Estimation of Postaverage SNR from Evoked Responses Under Nonstationary Noise

Ikaro Silva*, *Member, IEEE*

***Abstract*—In any measure of event-related potentials, it is important to be able to estimate the postaverage signal-to-noise ratio (SNR) in order to assess the quality of the measured signals. The estimated postaverage SNR can be an important detection criteria (as in infant hearing–screening of evoked auditory potentials) and a control factor when comparing signals obtained during different conditions (accounting for residual noise variability). Standard SNR estimation methods, such as the fixed-single-point (Fsp) statistic (C. Elberling and M. Don, "Quality estimation of averaged auditory brainstem responses," *Scandinavian Audiol.*, vol. 13, pp. 187–197, 1984), assume a single-stationary noise source, with the postaverage SNR increasing proportionally to the number of trials averaged. This study proposes a modified version of the Fsp statistic, the nonstationary fixed-multiple-point (NS Fmp), that can account for a discrete number of noise sources of different power, and can also be modified for weighted averaging (WNS Fmp). A new noise segmentation procedure is also proposed that dynamically partitions contiguous trials based on their noise power estimates and a series of *F*-tests. Results from computer simulation and real data from auditory brain stem recordings show that the NS Fmp method yields lower mean square error than do the Fsp, and that the WNS Fmp has higher receiver-operating-curve area than do the standard Fsp procedure.**

***Index Terms*—Bioelectric potentials, noise measurement, nonstationary analysis, weighted averaging.**

I. Introduction

In ANY measure of event-related potentials, it is crucially important to be able to estimate the postaverage signal-to-noise ratio (SNR) in order to assess the quality of the measured signals. The estimated postaverage SNR of evoked auditory potentials is important in: detection criterion in identifying infant hearing loss [1]–[3], comparing responses obtained across different conditions, and as a quality metric in comparing different methods attempting to improve the signal acquisition process [4]–[7]. It is also an important criterion in determining if a sub average is of sufficient quality for use as a reference signal in adaptive noise cancellation in evoked response measurements [8], [9]. Thus it is crucial that every study or protocol that uses a metric related to the amplitude (or power) of the averaged evoked response also takes into consideration the residual noise level.

Typically, the measured evoked response $x(t,k)$ at trial k and time t is modeled as a signal of interest $s(t)$ plus some stochastic noise

$$x(t,k) = s(t) + n_k(t) \tag{1}$$

where the noise source $n_k(t)$ is typically assumed to be a zero-mean Gaussian process [5]. It has been observed that for evoked responses, such as evoked auditory brainstem response (ABR), the power of the noise source $n_k(t)$ is nonstationary, even across several nearby trials [10], [11]. The noise source in (1) is, however, assumed to be *locally* stationary as a function of trial *k*. Following a matrix notation similar to that of [12], the ensemble data can be rewritten as

$$S(t) = \begin{bmatrix} s_1(t) \\ s_2(t) \\ \vdots \\ s_M(t) \end{bmatrix} \qquad N(t) = \begin{bmatrix} \eta_1(t) \\ \eta_2(t) \\ \vdots \\ \eta_M(t) \end{bmatrix}$$

$$X(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_M(t) \end{bmatrix} = S(t) + N(t) \tag{2}$$

where $S(t)$, $N(t)$, and $X(t)$ are all M-by-T matrices (T being the trial length). The postaverage SNR expressed in terms of (2) becomes

$$\text{SNR} = \frac{E\left\{\bar{w}^T S(t) S(t)^T \bar{w}\right\}}{E\left\{\bar{w}^T N(t) N(t)^T \bar{w}\right\}} = \frac{\bar{w}^T R_s \bar{w}}{\bar{w}^T R_\eta \bar{w}} \tag{3}$$

where $E\{\ \}$ denotes the expectation operation, $\bar{w}$ is an *M*-by-1 row vector of weights corresponding to the type of averaging being used, the matrices R_s and R_η are *M*-by-*M* covariance matrices for the signal and noise, respectively. If the signal is deterministic, all elements of its covariance matrix R_s are constant and equal to the signal's power over the trial σ_s^2; the signal's covariance matrix R_s can then be written as

$$R_s = \sigma_s^2 \begin{bmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{bmatrix}. \tag{4}$$

The assumption that the noise source is locally stationary and independent across trials limits the noise covariance matrix R_η in (3) to be diagonal. The noise can be modeled as $\hat{P}$ independent noise sources that are stationary across a variable number of trials. The noise covariance matrix R_η will then have a diagonal structure with $\hat{P}$ clusters of repeated elements,

$$R_\eta = \begin{bmatrix} R_{\eta 1} & & 0 \\ & R_{\eta 2} & \\ & & \ddots \\ 0 & & R_{\eta \hat{P}} \end{bmatrix} \qquad \text{where } R_{\eta i} = \sigma_i^2 I \tag{5}$$

and $R_{\eta i}$ is a variable size diagonal covariance matrix for source *i* (i.e., the number of trials over which a noise source is stationary

Manuscript received July 22, 2008; revised January 28, 2009 and March 9, 2009. First published April 28, 2009; current version published July 15, 2009. This work was supported by The Capita Foundation. *Asterisk indicates corresponding author.*

The author is with the Northeastern University, Boston, MA 02115 USA (e-mail: silva.i@neu.edu).

Digital Object Identifier 10.1109/TBME.2009.2021400

is variable), *I* is the identity matrix, and the $\hat{P}$ is the estimated number of distinct, locally stationary noise sources.

