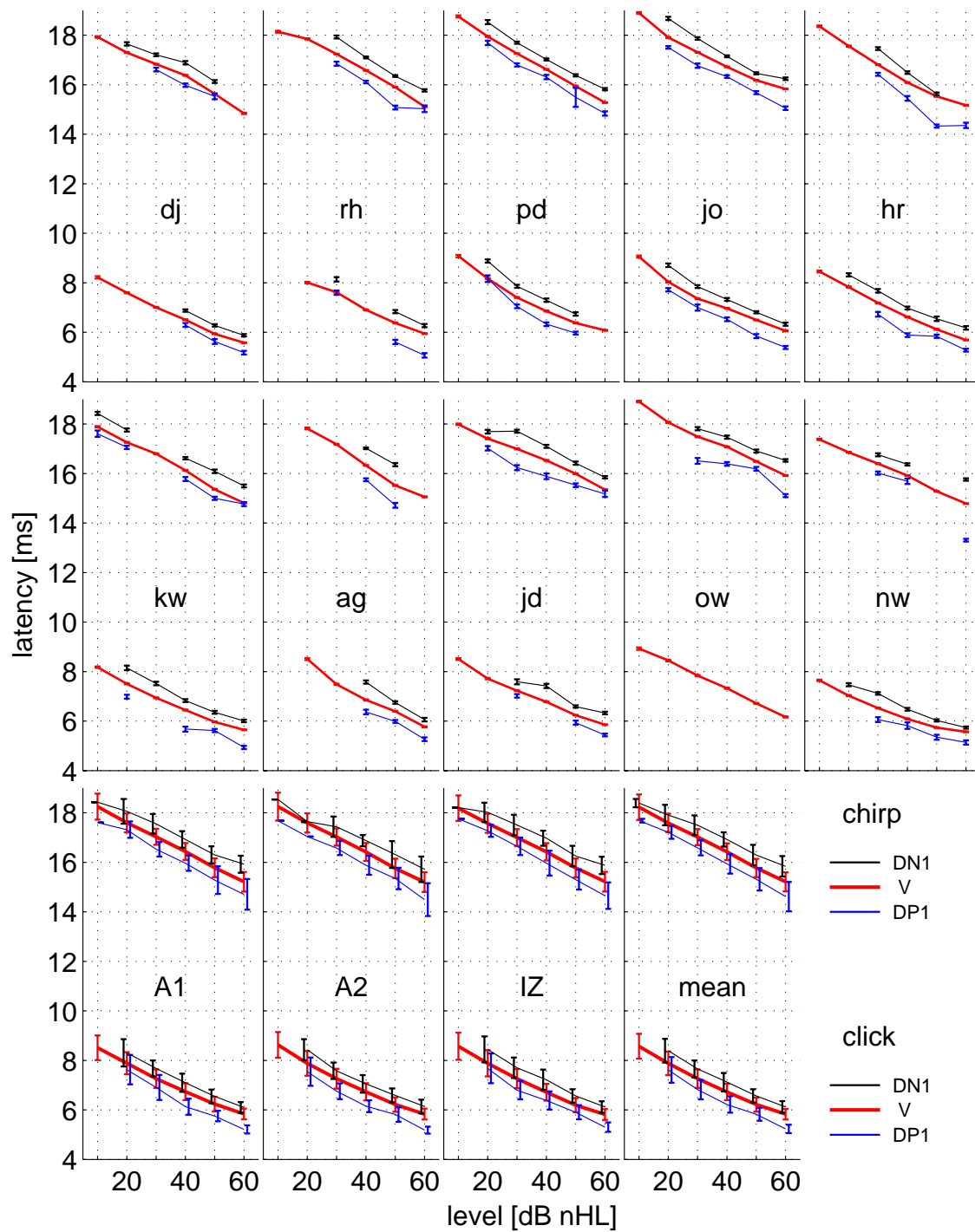
Fig. 5.7: Latencies of wave V (thick lines) and BD waves DP1 and DN1 (thin lines) as a function of the stimulus level. **Top and middle row:** Data for single subjects from channel A1 with intraindividual standard errors. **Bottom row:** Data averaged over subjects with interindividual standard deviations, channels A1, A2, IZ and mean over channels.
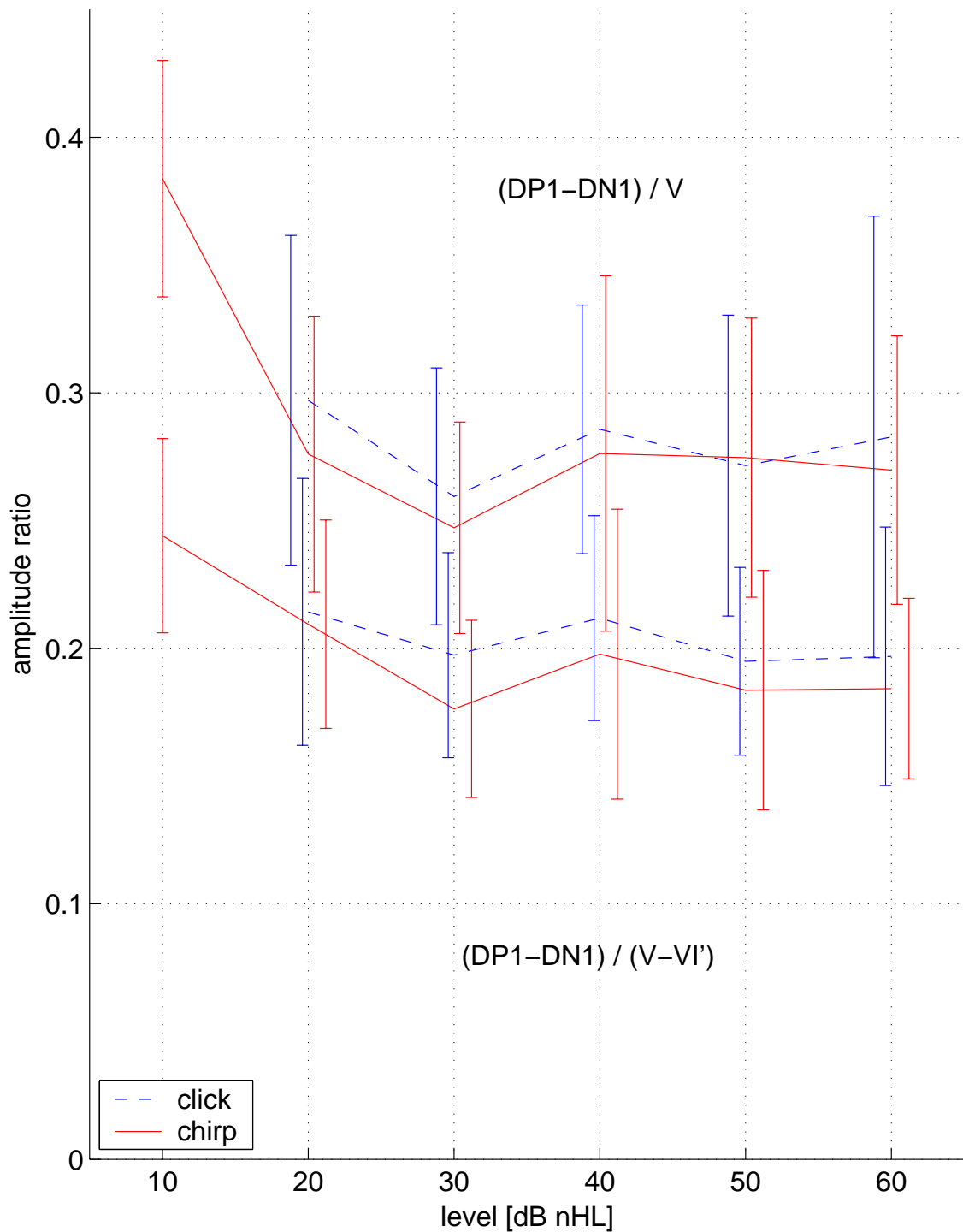
Fig. 5.8: Amplitude ratio of the BD to the binaural response, mean data and standard deviations over 10 subjects and 3 channels. Dashed lines are for clicks, solid lines for chirps. In the upper two curves wave V amplitude is measured baseline-to-peak, in the lower curves peak-to-peak from wave V to VI' (the trough following wave VI).