A. Nonstationary SNR Estimation Under Normal Averaging

Under normal ensemble averaging conditions the *M*-by-1 weight vector in (3) is held constant across trials and set to

$$\bar{w}_{\rm Nave}^T = \begin{bmatrix} \frac{1}{M} & \frac{1}{M} & \dots & \frac{1}{M} \end{bmatrix} \quad (6a)$$

$$\hat{s}_{\rm Nave}(t) = \bar{w}_{\rm Nave}^T X(t). \quad (6b)$$

Under the assumption of a deterministic signal and locally stationary noise sources, we can use (4) and (5) to simplify the SNR estimate in (3) to

$$\text{SNR}_{\rm Nave} = \frac{\bar{w}_{\rm Nave}^T R_s \bar{w}_{\rm Nave}}{\bar{w}_{\rm Nave}^T R_\eta \bar{w}_{\rm Nave}} = \frac{\sigma_s^2 \left(\sum_{i=1}^{M} \frac{1}{M}\right)^2}{\sigma_{\rm Nres}^2} = \frac{\sigma_s^2}{\sigma_{\rm Nres}^2} \quad (7)$$

where the residual noise power $\sigma_{\rm Nres}^2$ from $\hat{P}$ locally stationary and independent noise sources are given by [using (5)]

$$\sigma_{\rm Nres}^2 = \bar{w}_{\rm Nave}^T R_\eta \bar{w}_{\rm Nave} = \frac{1}{M^2}\text{tr}(R_\eta) = \frac{1}{M^2}\sum_{i=1}^{\hat{P}} \hat{M}_i \sigma_{\eta i}^2 \quad \text{with} \sum_{i=1}^{\hat{P}} \hat{M}_i = M \quad (8)$$

where $\hat{M}_i$ is the number of trials collected under the noise source power $\sigma_{\eta i}^2$ (note that $\hat{M}_i$ is random under evoked response experiments and is estimated from a segmentation procedure) and $tr(\,)$ is the trace operator. Equation (8) reduces to the classical $1/M$ reduction in the noise power under normal averaging if all $\hat{P}$ noise sources are equal, when this is the case (8) also yields the first term of (56) proposed in [13] for a single stationary noise source independent of the signal (in this paper Bendat set the signal to be nonstationary and the noise to be stationary). The SNR estimation under standard averaging can be done by using the final average and an estimate of the residual noise power

$$\hat{\text{SNR}}_{\rm Nave} = \frac{\hat{\sigma}_s^2}{\hat{\sigma}_{\rm Nres}^2} = \frac{\sigma_{\rm Nave}^2}{\hat{\sigma}_{\rm Nres}^2} - 1 \quad (9)$$

with

$$\sigma_{\rm Nave}^2 \equiv \text{var}(\hat{s}_{\rm Nave}(t)) \quad (10)$$

where $\sigma_{\rm Nave}^2$ is the variance of the averaged trials. The estimation of the residual noise power $\hat{\sigma}_{\rm Nres}^2$ in (9) requires segmentation of trials into stationary regions and the estimation of the noise variances within these regions. Assuming that the trials have been segmented into $\hat{P}$ sections, the locally stationary noise sources power $\sigma_{\eta i}^2$ can be estimated in the same manner as that proposed by [11] for weighted (Bayesian) averaging. The noise variance is estimated by selecting *L* fixed points with respect to the stimulus onset time and measuring the variability of these points across $\hat{M}_i$ trials

$$\hat{\sigma}_{\eta i}^2 = \frac{1}{L}\sum_{n=0}^{L-1} \frac{1}{M_i - 1} \sum_{m=k_i}^{k_i+\hat{M}_i} (x(n\Delta, m) - \overline{x(n\Delta)})^2 \quad (11a)$$

$$\hat{\sigma}_{\rm Nres}^2 = \frac{1}{M^2}\sum_{i=1}^{\hat{P}} \hat{M}_i \hat{\sigma}_{\eta i}^2 \text{ with } \sum_{i=1}^{\hat{P}} \hat{M}_i = M \quad (11b)$$

where, k_i is number of the first trial collected under the noise source power $\hat{\sigma}_{\eta i}^2$, $\overline{x(n\Delta)}$ is average value of a fixed point within a set of trials, and Δ is an integer corresponding to the number of samples such that the fixed multiple points are sufficiently far apart to be considered independent and uncorrelated. Points that are close together will be inherently correlated due to filtering and the autoregressive nature of the noise and signal as well as the sampling rate, which can result in overestimation of noise variance [10], [11], [14], [15].

B. Segmentation of Noise Sources

The total number of independent noise sources $\hat{P}$ in (8) can be estimated recursively by a series of *F*-tests. The current noise variance $\hat{\sigma}_{\eta i}^2$ is estimated by using (11a) and a parameter that determines the minimum number of trials $M_{\min}$ for all sources. The ratio between the current and previous noise variances are then used in an *F*-test to determine the likelihood that the current noise power lies within a preset confidence interval of the previous noise power

$$F(D_{\eta(i-1)}, D_{\eta i}) = \frac{\hat{\sigma}_{\eta(i-1)}^2}{\hat{\sigma}_{\eta i}^2} \quad (12)$$

where the degrees of freedom of the statistic in (12) are related to the number of *L* fixed samples used to estimate the respective noise sources in (11a)

$$D_{\eta i} = LM_{\min} - 1 \qquad D_{\eta(i-1)} = L\hat{M}_{\eta(i-1)} - 1. \quad (13)$$

If the test statistic in (12) is within the confidence interval of the previous noise source power, the current estimate is updated using recursive averaging

$$\hat{\sigma}_{\eta i}^{2\prime} = \frac{Q\hat{\sigma}_{\eta(i-1)}^2 + \hat{\sigma}_{\eta i}^2}{Q+1} \qquad M'_{\eta i} = M_{\eta i} + M_{\eta(i-1)} \quad (14)$$

where *Q* is the total number of variance estimates used to generate the grand average variance estimate of the previous block (i.e., $Q = \hat{M}_{\eta(i-1)}/M_{\min}$). Segmenting the data in this manner allows the noise sources to have a dynamic region of stationarity (but with minimal segment size of $M_{\min}$). The confidence parameter *p* for the statistical test of (12) acts as the smoothing parameter of the segmentation. The extreme value of $p = 0$ yields maximum smoothing (all noise source are averaged to a single noise source, similar to current standard methods [5]) and a value of $p = 1$ yields the maximum segmentation ($\hat{P} = M/M_{\min}$).

C. Nonstationary SNR Estimation Under Weighted Averaging

Previous studies have shown that weighted (Bayesian) averaging can significantly improve the SNR of ABR evoked potentials compared to simple artifact rejection (where trials are discarded if a preset amplitude limit is reached) [10], [11], [16]. The *M*-by-1 weighting vector $\bar{w}_{\mathrm{Wave}}$ under the weighted average scheme is given by [17]

$$\bar{w}_{\mathrm{Wave}} = \frac{\hat{R}_\eta^{-1}\bar{1}}{\bar{1}^T\hat{R}_\eta^{-1}\bar{1}} \qquad \text{where } \bar{1}^T = [1 \quad 1 \quad \ldots \quad 1] \tag{15a}$$

$$\hat{s}_{\mathrm{Wave}}(t) = \bar{w}_{\mathrm{Wave}}^T X(t). \tag{15b}$$

Using (15) and the covariance for a deterministic signal (4), the estimated SNR from (3) becomes

$$\mathrm{SNR}_{\mathrm{Wave}} = \frac{\bar{w}_{\mathrm{Wave}}^T R_s \bar{w}_{\mathrm{Wave}}}{\bar{w}_{\mathrm{Wave}}^T R_\eta \bar{w}_{\mathrm{Wave}}} = \frac{\sigma_s^2\left(\sum_{i=1}^{M} w_{\mathrm{Wave}}(i)\right)^2}{\sigma_{\mathrm{Wres}}^2} = \frac{\sigma_s^2}{\sigma_{\mathrm{Wres}}^2} \tag{16a}$$

$$\sigma_{\mathrm{Wres}}^2 = \bar{w}_{\mathrm{Wave}}^T R_\eta \bar{w}_{\mathrm{Wave}} = \frac{\bar{1}^T\hat{R}_\eta^{-1}R_\eta\hat{R}_\eta^{-1}\bar{1}}{(\bar{1}^T\hat{R}_\eta^{-1}\bar{1})^2} \tag{16b}$$

where the last equality in (16a) follows from the normalization constraint of the weights in (15a).

From (16b), we see that in the case of exact estimation of the noise covariance matrix R_η the SNR under weighted averaging is related to the signal power by

$$\mathrm{SNR}_{\mathrm{Wave}}^{\mathrm{opt}} = \sigma_s^2\frac{(\bar{1}^T R_\eta^{-1}\bar{1})^2}{\bar{1}^T R_\eta^{-1}R_\eta R_\eta^{-1}\bar{1}} = \sigma_s^2(\bar{1}^T R_\eta^{-1}\bar{1}) = \sigma_s^2\mathrm{tr}(R_\eta^{-1}) \tag{17}$$

where the last equality follows from (5). Thus, the optimal SNR under weighted averaging will increase proportionally to the sum of the inverse of the power of the noise sources (if the environment is stationary, that value converges to a $1/M$ reduction in noise variance). For a fixed set of known trials (i.e., offline processing), where the trial weights are known ahead of time, the SNR curve in (17) is a monotonically increasing function of trial (see Appendix A). With an exact estimate of the noise covariance matrix, weighted averaging is always superior to normal averaging (with the exception of stationary noise conditions, in which case both techniques are equivalent) (see Appendix B).

One method to estimate the diagonal elements $\hat{\sigma}_{\eta i}^2$ of the noise covariance matrix is through the fixed multiple point method from (11a). The diagonal elements of R_η are estimated from (11a) and then averaged across statistically similar trials by using the segmentation procedure (14) (this assumes the noise is stationary for at least $M_{\min}$ trials). The estimated SNR under weighted averaging then becomes

$$\hat{\mathrm{SNR}}_{\mathrm{Wave}} = \frac{\hat{\sigma}_{\mathrm{Wave}}^2}{\hat{\sigma}_{\mathrm{Wres}}^2} - 1 \tag{18a}$$

$$\hat{\sigma}_{\mathrm{Wave}}^2 = \mathrm{var}(\hat{s}_{\mathrm{Wave}}) \tag{18b}$$

$$\hat{\sigma}_{\mathrm{Wres}}^2 \approx \frac{1}{\bar{1}^T\hat{R}_\eta^{-1}\bar{1}} = \left(\sum_{i=1}^{\hat{P}}\frac{\hat{M}_i}{\hat{\sigma}_{\eta i}^2}\right)^{-1}. \tag{18c}$$