5.4 Discussion

Auditory brain stem responses and binaural difference potentials were measured using interleaved recording of left, right and binaural responses and averaging a large number of sweeps, in order to obtain a high SNR, for clicks and flat spectrum chirps. The results clearly demonstrate that chirps evoke larger binaural responses as well as larger binaural difference potentials than clicks. Hence, the improved neural synchronization obtained with the monaural chirps (Dau et al., 2000) is also found for binaural chirps and propagates into an enlarged BD. This qualifies the chirp stimulus for further research work on binaural interaction with evoked potentials. However, the level range over which the chirp stimulus provides advantages over the click stimulus is limited to low and intermediate levels.

In the chirp response, a small positive deflection in the latency range from 7 to 10 ms was observed for all subjects for a stimulus level of 60 dB nHL (see, e.g., Fig. 5.2). This deflection is a consequence of a slight discontinuity of the chirp at time = 0 ms (Fig. 1, top right) caused by a shortcoming in the stimulus generation program. The magnitude of this onset transient was about -38 dB compared to the maximum amplitude of the stimulus. A control measurement using a smoothed version of the chirp was performed at levels 40 and 60 dB nHL for three subjects. No differences in the responses were detected for 40 dB nHL. However, at 60 dB nHL the early positive deflection disappeared, indicating that it was a response to the onset discontinuity. The onset transient did not mask the chirp response. Wave V amplitudes of the smoothed chirp tend to be slightly smaller compared to those obtained with the original chirp. This implies that the nonmonotonicity of the chirp response could be even more pronounced, i.e., exhibiting a maximum at 40 dB nHL rather than levelling off.

The observation that the gain is largest at medium levels is consistent with the results of Dau et al. (2000), see their Fig. 3. Their average monaural data showed a significantly larger wave V amplitude for the chirp than for the click at all levels. But only at medium levels (40 dB nHL) the difference was highly significant.

From animal studies it is known that the superior olive (SO) is the first stage at which binaural interaction occurs (e.g., Irvine, 1992). Goldberg and Brown (1969) classified the SO cells by the type of their input: IE-cells predominantly found in the lateral SO receive contralateral inhibitory input via the medial nucleus of the trapezoid body (MNTB) and excitatory input from the ipsilateral cochlear nucleus (CN), they are believed to code the interaural level difference (ILD) at high frequencies. EE-cells in the medial part of the SO obtain excitatory inputs from both CNs and are believed to code the interaural time

difference (ITD) at low frequencies.

In evoked response studies the sum of the monaural responses is always found to be larger than the binaural response. According to Gaumond and Psaltikidou (1991) this can be explained by two mechanisms: (i) contralateral inhibition in IE-cells reduces the binaural response. (ii) EE-cells get already saturated by monaural stimulation and can therefore not reach the summed firing rate of both monaural responses. Gaumond and Psaltikidou (1991) analyzed an EE- and an IE-model of binaural interaction. The constancy of the amplitude ratio $A_{\mathrm{DP1-DN1}}/A_{\mathrm{V}}$ across levels is easily explained by the IE-model. Within the EE-model special assumptions about the compressive non-linearity describing the saturation are needed to preserve the constant amplitude ratio. Therefore Gaumond and Psaltikidou (1991) conclude that IE-cells play a larger role in BD generation.

From studies involving models of cochlear mechanics and evaluating the dispersion along the basilar membrane (e.g., de Boer, 1980; Dau et al., 2000) it is known that a click stimulus exhibits a high synchronization in hair cell deflection in the basal portion of the cochlea (corresponding to high frequencies) whereas the response to low frequencies shows a considerable temporal smearing of the deflection along different places on the basilar membrane. Hence, ABRs with clicks mainly reflect the response to high frequencies. Therefore it can be presumed that the BD measured in response to clicks represents to a larger portion the activity of IE-cells processing ILDs than of EE-cells processing ITDs. Consequently, contributions from EE-cells due to saturation are not seen in the evoked response because the responses from low frequencies are not synchronized for clicks. This conclusion is in accordance with the results from Levine and Davis (1991). They measured the BD to clicks in the presence of a highpass noise masker and showed by means of the derived response technique that the BD is principally due to the high frequency components of the click.

The improved synchronisation at lower frequencies using chirp stimuli yields contributions from both low and high frequency neurons augmenting the ABR and the BD. From comparing the relation of the binaural interaction and the monaural response between the click and the chirp stimulus it thus can be assessed if the lower frequencies contribute to the BD in a different way than the high frequencies. If one assumes that IE-cells primarily produce the BD at high frequencies and EE-cells are mainly active at low frequencies, this comparison can also be used to assess the relative contribution of both types of cell populations. Any contribution to the BD from EE-cells must originate from a non-linearity after summation of the activities from the left and right side. Without such a non-linearity the contribution from EE-cells would be zero. A positive contribution from EE-cells to

the BD, i.e., B > L + R, can only occur with an expansive non-linearity. However, an expansion after the summation is not very feasible physiologically. With the more plausible assumption of a compressive non-linearity describing the saturation (Gaumond and Psaltikidou, 1991), a negative contribution of the EE-cells to the BD, i.e. B < L + R, is obtained. Within this view both IE- and EE-cells would contribute to the BD in the same way. However, both types of non-linearity after the EE-cells would alter the amplitude ratio $A_{\mathrm{DP1-DN1}}/A_{\mathrm{V}}$ as a function of the stimulus level contradicting the experimental findings. A compressive non-linearity would result in a smaller amplitude ratio at higher stimulus levels, an expansive non-linearity in a larger amplitude ratio. A possible interpretation is that binaural interaction in the ABR at low frequencies does not differ in principle from the one at high frequencies and is therefore predominantly of the IE-type.

Another conclusion from the current data is that the rising frequency chirp constitutes a stimulus which allows to analyze the binaural difference potential with higher SNR and larger dynamic range than the conventional click stimulus. For example, the frequency specificity of the BD can be investigated by delivering chirps with different frequency contents to the two ears, e.g., a low frequency chirp to the left and a high frequency chirp to the right ear. Another useful application in future research will be the analysis of the correlation between the spatial position of a stimulus and its corresponding binaural difference potential.

Chapter 6

Summary and outlook

The analysis of evoked potentials elicited by spatial stimuli commenced with a methodological investigation of averaging methods that are used to compute the evoked potential from raw EEG data. In chapter 2 it was shown that the incorporation of single sweep information provides far more reliable estimates of the residual noise and the SNR of an evoked potential. This result was supported by the comparison of the variances of the average-based and the single-sweep-based noise estimate given in appendix A. Using the superior noise estimate, different averaging methods to calculate the evoked potential were compared. A new averaging method, termed iterated weighted averaging, was introduced. As validated by the simulation study, it allows for the best reconstruction of the signal and the most accurate estimation of the residual noise. Additionally, it has the advantage that an a-priori artifact threshold has not to be determined. Even though the hard disk space needed to record the raw data exceeds the space needed for the online average by a factor given by the number of sweeps, the benefits of using single-sweep information by far outweigh the inconvenience of large data files.

In the third chapter the auditory brain stem response (ABR) and the binaural difference potential (BD) were analyzed in dependence on interaural time and level difference. The amplitude of Jewett wave V and the amplitude difference of peaks DP1 and DN1 of the BD were found to correlate with the lateralization of the stimuli. Largest amplitudes were observed for centrally perceived stimuli, smallest amplitudes for the most lateral stimuli, respectively. It was also shown that the processing of ITD and ILD is not independent indicating a representation of lateralization rather than a representation of interaural cues within the brain stem.

Chapter 4 concentrated on dipole source analysis of multi-channel ABRs. Important

methodological improvements were possible by accounting for the noise covariance matrix derived from raw data. The inverse problem was solved by means of generalized maximum likelihood estimation. Additionally, confidence regions for dipole parameters could be specified as well as the goodness-of-fit based on a chi-square test. As yet the correct determination of fit quality is not possible in commercial software programs, e.g., ASA (Advanced Source Analysis) and BESA (Brain Electrical Source Analysis), because they do not incorporate the noise information in their strategies to solve the inverse problem.