D. Predicting Number of Trials Required for Given Residual Noise Level

An issue of practical interest is the ability to predict how many trials are necessary to achieve a minimum postaverage residual noise level (this also applies to procedures that use a subaverage as a reference signal for adaptive noise cancellation). Under nonstationary noise sources, one way to make this prediction is by assuming that the current noise source will remain stationary for an infinite period ($\hat{M}_{\hat{P}}$ is incremented by the number of trials to predict). Using (8) for forecasting, with *M* trials collected prior to the current noise source and θ trials under the current noise source with normal averaging, the background noise level $\mathrm{BN}(\theta)$ as a function of trial θ is given by

$$\mathrm{BN}(\theta) = \sum_{i=1}^{\hat{P}-1}\frac{\hat{M}_{i-1}}{(M+\theta)^2}\hat{\sigma}_{\eta(i-1)}^2 + \frac{\theta}{(M+\theta)^2}\hat{\sigma}_{\eta\hat{P}}^2 = \frac{c+\theta\hat{\sigma}_{\eta\hat{P}}^2}{(M+\theta)^2} \tag{19a}$$

where the constant *c* is given by

$$c = \sum_{i=1}^{\hat{P}-1}\hat{M}_{i-1}\hat{\sigma}_{\eta(i-1)}^2. \tag{19b}$$

The solution to find the number of trials θ to reach a given residual background power level $\mathrm{BN}_{\mathrm{TH}}$ is obtained by applying the quadratic equation to (19)

$$\theta = \frac{\frac{\sigma_{\hat{P}}^2}{\mathrm{BN}_{\mathrm{TH}}} - 2M \pm \frac{1}{\mathrm{BN}_{\mathrm{TH}}}\sqrt{\hat{\sigma}_{\hat{P}}^4 - 4\mathrm{BN}_{\mathrm{TH}}(M\hat{\sigma}_{\hat{P}}^2 - c)}}{2}. \tag{20}$$

A negative or zero value for θ means that the current residual background power is already lower than the desired $\mathrm{BN}_{\mathrm{TH}}$. Under a single stationary noise source, $c = M\hat{\sigma}_{\hat{P}}^2$ in (20), and the results converge to the classical linear form ($\theta = (\sigma_{\hat{P}}^2/\mathrm{BN}_{\mathrm{TH}}) - M$).

Determining the number of trials necessary to reach a previous residual power level once an artifact perturbation has occurred might also be of particular interest to clinicians (i.e., $\mathrm{BN}_{\mathrm{TH}} = \frac{\sigma_1^2}{M}, \hat{\sigma}_{\hat{P}}^2 = \sigma_2^2, c = M\sigma_1^2$). Equation (20) with two noise sources then becomes

$$\theta = M\left(\frac{\sigma_2^2}{\sigma_1^2} - 2\right). \tag{21}$$

From a practical point of view, (21) might be used as a rough estimate of (20) by replacing σ_1^2 with the residual noise power level from (11) after any transients have decayed. For the case of weighted averaging, a closed form solution is not so straightforward to obtain; however, a numerical approach can be implemented by using (18) with the assumption that the power of the current noise source $\hat{\sigma}_{\hat{P}}^2$ remains constant and incrementing $\hat{M}_{\hat{P}}$.

II. Methods

A. Introduction

The analysis of the quality of the different SNR estimators was done using two experiments. In the first part, an appropriate range for the smoothing parameter *p*, used by the segmentation procedure (the *F*-test significance level), was determined. The first part, also aimed at comparing the mean square error (mse) of the current standard ABR SNR estimator, the *Fsp* [2], with the proposed estimator, nonstationary fixed-Multiple-Point (*NS Fmp*), which consists of (9)–(11); both estimators ignored the offset by negative one in the equations cited. The second experiment compared the receiver operating curves (ROC) for these estimators based on real data (in similar manner to that proposed by Gentiletti-Faenze *et al*. [3]). The choice in deciding to compare the different SNR estimation in terms of detection parameters (ROC curves) is based on the importance that these SNR estimators have in the current hearing–screening procedures. For all experiments, the *Fsp* and weighted *Fsp* (*WFsp*) noise power estimation $\hat{\sigma}_n^2$ was updated using one fixed point on blocks of 256 trials in (11a), as originally suggested by [2]. The *Fsp* at any given trial *m* was calculated by

$$\mathrm{Fsp}(m) = \frac{m\mathrm{var}(\overline{x(tm)})}{\hat{\sigma}_n^2}. \tag{22}$$

The fixed point location was arbitrarily selected to be the first data point of each waveform. The noise estimation for *NS Fmp* and *WNS Fmp* was done using (11) with a minimum block length of 32 trials with 8 samples equally spaced per trial and with. For both the *NS Fmp* and *WNS Fmp*, the first point used was arbitrarily selected to be the first data point on each waveform (the same as that of the Fsp procedure), with subsequent points spaced by 50 samples apart. These values were chosen based on [8] as a result of a tradeoff between resolution of the block size, accuracy of estimation, and correlation between samples within a trial. The variances of the individual noise sources were updated accordingly whenever blocks of trials were merged, and the SNR was recalculated. The noise segmentation for both the *NS Fmp* and *WNS Fmp* was done with the smoothing parameter set to $p = 0.0005$. Notice that the *Fsp* the procedure is equivalent to using the *NS Fmp* procedure with the following parameters: number of points per waveform = 1, number of trials per block = 256, and smoothing parameter $p = 0$.

B. ABR Measurements

The data from the ABR measurements were from the Gentiletti-Faenze *et al*. [3] study (courtesy of the original authors). The ABR measurements were done in five normal listeners at 0, 20, 30, 40, 60, and 80 dB sound pressure level (SPL). Two records at each level were made (except at 0 SPL). The stimuli consisted of monaural rarefaction clicks at a stimulus rate of 17.5 Hz (with a 5 ms prestimulus period followed by a 15 ms recording). A total of 4000 trials were obtained at each level. The signal was sampled at 20 kHz and filtered between 100 and 3000 Hz. The electrodes were place on channels Cz-A2, Cz-A1, and Fz for the ground.

C. Experiment I: Mean Square Error of Residual Noise Estimation

In this experiment, the mse of the residual noise estimators from (3) and (5) were measured according to the following procedure. For each of the five listeners, an estimate of the noise variance on a trial-by-trial basis was obtained by measuring the variance for each trial under the 0 dB SPL condition. The estimated noise variances of the individual trials were used to generate Monte Carlo simulations of 50 different sets of experiments for each listener. The Monte Carlo simulation allowed a measurement of the average decrease in residual noise level in the postaverage waveform as a function of the number of trials being added, and under changes in noise variance similar to those observed from real data. The mse for (3) and (5) was measured by computing the mse between the residual noise estimation for each of the 50 simulations and the true simulated value.

D. Experiment II: ROC Analysis

The ROC analysis for the four procedures (*Fsp*, *WFsp*, *NS Fmp*, and *WNS Fmp)* was done in the same way as suggested by [3]. Briefly, the ABR measurements obtained at 0 SPL were treated as indicative of a hearing loss. The ROC curve for each level was calculated by treating the ABR signal generated under 0 SPL as the nosignal condition (H0) and the ABR signal generated under a higher level (20, 30, 40, 60, or 80 dB SPL) as the present signal (H1) condition. For example, if the *Fsp* yielded the following values under H0 (0 dB SPL): 1.93, 1.60, 1.93, and 1.5; with the following values for $\mathrm{H1}_{20\mathrm{SPL}}$ (20 dB SPL): 2.74, 1.90, 1.75, and 2.41; and the following values for $\mathrm{H1}_{30\mathrm{SPL}}$ (30 dB SPL): 6.51, 7.09, 6.51, and 6.77; then the ROC areas for the *Fsp* procedure at 20 and 30 dB SPL would be 0.75 and 1.00, respectively.

Hence the ROC area quantifies how well the procedures can discriminate between the ABR measurement at 0 dB SPL and ABR measurement obtained at higher levels. In order to generate more points for the ROC analysis, the data for each level were broken into three different segments. These three segments were assumed to be independent from each other and were permutated in six different ways in order to simulate six more "experiments" for each level. The performance of each estimator was quantified in terms of the area of the ROC curve (1 being the maximum and floor of 0.5 equivalent to chance). The median ROC area for all levels (20, 30, 40, 60, and 80 dB SPL) was then calculated to yield a final value of performance; the median was used instead of the mean because of ceiling effects occurring.

The ROC performance of all four procedures was examined under two parameters: artifact rejection threshold level and a minimum number of accepted trials required to start the estimation procedure. The artifact rejection parameter had two values: 10 and ∞ μV (the second value being equivalent to not using artifact rejection and accepting all trials). The minimum

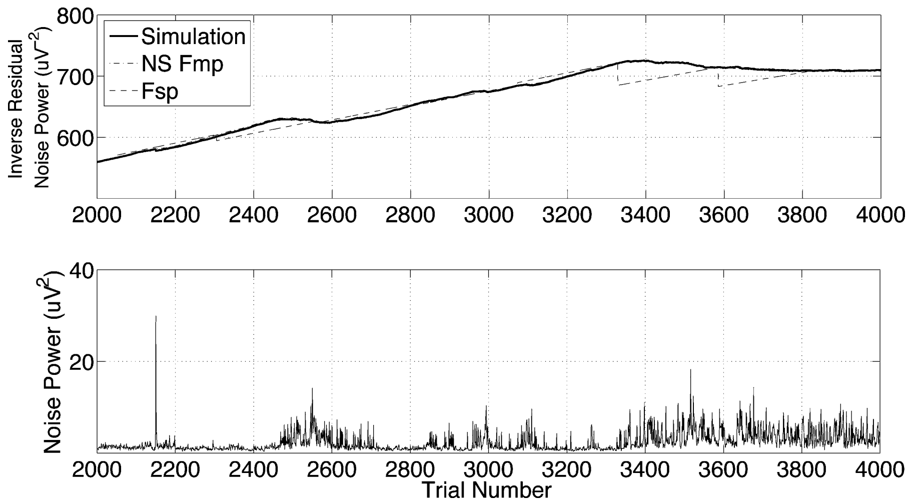

Fig. 1. Example of residual noise estimation. The top graph shows the estimated inverse residual noise as a function of number of trials added to the average. There is very close agreement between the residual noise estimation in the *NS Fmp* procedure and the average results from the Monte Carlo simulation, making their distinction difficult. The bottom graph shows the background noise power present in each of the individual trials used for the simulation. For this particular listener, the background power is clearly nonstationary (with changes in noise power of more than one order of magnitude).

number of accepted trials was varied from 256 to 3840 trials in step of 256 trials. This parameter determines the minimum number of accepted trials that the procedure should have prior to start estimating the SNR. Controlling the minimum number of accepted trials prior to any SNR estimation in a procedure allows for minimization of any initial transients affects, as observed by [3]. The final SNR estimate value was taken as the maximum value on the SNR curve between the minimum number of accepted trials and the final accepted trial.

III. Results

A. Experiment I: Mean Square Error of Residual Noise Estimation

An example of the simulations based on real data is shown in Fig. 1. The top graph shows the average residual noise level as a function number of trials from the Monte Carlo simulation and the results from the *Fsp* and the *NS Fmp*. The bottom graph shows the estimated noise variance from each trial that was used to generate the Monte Carlo simulations. Fig. 2 shows the mse for the residual noise estimation in the *Fsp* and *NS Fmp* as a function of the smoothing segmentation parameter p. The mse of the *Fsp* procedure, which assumes a single noise source, is constant as a function of this parameter. The *Fsp* yielded a higher mse than did the *NS Fmp*.