Three alternative dipole models to explain the 2-ms interval centered around the binaurally evoked wave V for 15 stimulus conditions were presented. The rotating dipole model was preferred over the moving dipole model because it achieves a similar fit quality with fewer parameters. The physiologically motivated model using two constrained fixed dipoles is also capable of explaining the data, but results in a worse fit quality. In addition, the separation of the dipoles had to be set manually since the fit algorithm converged to solutions with nearly identical source locations. Furthermore, although the amplitudes of the sources looked reasonable, the physiological plausibility of the model remained questionable because for some stimulus conditions the contralateral dipole exhibited a smaller latency than the ipsilateral dipole. The comparison of the forward models of the rotating and the pair of fixed dipoles exemplified the ambiguity of the inverse problem. The models were scarcely distinguishable considering the residual noise of data. Poor spatial sampling played a minor role because deep sources were considered. However, the rotating dipole model was best suited to demonstrate the striking correlation between dipole parameters and psychophysical lateralization. For monaural clicks dipole locations contralateral to the side of stimulation were found while for binaural stimulation the dipole locations resided in the sagittal plane. This proves that nearby brain stem nuclei can be distinguished by dipole source analysis. The dipole moment trajectories did not correlate with a single interaural cue (ITD or ILD) alone, but with the psychophysical lateralization caused by the combination of both cues indicating an early extraction of the lateralization of a sound source. Using the mismatch negativity (MMN) paradigm, Schröger (1996) analyzed tonal stimuli with ITDs and ILDs. The MMN was computed by subtracting the response to the frequent diotic standard stimulus from the response to a rare dichotic deviant stimulus. Schröger et al. (1997) found an enlarged MMN to an antagonistic combination of ITD and ILD compared to the MMN evoked by a single deviant, i.e., ITD or ILD. They concluded that at least partly separate neural codes representing ITDs and ILDs exist at a cortical level of auditory information processing. At a first sight this finding contradicts with the current result of a trading relation of ITD and ILD at the

brain stem level. However, although it is reasonable to assume that information about the location of an acoustic stimulus is extracted at an early stage of the auditory system, this does not necessarily imply that the information about the ITD and the ILD is completely discarded. Instead, it may be used to create an exact internal representation of auditory objects at a higher stage. At the brain stem level, a rough object representation may be present where the trading of ITD and ILD already provides a decent estimate of the object's location.

In chapter 5 binaural ABRs and BDs to clicks and rising frequency chirps were compared in dependence on the stimulus level. The chirp is known to elicit larger monaural responses due to the improved neural synchronization compared to the click (Dau et al., 2000). This advantage directly transfers into larger binaural ABRs and larger BDs. The advantage of the chirp was maximal at intermediate levels (40 dB nHL). The amplitude ratio between the BD and binaural ABR was found to be approximately constant over stimulus level and stimulus type. A hypothesis explaining this finding is that binaural processing is predominantly determined by the inhibitory-excitatory type of interaction.

To further elucidate the dependence of evoked potentials on spatial stimuli a variety of improvements and refinements are feasible. First, the spherical head model employed in this work can be improved by taking the attenuation of the skull into account. A three-shell model is currently being implemented. To allow for absolute instead of relative localization of sources, the incorporation of anatomical information is necessary. Nowadays, commercial software programs offer the possibility to create realistic volume conductor models by the segmentation of MRI scans. The head model used to solve the inverse problem is defined by boundary elements describing realistically shaped shells that represent brain, skull and skalp.

Second, a variety of alternative methods to solve the inverse problem has been proposed in the literature. The so-called minimum norm localization uses a distributed-source model (Pascual-Marqui et al., 1994). Multiple signal classification (MUSIC) has been suggested by Mosher et al. (1992) and was improved by Koles and Soong (1998). Uhl et al. (2001) improved source localization by means of 'dynamical systems based modeling'. Reviews on recent developments in source localization are given for example by Koles (1998) and Pascual-Marqui (1999).

Third, stimuli with ILDs and ITDs independent from frequency do not occur in nature. In headphone experiments they generate a perceived auditory image inside the head, therefore these stimuli are characterized by producing a lateralization rather than a localization.

Free field EEG measurements using directional stimuli are difficult and inconvenient because the head position of the subjects would have to be fixed during the measurement. Hence, a contamination of evoked potentials with myogenic responses seems unavoidable. A promising alternative is the use of virtual acoustics. Before performing EEG measurements, head related transfer functions (HRTFs) of the desired spatial positions have to be recorded for all subjects. The convolution with the stimulus of interest produces an extracranial acoustical image. Experiments using the rising frequency chirp convolved with individual HRTFs are currently in progress.

Fourth, in the present thesis the dependence of the ABR and the BD on both ITD and ILD was investigated in detail. However, due to the limited measurement time only three values of the ITD and the ILD could be considered. BD measurements for ITDs in the range from 0 to 2 ms with a small spacing are especially interesting because the latency of the BD components as function of the ITD allows to test the coincidence-detection model of binaural interaction (Ungan et al., 1997). At present, experiments using the chirp stimulus with 17 ITDs presented in an interleaved fashion are performed.

Last but not least, measurements of evoked potentials to spatial stimuli should be extended to the middle latency auditory evoked potentials (MAEPs) which represent the activity of the primary auditory cortex. Optimized stimulation procedures to allow for the simultaneous recording of ABRs and MAEPs are currently tested.

Even though not all problems of recording binaural interaction from the brain stem could be solved here, the current thesis demonstrates that a rough neural correlate of localization in humans is already present on the brain stem level, while a more refined analysis of spatially separated acoustic objects seems to take place at higher stations of the auditory pathway.

Appendix A

Derivation of the variances of the average-based and the single-sweep-based noise estimate

In this appendix all quantities are treated as time independent, i.e., for one time sample only. However, the results are the same for each sample in time. Similar results hold for the time averaged quantities.

To prove equations (2.17), (2.18), and (2.19) the following assumptions have to be fulfilled:

- From the measured values x_j, $j = 1, \ldots J$, a signal estimate s, cf. eq. (2.15), is obtained, which is a random variable. From s, the corresponding noise estimates $n_j = x_j - s$, $j = 1, \ldots J$ can be derived. They are also random variables.

- The measured values x_j consist of a reproducible signal S and a random additive noise $N_j : x_j = S + N_j$.

- The noise values are realisations of a random variable N_j that obey a Gaussian distribution with zero mean and variance σ_0^2.

- Noise values N_j and N_k are uncorrelated for all epochs with $j \neq k$.

According to eq. (2.14) the average-based estimate of the noise variance is

$$\sigma_{\text{oe}}^2 = \frac{1}{4}(s_o - s_e)^2 = \frac{1}{4}\Big(S + \frac{2}{J}\sum_{o=1,3,\ldots}^{J-1} N_o - S - \frac{2}{J}\sum_{e=2,4,\ldots}^{J} N_e\Big)^2 = \frac{1}{J^2}\Big(\sum_o N_o - \sum_e N_e\Big)^2. \quad \text{(A.1)}$$

Given the independence of the noise values, i.e., $E(N_j N_k) = E(N_j)E(N_k)$ the expectation value of σ_{oe}^2 can be written as

$$
\begin{aligned}
E(\sigma_{oe}^2) &= \frac{1}{J^2}\left\{ E\left(\left(\sum_o N_o\right)^2\right) - 2E\left(\sum_o N_o \sum_e N_e\right) + E\left(\left(\sum_e N_e\right)^2\right)\right\} \\
&= \frac{1}{J^2}\left\{ \sum_o E(N_o^2) - 2E\left(\sum_o N_o\right)E\left(\sum_e N_e\right) + \sum_e E(N_e^2)\right\} \\
&= \frac{1}{J^2}\left\{ \frac{J}{2}\sigma_0^2 - 0 + \frac{J}{2}\sigma_0^2\right\} = \frac{\sigma_0^2}{J} \; .
\end{aligned}
\tag{A.2}
$$

According to eq. (2.16) the single-sweep-based estimate of the noise variance is

$$
\sigma^2 = \frac{1}{J(J-1)}\sum_{j=1}^{J}(x_j - s)^2 = \frac{1}{J(J-1)}\sum_{j=1}^{J}n_j^2 \; .
\tag{A.3}
$$

The division by $J(J-1)$ is necessary because s is the signal estimate, not the true signal S. Without loss in generality, however, it can be assumed, that the true signal S is known. This provides an extra degree of freedom and leads to

$$
\sigma^2 = \frac{1}{J^2}\sum_{j=1}^{J}(x_j - S)^2 = \frac{1}{J^2}\sum_{j=1}^{J}N_j^2 \; .
\tag{A.4}
$$

The expectation value of the single-sweep-based estimate of the noise variance is

$$
E(\sigma^2) = \frac{1}{J^2}E\left(\sum_{j=1}^{J}N_j^2\right) = \frac{1}{J^2}\sum_{j=1}^{J}E(N_j^2) = \frac{1}{J^2}\sum_{j=1}^{J}\sigma_0^2 = \frac{\sigma_0^2}{J} \; .
\tag{A.5}
$$

Herewith the equality of the expectation values of both noise estimates (eq. (2.17)) is proven. The result reflects the well known fact that during the averaging process the variance of the noise is reduced by the number of epochs entering the average.