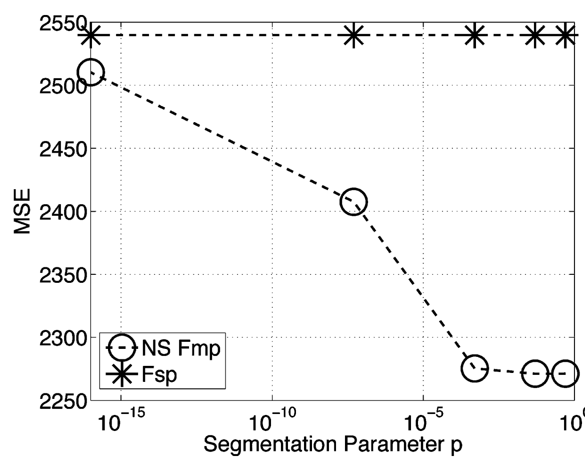

Fig. 2. Mean square error between the estimated residual noise and the Monte Carlo simulations as a function of the segmentation parameter p. The mse of the *Fsp* procedure is plotted for reference. The mse of the *NS Fmp* procedure is a monotonic function of p. As p approaches 0, the number of noise sources estimated decrease, with *NS Fmp* algorithm being equivalent to the *Fsp* at $p = 0$.

B. Experiment II: ROC Analysis

The results of the ROC analysis are shown in Figs. 3 and 4 and Table I. Fig. 3 shows the performance of all procedures as a function of minimum number of accepted trials and artifact rejection threshold. In general, the weighted averaging procedures (*WFsp* and *WNS Fmp*) was superior the normal averaging procedures on both artifact rejection levels. In similar fashion, the *NS Fmp* was superior to the *Fsp* on both artifact rejection levels as well. The performance of all procedures increases as the artifact rejection level was lowered (at the cost of more trials being discarded) and as the minimum number of accepted trials required for estimation increased. Although the performance for the weighted averaging procedures seem identical, a more careful look at Fig. 3 and Table I shows that the *WNS Fmp* procedure has a small but consistent advantage when the minimum number of accepted trials $\leq$1000 for either rejection levels.

IV. Discussion

The results from Experiment I show that changes in the ABR background noise power can be of more than an order of magnitude (Fig. 1 bottom graph). This corroborates previous literature on the nonstationary behavior of the ABR background noise and

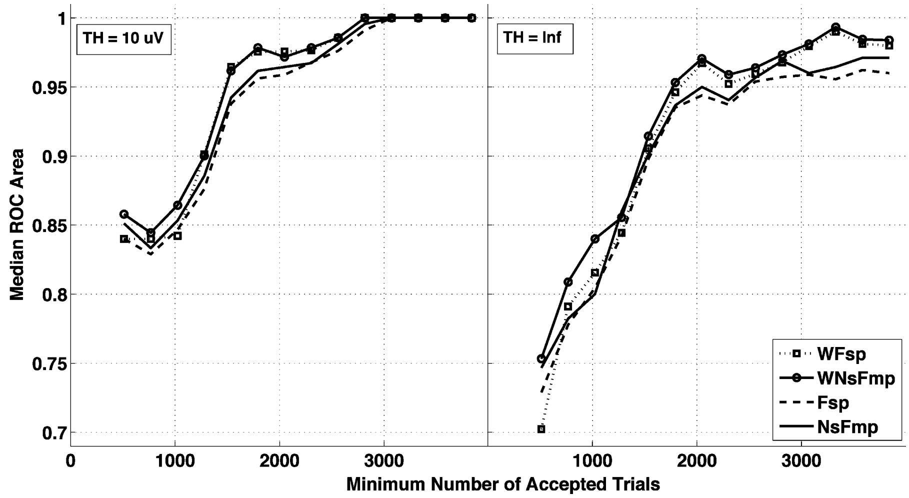

Fig. 3. Median ROC area for all four procedures as a function of a minimum number of accepted trials before any estimation begins, and artifact rejection level. The minimum number of accepted trials attempts to control for initial transient effects in the estimation of the SNR curve.

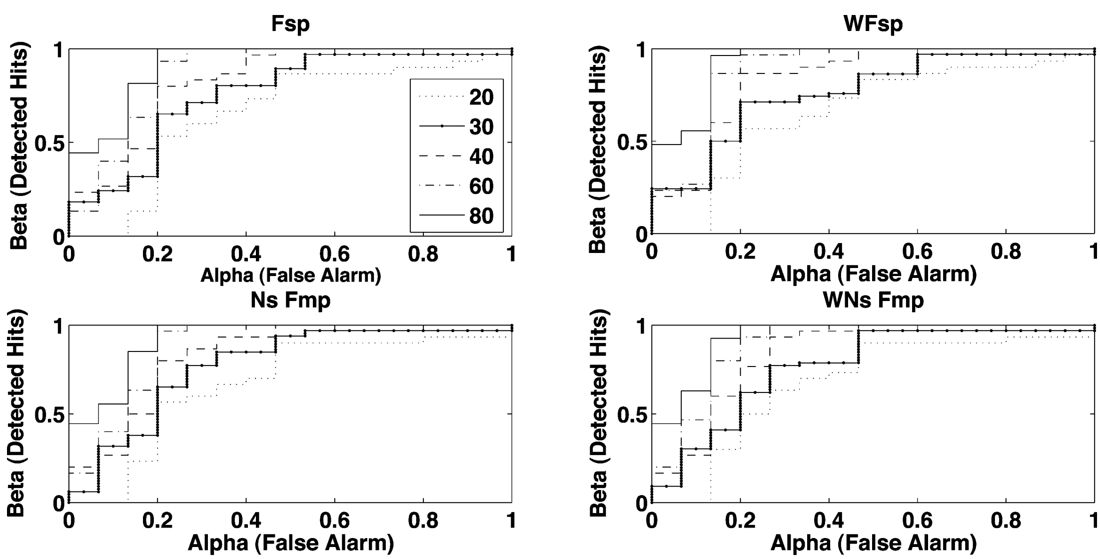

Fig. 4. ROC curves for all four procedures with artifact rejection level set to 10 μVand minimum number of accepted trials required for estimation set to 1024.