The crucial difference between σ_{oe}^2 and σ^2 is their respective variance. For the derivation of these quantities, the kurtosis of the sum of normal distributed random variables with variance σ_0^2 is needed. With the identity

$$
\int_0^\infty x^m e^{-ax^2}dx = \frac{\Gamma(\frac{m+1}{2})}{2a^{(\frac{m+1}{2})}} \quad ,a > 0, \; m > -1
\tag{A.6}
$$

and setting $a = 1/(2\sigma_0^2)$ and $m = 4$, it follows for $J = 1$

$$E(N^4) = \frac{1}{\sqrt{2\pi}\sigma_0} \int_{-\infty}^{\infty} N^4 e^{-\frac{N^2}{2\sigma_0^2}} dN = \frac{2}{\sqrt{2\pi}\sigma_0} \Gamma\left(\frac{5}{2}\right) \frac{1}{2} (2\sigma_0^2)^{5/2} = 3\sigma_0^4. \qquad (A.7)$$

Since only quadratic terms in N_j contribute to the expectation value, the sum of $J/2$ terms is

$$
\begin{aligned}
E\left\{\left(\sum_{j=1}^{J/2} N_j\right)^4\right\} &= E\left\{\sum_{j=1}^{J/2} N_j^4 + \sum_{j=1}^{J/2}\sum_{\substack{k=1\\k\neq j}}^{J/2} \binom{4}{2} N_j^2 N_k^2\right\} \\
&= \sum_{j=1}^{J/2} E(N_j^4) + 6 \sum_{j=1}^{J/2}\sum_{\substack{k=1\\k\neq j}}^{J/2} E(N_j^2)E(N_k^2) \\
&= \frac{J}{2} 3\sigma_0^4 + 6\frac{J}{2}\left(\frac{J}{2}-1\right)\sigma_0^2\sigma_0^2 \\
&= \frac{3}{2}J(J-1)\sigma_0^4.
\end{aligned}
\qquad (A.8)
$$

With this result and assuming $J \gg 1$, the variance of the average-based estimate of the noise variance can be calculated:

$$
\begin{aligned}
\mathrm{var}(\sigma_{\mathrm{oe}}^2) &= E((\sigma_{\mathrm{oe}}^2)^2) - (E(\sigma_{\mathrm{oe}}^4))^2 = E(\sigma_{\mathrm{oe}}^4) - \frac{\sigma_0^4}{J^2} \\
&= \frac{1}{J^4} E\left\{\left(\sum_o N_o - \sum_e N_e\right)^4\right\} - \frac{\sigma_0^4}{J^2} \\
&= \frac{1}{J^4}\left\{E\left(\left(\sum_o N_o\right)^4\right) - 4E\left(\left(\sum_o N_o\right)^3\left(\sum_e N_e\right)\right) + 6E\left(\left(\sum_o N_o\right)^2\left(\sum_e N_e\right)^2\right)\right. \\
&\quad \left. -4E\left(\left(\sum_o N_o\right)\left(\sum_e N_e\right)^3\right) + E\left(\left(\sum_e N_e\right)^4\right)\right\} - \frac{\sigma_0^4}{J^2} \\
&= \frac{1}{J^4}\left\{\frac{3}{2}J(J-1)\sigma_0^4 + 6\left(\frac{J}{2}\sigma_0^2\right)\left(\frac{J}{2}\sigma_0^2\right) + \frac{3}{2}J(J-1)\sigma_0^4\right\} - \frac{\sigma_0^4}{J^2} \\
&= \frac{\sigma_0^4}{J^2}\left(\frac{7}{2} - \frac{3}{J}\right) \approx \frac{7}{2}\frac{\sigma_0^4}{J^2}.
\end{aligned}
\qquad (A.9)
$$

On the other hand, the variance of the single-sweep-based estimate of the noise variance is

$$
\begin{aligned}
\mathrm{var}(\sigma^2) &= \mathrm{E}((\sigma^2)^2) - (\mathrm{E}(\sigma^2))^2 \\
&= \frac{1}{J^4}\,\mathrm{E}\Big\{(\sum_{j=1}^{J} N_j^2)(\sum_{k=1}^{J} N_k^2)\Big\} - \frac{\sigma_0^4}{J^2} \\
&= \frac{1}{J^4}\Big(\sum_{j=1}^{J}\mathrm{E}(N_j^4) + \sum_{j=1}^{J}\sum_{\substack{k=1 \\ k\neq j}}^{J}\mathrm{E}(N_j^2)\,\mathrm{E}(N_k^2)\Big) - \frac{\sigma_0^4}{J^2} \\
&= \frac{1}{J^4}\Big(3J\sigma_0^4 + J(J-1)\sigma_0^2\sigma_0^2\Big) - \frac{\sigma_0^4}{J^2} \\
&= 2\,\frac{\sigma_0^4}{J^3}\,.
\end{aligned}
\tag{A.10}
$$

With these derivations equations (2.18) and (2.19) are proven. Expectation value and variance of the single-sweep-based noise variance estimate can also be derived by noting that

$$
\sigma^2 = \frac{\sigma_0^2}{J^2}\sum_{j=1}^{J}\left(\frac{N_j}{\sigma_0}\right)^2 .
\tag{A.11}
$$

The sum in the above equation obeys a χ^2-distribution with $\nu = J$ degrees of freedom and therefore has expectation value J and variance $2J$ which is consistent with equations (A.5) and (A.10).

Appendix B

Transformation of the noise covariance matrix into average reference

In this appendix equations for the transformation of average data and for the noise covariance matrix $\mathbf{S}$ from recording reference to average reference and vice versa are derived. The transformations do not depend on time and therefore are given for a single time sample.

Consider an EEG measurement with C channels, $c = 1, \ldots, C$ containing J single sweeps, $j = 1, \ldots, J$. Let x_{jc} be the EEG-signal of sweep j and channel c. The evoked potential of a channel c is obtained by averaging the x_{jc}. A conventional, i.e., non-weighted average is assumed here:

$$\mu_c = \frac{1}{J} \sum_{j=1}^{J} x_{jc}. \tag{B.1}$$

The elements of the noise covariance matrix $\mathbf{S}$ are defined as

$$\sigma_{cd} = \frac{1}{J(J-1)} \sum_{j=1}^{J} (x_{jc} - \mu_c)(x_{jd} - \mu_d) = \frac{1}{J(J-1)} \sum_{j=1}^{J} x_{jc} x_{jd} \; - \; \frac{1}{J-1} \mu_c \mu_d \tag{B.2}$$

with c and d being the indices of two channels. On the diagonal of $\mathbf{S}$ the noise variances, i.e., the squared standard errors $\sigma_c^2 = \sigma_{cc}$ are found.

Transformation from recording reference to average reference

The transformation of a signal into average reference is performed in two steps: First an additional channel, i.e., the reference channel with number $C + 1$, is explicitly introduced and initialized with the value zero. Second, the mean over all channels including the (new) reference channel is subtracted to obtain the signal $x_{jc'}$ in average reference:

$$x_{jc'} = x_{jc} - \frac{1}{C+1} \sum_{c=1}^{C+1} x_{jc}, \quad c' = 1, \ldots, C+1. \tag{B.3}$$

It is easily derived that the transformation of the average signal μ_c into average reference is analogous:

$$\mu_{c'} = \mu_c - \frac{1}{C+1} \sum_{c=1}^{C+1} \mu_c, \quad c' = 1, \ldots, C+1. \tag{B.4}$$

Before the transformation of the noise covariance matrix $\mathbf{S}$ into average reference, the reference channel is explicitely added and initialized with zeros:

$$\sigma_{c,C+1} = \sigma_{C+1,c} = 0, \qquad c = 1, \ldots, C+1. \tag{B.5}$$

The elements of the covariance matrix $\mathbf{S}'$ in average reference are given by

$$J(J-1)\sigma_{c'd'}$$

$$= \sum_{j=1}^{J} x_{jc'} x_{jd'} - J\mu_{c'}\mu_{d'}$$

$$= \sum_{j=1}^{J} \left(x_{jc} - \frac{1}{C+1} \sum_{c=1}^{C+1} x_{jc} \right)\left(x_{jd} - \frac{1}{C+1} \sum_{d=1}^{C+1} x_{jd} \right) - J\mu_{c'}\mu_{d'}$$

$$= \sum_{j=1}^{J} x_{jc} x_{jd} - \frac{1}{C+1} \sum_{c=1}^{C+1} \sum_{j=1}^{J} x_{jc} x_{jd} - \frac{1}{C+1} \sum_{d=1}^{C+1} \sum_{j=1}^{J} x_{jc} x_{jd}$$

$$+ \frac{1}{(C+1)^2} \sum_{c=1}^{C+1} \sum_{d=1}^{C+1} \sum_{j=1}^{J} x_{jc} x_{jd} - J\left(\mu_c - \frac{1}{C+1} \sum_{c=1}^{C+1} \mu_c \right)\left(\mu_d - \frac{1}{C+1} \sum_{d=1}^{C+1} \mu_d \right)$$

$$= J(J-1)\sigma_{cd} + J\mu_c\mu_d - \frac{J(J-1)}{C+1} \sum_{c=1}^{C+1} \sigma_{cd} - \frac{J}{C+1}\mu_d \sum_{c=1}^{C+1} \mu_c - \frac{J(J-1)}{C+1} \sum_{d=1}^{C+1} \sigma_{cd}$$

$$- \frac{J}{C+1}\mu_c \sum_{d=1}^{C+1} \mu_d + \frac{J(J-1)}{(C+1)^2} \sum_{c=1}^{C+1} \sum_{d=1}^{C+1} \sigma_{cd} + \frac{J}{(C+1)^2} \sum_{c=1}^{C+1} \mu_c \sum_{d=1}^{C+1} \mu_d$$

$$- J\mu_c\mu_d + \frac{J}{C+1}\mu_d \sum_{c=1}^{C+1} \mu_c + \frac{J}{C+1}\mu_c \sum_{d=1}^{C+1} \mu_d - \frac{J}{(C+1)^2} \sum_{c=1}^{C+1} \mu_c \sum_{d=1}^{C+1} \mu_d.$$

All terms containing the means vanish. The transformation of $\mathbf{S}$ from recording to average reference $\mathbf{S}'$ is

$$\sigma_{c'd'} = \sigma_{cd} - \frac{1}{C+1}\sum_{c=1}^{C+1}\sigma_{cd} - \frac{1}{C+1}\sum_{d=1}^{C+1}\sigma_{cd} + \frac{1}{(C+1)^2}\sum_{c=1}^{C+1}\sum_{d=1}^{C+1}\sigma_{cd}. \tag{B.6}$$

To get an element $\sigma_{c'd'}$ in average reference, take the original value σ_{cd} in recording reference, subtract the mean of the values in row c and column d and finally add the mean of all elements in the matrix.

In the case that $\mathbf{S}$ in recording reference is diagonal ($\sigma_{cd} = \delta_{cd}\sigma_c^2$), i.e., only the variances are considered, the transformation reduces to

$$\sigma_{c'}^2 = \sigma_c^2 - \frac{2}{C+1}\sigma_c^2 + \frac{1}{(C+1)^2}\sum_{c=1}^{C+1}\sigma_c^2. \tag{B.7}$$

Transformation from average reference to recording reference

The transformation back to recording reference is rather simple. For single sweep data and averages, the last channel has to be subtracted from all channels. After this operation the last channel, i.e., the reference channel, contains zeros. The reference channel is normally not written in recording reference. The backtransformation of the single sweep is

$$x_{jc} = x_{jc'} - x_{j,C'+1}. \tag{B.8}$$

An analogue formula holds for the backtransformation of the mean:

$$\mu_c = \mu_{c'} - \mu_{C'+1}. \tag{B.9}$$

The backtransformation of $\mathbf{S}'$ to recording reference is

$$
\begin{aligned}
J(J-1)\sigma_{cd} &= \sum_{j=1}^{J} x_{jc'}x_{jd'} - J\mu_{c'}\mu_{d'} \\
&= \sum_{j=1}^{J}\left(x_{jc'} - x_{j,C'+1}\right)\left(x_{jd'} - x_{j,C'+1}\right) - J\left(\mu_{c'} - \mu_{C'+1}\right)\left(\mu_{d'} - \mu_{C'+1}\right) \\
&= \sum_{j=1}^{J} x_{jc'}x_{jd'} - \sum_{j=1}^{J} x_{jc'}x_{j,C'+1} - \sum_{j=1}^{J} x_{jd'}x_{j,C'+1} + \sum_{j=1}^{J} x_{j,C'+1}^2 \\
&\quad - J\mu_{c'}\mu_{d'} + J\mu_{c'}\mu_{C'+1} + J\mu_{d'}\mu_{C'+1} - J\mu_{C'+1}^2.
\end{aligned}
$$

Therefore, the transformation formula from average reference to recording reference is

$$\sigma_{cd} = \sigma_{c'd'} - \sigma_{c',C'+1} - \sigma_{d',C'+1} + \sigma_{C'+1,C'+1} \tag{B.10}$$

In the case of vanishing off-diagonal elements of $\mathbf{S}'$, i.e., $\sigma_{c'd'} = \delta_{c'd'}\sigma_{c'}^2$, the backtransformation reduces to

$$\sigma_c^2 = \sigma_{c'}^2 + \sigma_{C'+1}^2 \tag{B.11}$$

Notes and implications

1. For non-weighted averages, data from single sweeps that entered the average is not needed for the transformation. The mean and noise covariance in average reference are only functions of the mean and the noise covariance in recording reference, respectively.

2. It is not possible to transform a weighted average and the corresponding weighted noise covariance matrix into average reference without the information of the weightings of the single sweeps.

3. In recording reference the voltage at the reference electrode is zero by definition and therefore usually not mentioned. By transforming to average reference, an 'additional' channel $C+1$ with non-vanishing elements is introduced. This is possible because only a voltage difference, i.e., the electric field, is a meaningful physical quantity. Therefore, in any reference C voltage differences are described.

4. An experimentally derived noise covariance matrix $\mathbf{S}$ usually has full rank C, condition values for $C = 32$ channels typically vary between 10 and 30. $\mathbf{S}$ must be invertible to solve the inverse problem (cf. eq. 4.4 and eq. 4.7).

5. The elements of $\mathbf{S}'$ are linear combinations of the elements of $\mathbf{S}$ (cf. eq. B.6). Therefore $\mathbf{S}'$ does not have full rank $C + 1$, but rank C. This implicates that the inverse problem cannot be solved in average reference if the complete noise covariance matrix is used.