TABLE I
MEDIAN ROC AREA AS A FUNCTION PROCEDURE, ARTIFACT REJECTION LEVEL (T_h), AND MINIMUM NUMBER OF ACCEPTED TRIALS (M)

VI. PROCED URE	VII. TH=10,M=5 12	VIII. TH=10,M=1 024	**Th=Inf,M=512**	**Th=Inf,M=1024**
IX. FSP	84.0%	84.6%	72.8%	80.4%
X. NS FMP	85.1%	84.2%	74.6%	80.0%
XI. WFSP	84.0%	85.3%	70.2%	81.5%
XII. WNS FMP	85.7%	86.4%	75.3%	84.0%

the necessity of accounting for residual noise levels while analyzing the averaged data. Computer simulations of averaging under nonstationary noise sources show that the mse of the SNR curve obtained under a standard procedure, the *Fsp*, is higher than the *NS Fmp* (which can account for differences in noise power across trials). The *NS Fmp*'s smoothing parameter *p* can be adjusted in such a way that the *NS Fmp* performance will asymptote to that of the *Fsp*.

In the second experiment, the performance of all procedures was examined as function of artifact rejection threshold and a minimum number of accepted trials required for estimation. On average, the *WNS Fmp* procedure had a small but consistent overall best performance. The *WNS Fmp* advantage was greatest at small values for the minimum required number of accepted trials ($\leq$1000), which is precisely where transients are more likely to have an effect. In general, as the number of accepted trials increase the estimation of the signal power component of the SNR (10) or (18b) converges, thus becoming robust to background noise perturbations and improving the overall SNR estimation. Therefore, as the minimum number of trials required for estimation increased, all procedures showed a general increase in performance regardless of the artifact rejection level. When the artifact rejection level is taken into account, all procedures showed improvement at lower artifact rejection level (with all performances converging at high numbers of minimum accepted trials). The biggest difference between performances and across threshold levels occur for small values of minimum required number of accepted trials, although differences at both extremes are clearly visible. From Table I, we see that at a value of 512 for the number of minimum accepted trials the average performance across procedures was 84.7% for a threshold of 10 μV, this average performance drops to 73.2% when no threshold is applied ($\infty\,\mu$V). As the number of trials increase to 1024, the estimation becomes more robust to perturbations and the drop is less substantial (from 85.1% at 10 μV to 81.4% at $\infty\,\mu$V).

V. Conclusion

This paper proposes two new methods (*NS Fmp* and *WNS Fmp*) that deals with nonstationary noise sources based on a modification to a widely used procedure (*Fsp*) for estimating SNR. The new method attempts to account for the nonstationary background by assuming $\hat{P}$ discrete locally stationary noise sources. A noise segmentation procedure is also introduced that dynamically partitions the background noise into discrete segments of different noise power based on a series of *F*-tests. The combined *NS Fmp* and segmentation parameters values can be set to result in the standard *Fsp* estimation. Further modifications to the *NS Fmp* were also done to account for weighted averaging (*WNS Fmp*).

Appendix A

The final (i.e., offline) SNR curve as a function of trial under weighted averaging is monotonically increasing. Using (17), we rewrite the SNR as a function of trial in piecewise syntax shown in (A1) at the bottom of this page.

From (A5) it is clear that

$$\mathrm{SNR}^{\mathrm{opt}}_{\mathrm{Wave}}(\theta) = \mathrm{SNR}^{\mathrm{opt}}_{\mathrm{Wave}}(\theta-1) + \frac{\sigma_s^2(\theta - \sum_{i=1}^{M-1}\hat{M}_i)}{\sigma_{i\eta}^2} > \mathrm{SNR}^{\mathrm{opt}}_{\mathrm{Wave}}(\theta-1). \tag{A2}$$

Appendix B

If we define the ratio between the optimal SNR under weighted averaging (17) and the SNR under normal averaging (7) we see that

$$\delta = \frac{\mathrm{SNR}^{\mathrm{opt}}_{\mathrm{Wave}}}{\mathrm{SNR}_{\mathrm{Nave}}} = \left(\frac{\mathrm{tr}(R_\eta)}{\sigma_s^2 M^2}\right)\sigma_s^2\,\mathrm{tr}(R_\eta^{-1}) = \frac{\mathrm{tr}(R_\eta)\mathrm{tr}(R_\eta^{-1})}{M^2} \geq 1 \tag{B1}$$

with the assumption that the noise covariance matrix is diagonal with positive elements than the trace operator is related to the 2-norm of a vector whose elements are the square root of the diagonal elements of the covariance matrix

$$\mathrm{tr}(R_\eta) = \bar{\sigma}_\eta^T\bar{\sigma}_\eta \tag{B2}$$

where $\bar{\sigma}_\eta^T = [\sigma_{\eta 1}, \sigma_{\eta 2}, \ldots, \sigma_{\eta M}]$, and $\bar{\sigma}_\eta^T\bar{\sigma}_\eta$ is the vector dot product. Using Cauchy-Schwarz's theorem, we can then derive a lower bound to (B1)

$$\left(\bar{\sigma}_\eta^T\bar{\sigma}_\eta\right)\left(\left(\bar{\sigma}_\eta^{-1}\right)^T\bar{\sigma}_\eta^{-1}\right) \geq \left|\bar{\sigma}_\eta^T\bar{\sigma}_\eta^{-1}\right|^2 \tag{B3}$$

that from (B2) and (B3) we have,

$$\mathrm{tr}(R)\mathrm{tr}(R^{-1}) \geq M^2. \tag{B4}$$

Using (B4) in the numerator of (B1) and canceling the *M* terms, we get the final inequality.