Bibliography

Achor, L. J., Starr, A., 1980. Auditory brain stem responses in the cat. I. Intracranial and extracranial recordings. Electroenceph. clin. Neurophysiol. 48: 154–173. 42

Ainslie, P. J., Boston, J. R., 1980. Comparison of brain stem auditory evoked potentials for monaural and binaural stimuli. Electroenceph. clin. Neurophysiol. 49: 291–302. 42, 43

Ary, J. P., Klein, S. A., Fender, D. H., 1981. Location of sources of evoked scalp potentials: corrections for skull and scalp thicknesses. IEEE Trans. Biomed. Eng. 28(6): 447–452. 102

Berger, H., 1929. Über das Elektroenkephalogramm des Menschen. Arch. Psychiat. Nervenkr. 87: 527–570. 2

Bertrand, O., Perrin, F., Pernier, J., 1985. A theoretical justification of the average reference in topographic evoked potential studies. Electroenceph. clin. Neurophysiol. 62(6): 462–464. 100

Brantberg, K., Fransson, P. A., Hansson, H., Rosenhall, U., 1999a. Measures of the binaural interaction component in human auditory brainstem response using objective detection criteria. Scand. Audiol. 28(1): 15–26. 42, 43, 46, 66, 69, 106

Brantberg, K., Hansson, H., Fransson, P. A., Rosenhall, U., 1999b. The binaural interaction component in human ABR is stable within the 0- to 1-ms range of interaural time differences. Audiol. Neurootol. 4: 88–94. 42, 62, 66, 106

Brody, D. A., Terry, F. H., Ideker, R. E., 1973. Eccentric dipole in a spherical medium: generalized expression for surface potentials. IEEE Trans. Biomed. Eng. 20(3): 141–143. 71

Caird, D., Klinke, R., 1983. Representation of sound frequency and laterality by units in the central nucleus of cat inferior colliculus. Exp. Brain. Res. 52: 385–399. 42, 66

Caird, D., Sontheimer, D., Klinke, R., 1985. Intra- and extracranially recorded auditory evoked potentials in the cat: I. Source location and binaural interaction. Electroenceph. clin. Neurophysiol. 61: 50–60. 42, 66

Cebulla, M., Stürzebecher, E., Wernecke, K.-D., 2000a. Objective detection of auditory brainstem potentials: comparison of statistical tests in the time and frequency domains. Scand. Audiol. 29(1): 44–51. 6, 14, 37

Cebulla, M., Stürzebecher, E., Wernecke, K.-D., 2000b. Untersuchung verschiedener SNR–Schätzer für den Nachweis von biologischen Reizantworten im Rauschen. Z. Audiol. 39(1): 14–22. 6, 14, 37

Colburn, H. S., Durlach, N. I., 1978. Models of binaural interaction. In: Carterette, E. C., Friedman, M. P., eds., Handbook of Perception – Hearing, vol. IV, chap. 11, pages 467–518. Academic Press, New York. 42, 66

Cone-Wesson, B., Ma, E., Fowler, C. G., 1997. Effect of stimulus level and frequency on ABR and MLR binaural interaction in human neonates. Hear. Res. 106: 163–178. 42, 66

Cuffin, B. N., Cohen, D., 1979. Comparison of the magnetoencephalogram and electroencephalogram. Electroenceph. clin. Neurophysiol. 47(2): 132–146. 70, 102

Cuffin, B. N., Schomer, D. L., Ives, J. R., Blume, H., 2001. Experimental tests of EEG source localization accuracy in spherical head models. Clin. Neurophysiol. 112: 46–51. 103

Damaschke, J., Granzow, M., Riedel, H., Kollmeier, B., 2000. Zur Äquivalenz von interauralen Zeit- und Pegelunterschieden bei kurzen Stimuli. Z. Audiol. 39: 40–52. 42

Dau, T., Wegner, O., Kollmeier, B., 2000. Auditory brain stem responses with optimized chirp signals compensating basilar-membrane dispersion. J. Acoust. Soc. Am. 107(3): 1530–1540. 3, 103, 105, 106, 108, 122, 123, 127

Dawson, G. D., 1951. A summation technique for detecting small signals in a large irregular background. J. Physiol. 115: 2. 2

Dawson, G. D., 1954. A summation technique for the detection of small evoked potentials. Electroenceph. clin. Neurophysiol. 6: 65–84. 2

de Boer, E., 1980. Auditory physics. Physical principles in hearing theory I. Phys. Rep. 62: 87–174. 123

Desmedt, J. E., Chalklin, V., Tomberg, C., 1990a. Emulation of somatosensory evoked potential (SEP) components with the 3-shell head model and the problem of 'ghost potential fields' when using an average reference in brain mapping. Electroenceph. clin. Neurophysiol. 77(4): 243–258.   100

Desmedt, J. E., Tomberg, C., 1990. Topographic analysis in brain mapping can be compromised by the average reference. Brain Topogr. 3(1): 35–42.   100

Desmedt, J. E., Tomberg, C., Noel, P., Ozaki, I., 1990b. Beware of the average reference in brain mapping. Electroenceph. clin. Neurophysiol. suppl. 41: 22–27.   100

Dobie, R. A., Berlin, C. I., 1979. Binaural interaction in brainstem-evoked responses. Arch. Otolaryngol. 105: 391–398.   42

Dobie, R. A., Norton, S. J., 1980. Binaural interaction in human auditory evoked responses. Electroenceph. clin. Neurophysiol. 49: 303–313.   42

Don, M., Elberling, C., 1994. Evaluating residual background noise in human auditory brain-stem responses. J. Acoust. Soc. Am. 96(5, pt.1): 2746–2757.   38

Don, M., Elberling, C., Waring, M., 1984. Objective detection of averaged auditory brainstem responses. Scand. Audiol. 13: 219–228.   6, 37

Durlach, N. I., Colburn, H. S., 1978. Binaural phenomena. In: Carterette, E. C., Friedman, M. P., eds., Handbook of Perception – Hearing, vol. IV, chap. 10, pages 365–466. Academic Press, New York.   1, 42, 66

Edelman, G. M., Gall, W. E., Cowan, W. M., 1988. Auditory Function: Neurobiological Bases of Hearing. John Wiley and Sons, New York.   1

Elberling, C., Don, M., 1984. Quality estimation of averaged auditory brainstem responses. Scand. Audiol. 13: 187–197.   6, 14, 37

Elberling, C., Wahlgreen, O., 1985. Estimation of auditory brainstem response, ABR, by means of Bayesian inference. Scand. Audiol. 14: 89–96.   7, 11

Feddersen, W. E., Sandel, T. T., Teas, D. C., Jeffress, L. A., 1957. Localization of high-frequency tones. J. Acoust. Soc. Am. 29: 988–991.   42, 66

Furst, M., Eyal, S., Korczyn, A. D., 1990. Prediction of binaural click lateralization by brainstem auditory evoked potentials. Hear. Res. 49(1-3): 347–359.   106

Furst, M., Levine, R. A., McGaffigan, P. M., 1985. Click lateralization is related to the β component of the dichotic brainstem auditory evoked potentials of human subjects. J. Acoust. Soc. Am. 78(5): 1644–1651. 42, 43, 45, 62, 66, 68, 106

Gaumond, R. P., Psaltikidou, M., 1991. Models for the generation of the binaural difference response. J. Acoust. Soc. Am. 89(1): 454–456. 63, 64, 123, 124

Gerull, G., Mrowinski, D., 1984. Brain stem potentials evoked by binaural click stimuli with differences in interaural time and intensity. Audiology 23: 265–276. 42, 66

Gevins, A. S., Rémond, A., 1987. Methods of Analysis of Brain Electrical and Magnetic Signals, vol. 1 of Handbook of Electroencephalography and Clinical Neurophysiology. Elsevier, Amsterdam. 6

Goldberg, J. M., Brown, P. B., 1968. Functional organization of the dog superior olivary complex: an anatomical and electrophysiological study. J. Neurophysiol. 31: 639–656. 66

Goldberg, J. M., Brown, P. B., 1969. Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: some physiological mechanisms of sound localization. J. Neurophysiol. 36: 157–178. 63, 66, 107, 122

Granzow, M., Riedel, H., Kollmeier, B., 2001. Single-sweep-based methods to improve the quality of auditory brain stem responses. Part I: Optimized linear filtering. Z. Audiol. 40(1): 32–44. 6, 37, 46, 62, 69, 109

Greenwood, D. D., 1990. A cochlear frequency-position function for several species – 29 years later. J. Acoust. Soc. Am. 87(6): 2592–2605. 106

Gummer, A. W., Zenner, H.-P., 1996. Central processing of the auditory information. In: Greger, R., Windhorst, U., eds., Comprehensive Human Physiology, vol. 1, chap. 36, pages 729–756. Springer, Berlin. 42, 66

Helmholtz, H., 1853. Über einige Gesetze der Vertheilung elektrischer Ströme in körperlichen Leitern mit Anwendung auf die thierisch-elektrischen Versuche. Annalen der Physik und Chemie 89: 211–233, 353–377. 3

Hoke, M., Ross, B., Wickesberg, R., Lütkenhöner, B., 1984. Weighted averaging - theory and application to electric response audiometry. Electroenceph. clin. Neurophysiol. 57: 484–489. 7, 10

Hoth, S., 1986. Reliability of latency and amplitude values of auditory-evoked potentials. Audiology 25: 248–257. 47, 110, 113

Irvine, D. R., 1992. Physiology of the auditory brainstem. In: Popper, A. N., Fay, R. R., eds., The Mammalian Auditory Pathway: Neurophysiology, vol. 2, chap. 4, pages 153–231. Springer, New York. 1, 63, 66, 122

Ito, S., Hoke, M., Pantev, C., Lütkenhöner, B., 1988. Binaural interaction in brainstem auditory evoked potentials elicited by frequency-specific stimuli. Hear. Res. 35: 9–20. 42, 46, 48, 66, 69, 106, 110

Jasper, H. H., 1957. The ten twenty electrode system of the international federation. Electroenceph. clin. Neurophysiol. 10: 371–375, appendix. 45, 68, 108

Jeffress, L. A., 1948. A place theory of sound localization. J. Comp. Physiol. Psychol. 41: 35–39. 63, 64

Jeffress, L. A., McFadden, D., 1971. Differences of interaural phase and level in detection and lateralization. J. Acoust. Soc. Am. 49(4): 1169–1179. 1, 66

Jewett, D. L., 1987. The 3-channel Lissajous' trajectory of the auditory brain-stem response. IX. Theoretical aspects. Electroenceph. clin. Neurophysiol. 68: 386–408. 66

Jewett, D. L., Romano, M. N., Williston, J. S., 1970. Human auditory evoked potentials: possible brain stem components detected on the scalp. Science 167: 1517–1518. 2

Jewett, D. L., Williston, J. S., 1971. Auditory-evoked far fields averaged from the scalp of humans. Brain 94(4): 681–696. 2

Jiang, Z. D., 1996. Binaural interaction and the effects of stimulus intensity and repetition rate in human auditory brain-stem. Electroenceph. clin. Neurophysiol. 100: 505–516. 42, 66, 106

John, E. R., Harmony, T., Valdes-Sosa, P., 1987. The use of statistics in electrophysiology. In: Gevins, A. S., Rémond, A., eds., Handbook of Electroencephalography and Clinical Neurophysiology, vol. 1, chap. 16, pages 497–540. Elsevier, Amsterdam. 72

Jones, S. J., van der Poel, J. C., 1990. Binaural interaction in the brain-stem auditory evoked potential: evidence for a delay line coincidence detection mechanism. Electroenceph. clin. Neurophysiol. 77: 214–224. 42, 46, 62, 66, 69, 106

Joseph, A. W., Hyson, R. L., 1993. Coincidence detection by binaural neurons in the chick brain stem. J. Neurophysiol. 69(4): 1197–1211.   42

Kavanagh, R. N., Darcey, T. M., Lehmann, D., Fender, D. H., 1978. Evaluation of methods for three-dimensional localization of electrical sources in the human brain. IEEE Trans. Biomed. Eng. 25(5): 421–429.   67

Kelly-Ballweber, D., Dobie, R. A., 1984. Binaural interaction measured behaviorally and electrophysiologically in young and old adults. Audiology 23: 181–194.   42

Koles, Z. J., 1998. Trends in EEG source localization. Electroenceph. clin. Neurophysiol. 106(2): 127–137.   127

Koles, Z. J., Soong, A. C., 1998. EEG source localization: implementing the spatio-temporal decomposition approach. Electroenceph. clin. Neurophysiol. 107(5): 343–352. 127

Levine, R. A., 1981. Binaural interaction in brainstem potentials of human subjects. Ann. Neurol. 9: 384–393.   42, 43, 46, 48, 62, 64, 66, 69, 106, 107, 110

Levine, R. A., Davis, P. J., 1991. Origin of the click-evoked binaural interaction potential, β, of humans. Hear. Res. 57(1): 121–128.   106, 123

Lütkenhöner, B., 1998a. Dipole source localization by means of maximum likelihood estimation I. Theory and simulations. Electroenceph. clin. Neurophysiol. 106(4): 314–321.   99

Lütkenhöner, B., 1998b. Dipole source localization by means of maximum likelihood estimation. II. Experimental evaluation. Electroenceph. clin. Neurophysiol. 106(4): 322–329.   99

Lütkenhöner, B., Hoke, M., Pantev, C., 1985. Possibilities and limitations of weighted averaging. Biol. Cybern. 52: 409–416.   7, 15, 37

McPherson, D. L., Starr, A., 1995. Auditory time-intensity cues in the binaural interaction component of the auditory evoked potentials. Hear. Res. 89: 162–171.   42, 43, 45, 46, 62, 66, 68, 69

McPherson, D. L., Tures, C., Starr, A., 1989. Binaural interaction of the auditory brainstem potentials and middle latency auditory evoked potentials in infants and adults. Electroenceph. clin. Neurophysiol. 74: 124–130.   42

Melcher, J. R., Kiang, N. Y. S., 1996. Generators of the brain-stem auditory-evoked potential in cat. 3. Identified cell-populations. Hear. Res. 93(1-2): 52–71.   64, 66

Mühler, R., von Specht, H., 1999. Sorted averaging – Principle and application to auditory brainstem responses. Scand. Audiol. 28(3): 145–149.   6, 7, 9

Mills, A. W., 1958. On the minimum audible angle. J. Acoust. Soc. Am. 30: 237–246.   42, 66

Mills, A. W., 1960. Lateralization of high-frequency tones. J. Acoust. Soc. Am. 32: 132–134.   42, 66

Moore, B. C. J., 1995. Hearing. Academic Press, San Diego.   1

Mosher, J. C., Lewis, P. S., Leahy, R. M., 1992. Multiple dipole modeling and localization from spatio-temporal MEG data. IEEE Trans. Biomed. Eng. 39(6): 541–557.   67, 127

Nelder, J. A., Mead, R., 1965. A simplex method for function minimization. The Computer Journal 8: 308–313.   73

Nieuwenhuys, R., Vogel, J., van Huijzen, C., 1988. The Human Central Nervous System – a Synopsis and Atlas. Springer, Berlin.   42, 66, 80, 101

Nordby, K., Rosness, R., Skotturi, B. C., 1982. Binaural time/intensity trading. J. Acoust. Soc. Am. 71(3): 132–134.   42

Onofrj, M., Fulgente, T., Malatesta, G., Locatelli, T., 1994. Focal abnormalities of P3 ERPs unveiled in patients with cortical lesions and primary progressive aphasia by average reference recordings. Brain Topogr. 6(4): 311–322.   100

Pascual-Marqui, R. D., 1999. Review of methods for solving the EEG inverse problem. Int. J. of Bioelectromagnetism 1(1): 75–86.   127

Pascual-Marqui, R. D., Michel, C. M., Lehmann, D., 1994. Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain. Int. J. Psychophysiol. 18: 49–65.   127

Polyakov, A., Pratt, H., 1994. Three-channel Lissajous' trajectory of the binaural interaction components in human auditory brain-stem evoked potentials. Electroenceph. clin. Neurophysiol. 92: 396–404.   67

Polyakov, A., Pratt, H., 1996. Evidence for spatio-topic organization of binaural processing in the human brainstem. Hear. Res. 94: 107–115.   67

Polyakov, A., Pratt, H., 1998. The effect of binaural masking noise disparity on human auditory brainstem binaural interaction components. Audiology 37: 17–26.   67

Popper, A. N., Fay, R. R., 1992. The Mammalian Auditory Pathway: Neurophysiology, vol. 2 of Springer Handbook of Auditory Research. Springer, New York.   1, 42

Pratt, H., Har'el, Z., Golos, E., 1983. Three-channel lissajous' trajectory of human auditory brain stem evoked potentials. Electroenceph. clin. Neurophysiol. 56: 682–688. 66

Pratt, H., Martin, W. H., Kaminer, M., Bleich, N., 1990. Three-channel lissajous' trajectories of auditory brain stem evoked potentials. Quantitative measures, effects of stimulus parameters and clinical promise. In: Grandori, F., Hoke, M., Romani, G. L., eds., Auditory Evoked Magnetic Fields and Electric Potentials, vol. 6, pages 331–356. Karger, Basel.   66