Acknowledgment

The author would like to thank G. G. Gentiletti-Faenze, O. Yanez-Suarez, and J. M. Cornejo-Cruz for providing him with their auditory brainstem response (ABR) data. The author

$$\mathrm{SNR}^{\mathrm{opt}}_{\mathrm{Wave}}(\theta) = \begin{cases} \dfrac{\sigma_s^2\theta}{\sigma_{1\eta}^2} & \text{for } 1 \leq \theta \leq \hat{M}_1 \\ \sigma_s^2\left(\dfrac{\hat{M}_1}{\sigma_{1\eta}^2} + \dfrac{\theta - \hat{M}_1}{\sigma_{2\eta}^2}\right) & \text{for } \hat{M}_1 < \theta \leq \hat{M}_1 + \hat{M}_2 \\ \sigma_s^2\left(\displaystyle\sum_i \frac{\hat{M}_i}{\sigma_{i\eta}^2} + \frac{\theta - \sum_{i=1}^{M-1}\hat{M}_i}{\sigma_{\hat{P}\eta}^2}\right) & \text{for } \displaystyle\sum_{i=1}^{M-1}\hat{M}_i < \theta \leq \sum_{i=1}^{\hat{P}}\hat{M}_i \end{cases} \tag{A1}$$

is also thankful to Dr. Michael Epstein, Dr. Dana Brooks, and Scott Hemphill for valuable communications and the Editor and two anonymous reviewers for helpful comments.

REFERENCES

[1] C. Elberling and M. Don, "Detection function for the human auditory brainstem response," *Scandinavian Audiol.*, vol. 16, pp. 89–92, 1987.

[2] M. Don, C. Elberling, and M. Waring, "Objective detection of averaged brainstem responses," *Scandinavian Audiol.*, vol. 13, pp. 219–228, 1984.

[3] G. G. Gentiletti-Faenze, O. Yanez-Suarez, and J. M. Cornejo-Cruz, "Evaluation of automatic identification algorithm for auditory brainstem response used in universal hearing loss screening," in *Proc. 25th Annu. Int. Conf. IEEE EMBS*, Cancun, Mexico, 2003, pp. 2857–2860.

[4] M. Don and C. Elberling, "Use of quantitative measures of auditory brainstem response peak amplitude and residual background noise in the decision to stop averaging," *J. Acoust. Soc. Amer.*, vol. 99, pp. 491–499, 1996.

[5] C. Elberling and M. Don, "Quality estimation of averaged auditory brainstem responses," *Scandinavian Audiol.*, vol. 13, pp. 187–197, 1984.

[6] A. Jacquin, E. Causevic, E. R. John, and L. S. Prichep, "Optimal denoising of brainstem auditory evoked response (BAER) for automatic peak identification and brainstem assessment," in *Proc. 28th IEEE EMBS Annu. Int. Conf.*, New York City, 2006, pp. 1723–1726.

[7] Y. S. Sininger, "Filtering and spectral characteristics of averaged auditory brain-stem response and background noise in infants," *J. Acoust. Soc. Amer.*, vol. 98, pp. 2048–2055, 1995.

[8] W. Qiu, F. H. Y. Chan, F. K. Lam, M. D. Noh, M. A. Howard, P. C. Garell, I. O. Volkov, R. A. Reale, J. E. Hind, and J. F. Brugge, "An adaptive approach for processing evoked potentials from the auditory cortex of man," in *Proc. 20th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 1998, pp. 1645–1648.

[9] F. H. Y. Chan, F. K. Lam, P. W. F. Poon, and W. Qiu, "Detection of brainstem auditory evoked potential by adaptive filtering," *Med. Biol. Eng. Comput.*, vol. 33, pp. 69–75, 1995.

[10] M. Don and C. Elberling, "Evaluating residual background noise in human auditory brain-stem responses," *J. Acoust. Soc. Amer.*, vol. 96, pp. 2746–2757, 1994.

[11] C. Elberling and O. Wahlgreen, "Estimation of auditory brainstem response, ABR, by means of bayesian inference," *Scandinavian Audiol.*, vol. 14, pp. 89–96, 1985.

[12] C. E. Davila and M. S. Mobin, "Weighted averaging of evoked potentials," *IEEE Trans. Biomed. Eng.*, vol. 39, no. 4, pp. 338–345, Apr. 1992.

[13] J. S. Bendat, "Mathematical analysis of average response values for nonstationary data," *IEEE Trans. Biomed. Eng.*, vol. BME-11, no. 3, pp. 72–81, Jul. 1964.

[14] J. A. McEwen and G. B. Anderson, "Modeling the stationarity and gaussianity of spontaneous electroencephalographic activity," *IEEE Trans. Biomed. Eng.*, vol. BME-22, no. 5, pp. 361–369, Sep. 1975.

[15] B. Lutkenhoner, M. Hoke, and C. Pantev, "Possibilities and limitations of weighted averaging," *Biol. Cybern.*, vol. 52, pp. 409–416, 1985.

[16] M. Hoke, B. Ross, B. Widkesberg, and B. Lutkenhoner, "Weighted averaging- theory and application to electric response audiometry," *Electroencephalography Clin. Neurophysiol.*, vol. 57, pp. 484–489, 1984.

[17] L. Sörnmo and P. Laguna, *Bioelectrical Signal Processing in Cardiac and Neurological Applications*. New York: Academic, 2005.

Ikaro Silva (M'08) received the B.S. degree in biomedical engineering from Boston University, Boston, in 2000, and the M.S. degree in electrical engineering from Northeastern University, Boston, in 2004. He is currently working toward the Ph.D. degree in electrical engineering at Northeastern University.

From 2000 to 2002, he was with The MathWorks, Natick, MA, His current research interests include biosignal processing, statistical signal processing, detection & estimation theory, auditory processing, hearing aid technology, and modeling of speech perception.