Pratt, H., Polyakov, A., Aharonson, V., Korczyn, A. D., Tadmor, R., Fullerton, B. C., Levine, R. A., Furst, M., 1998. Effects of localized pontine lesions on auditory brain-stem evoked potentials and binaural processing in humans. Electroenceph. clin. Neurophysiol. 108: 511–520.   67

Pratt, H., Polyakov, A., Kontorovich, L., 1997. Evidence for separate processing in the human brainstem of interaural intensity and temporal disparities for sound lateralization. Hear. Res. 108: 1–8.   67

Press, W. H., Teukolsky, S. A., Vetterling, W. T., Flannery, B. P., 1992. Numerical Recipes in C, the Art of Scientific Computing. Cambridge University Press, Cambridge. 75

Rayleigh, L. S. J. W., 1907. On our perception of sound direction. Philos. Mag. 13: 214–232.   63, 66

Riedel, H., Granzow, M., Kollmeier, B., 2001. Single-sweep-based methods to improve the quality of auditory brain stem responses. Part II: Averaging methods. Z. Audiol. 40(2): 62–85.   5, 43, 62, 109

Riedel, H., Kollmeier, B., 2002a. Auditory brain stem responses evoked by lateralized clicks: Is lateralization extracted in the human brain stem ? Hear. Res. 163(1-2): 12–26. 41, 66, 106

Riedel, H., Kollmeier, B., 2002b. Comparison of binaural auditory brain stem responses and the binaural difference potential evoked by chirps and clicks. Hear. Res. 169(1-2): 85–96. 105

Scherg, M., 1984. Spatio-temporal modelling of early auditory evoked potentials. Rev. Laryngol. Otol. Rhinol. 105(2, suppl.): 163–170. 67

Scherg, M., 1990. Fundamentals of dipole source potential analysis. In: Grandori, F., Hoke, M., Romani, G. L., eds., Auditory Evoked Magnetic Fields and Electric Potentials, vol. 6, pages 40–69. Karger, Basel. 67, 70

Scherg, M., 1991. Akustisch evozierte Potentiale: Grundlagen – Entstehungsmechanismen – Quellenmodell. Kohlhammer, Stuttgart. 64, 67, 70, 80

Scherg, M., von Cramon, D., 1985. A new interpretation of the generators of BAEP waves I-V: results of a spatio-temporal dipole model. Electroenceph. clin. Neurophysiol. 62: 290–299. 64, 67

Schimmel, H., 1967. The (±) reference: accuracy of estimated mean components in average response studies. Science 157: 92–93. 6, 13

Schröger, E., 1996. Interaural time and level differences: integrated or separated processing ? Hear. Res. 96(1-2): 191–198. 126

Schröger, E., Tervaniemi, M., Winkler, I., Wolff, C., Näätänen, R., 1997. Processing of interaural cues used for auditory lateralization as revealed by the mismatch negativity. In: Schick, A., Klatte, M., eds., Contributions to Psychological Acoustics. Results of the Seventh Oldenburg Symposium on Psychological Acoustics, pages 49–56. BIS, Oldenburg. 126

Sekihara, K., Ogura, Y., Hotta, M., 1992. Maximum-likelihood estimation of current-dipole parameters for data obtained using multichannel magnetometer. IEEE Trans. Biomed. Eng. 39(6): 558–562. 99

Semple, M. N., Aitkin, L. M., 1979. Representation of sound frequency and laterality by units in the central nucleus of cat inferior colliculus. J. Neurophysiol. 42(6): 1626–1639. 42

Sharbrough, F., Chatrian, G.-E., Lesser, R. P., Lüders, H., Nuwer, M., Picton, T. W., 1991. American Electroencephalographic Society Guidelines for Standard Electrode Position Nomenclature. J. Clin. Neurophysiol. 8(2): 200–202. 45, 68

Tomberg, C., Noel, P., Ozaki, I., Desmedt, J. E., 1990. Inadequacy of the average reference for the topographic mapping of focal enhancements of brain potentials. Electroenceph. clin. Neurophysiol. 77(4): 259–265.   100

Uhl, C., Hutt, A., Kruggel, F., 2001. Improvement of source localization by dynamical systems based modeling (DSBM). Brain Topogr. 13(3): 219–226.   127

Ungan, A., Yagcioglu, S., Özmen, B., 1997. Interaural delay-dependent changes in the binaural difference potential in cat auditory brainstem response: implications about the origin of the binaural interaction component. Hear. Res. 106: 66–82.   64, 128

van Adel, B. A., Kidd, S. A., Kelly, J. B., 1999. Contribution of the commissure of Probst to binaural evoked responses in the rat's inferior colliculus: interaural time differences. Hear. Res. 130: 115–130.   42, 66

von Békésy, G., 1960. Experiments in Hearing. McGraw Hill, New York.   106

Webster, D. B., Popper, A. N., Fay, R. R., 1992. The Mammalian Auditory Pathway: Neuroanatomy, vol. 1 of Springer Handbook of Auditory Research. Springer, New York. 1

Wong, P. K. H., Bickford, R. G., 1980. Brain stem auditory evoked potentials: the use of noise estimate. Electroenceph. clin. Neurophysiol. 50: 25–34.   6, 13

Wrege, K. S., Starr, A., 1981. Binaural interaction in human auditory brainstem evoked potentials. Arch. Neurol. 38: 572–580.   42, 43, 66

Yin, T. C., Chan, J. C., 1988. Neural mechanisms underlying interaural time sensitivity to tones and noise. In: Edelman, G. M., Gall, W. E., Cowan, W. M., eds., Auditory Function: Neurobiological Bases of Hearing, chap. 13, pages 385–430. John Wiley and Sons, New York.   63, 66

Yin, T. C. T., Chan, J. C. K., 1990. Interaural time sensitivity in medial superior olive of cat. J. Neurophysiol. 64(2): 465–488.   42, 63, 66

Yost, W. A., 1981. Lateral position of sinusoids presented with interaural intensive and temporal difference. J. Acoust. Soc. Am. 70(2): 397–409.   66

Yost, W. A., Gourevitch, G., 1987. Directional Hearing. Springer, New York.   1, 66

Erklärung

Hiermit erkläre ich, dass ich die vorliegende Dissertation selbstständig verfasst habe und nur die angegebenen Hilfsmittel verwendet habe.

Oldenburg, den 19. April 2002

Helmut Riedel

Danksagung

Lebenslauf

Am 23.7.1963 wurde ich als Sohn des Lehrers Joachim Riedel und der Lehrerin Gerda Riedel, geb. Szielenkewicz, in Lübeck geboren. Die Grundschulzeit verbrachte ich an der Gerhart-Hauptmann-Schule in Stockelsdorf bei Lübeck von 1969 – 1973. Anschliessend besuchte ich das Gymnasium am Mühlenberg in Bad Schwartau, wo ich am 18.5.1982 das Abitur absolvierte.

Nach der Grundwehrdienstzeit vom 1.7.1982 – 30.9.1983 nahm ich zum Wintersemester 1983/84 an der TU Braunschweig das Studium der Physik auf. Nach vier Semestern absolvierte ich am 18.10.1985 die Vordiplomsprüfungen und studierte zwei weitere Semester in Braunschweig. Aus Interesse an Biophysik und medizinischer Physik wechselte ich zum Wintersemester 1986/87 zur Universität Göttingen, an der ich mein Physikstudium am 30.10.1990 mit dem Diplom abschloss.

Die Diplomarbeit schrieb ich extern an der medizinischen Fakultät am Institut für Sinnes- und Neurophysiologie. In diesem Institut arbeitete ich vom 1.1. – 30.6.1991 als wissenschaftlicher Angestellter. Gleichzeitig nahm ich an der Universität Göttingen zum Sommersemester 1991 ein Zweitstudium der Philosophie und Musikwissenschaften auf. Im August 1994 begann ich, an der Universität Oldenburg in der Arbeitsgruppe 'Medizinische Physik' bei Prof. Dr. Dr. Kollmeier zu arbeiten. Ab 1.1.1995 wurde mir im Rahmen des Graduiertenkollegs Psychoakustik ein Stipendium bewilligt. Seit September 1998 arbeite ich im Sonderforschungsbereich 517 ('Neuronale Grundlagen kognitiver Leistungen').