Subspace Averaging of Steady-State Visual Evoked Potentials [1]

Carlos E. Davila* and Richard Srebro†

***Electrical Engineering Department, Southern Methodist University**
Dallas, Texas 75275-0338, †Department of Ophthalmology, University of Texas
Southwestern Medical Center, Dallas, Texas 75235-8592

E-mail: cd@seas.smu.edu
Phone: (214) 768-3197
Fax: (214) 768-3573

Abstract

A new algorithm for doing signal averaging of steady-state visual evoked potentials (VEP's) is described. The *subspace average* is obtained by finding the orthogonal projection of the VEP measurement vector onto the signal subspace, which is based on a sinusoidal VEP signal model. The subspace average is seen to out-perform the conventional average using a new SNR-based performance measure on simulated and actual VEP data.

Submitted to *IEEE Transactions on Biomedical Engineering*, July 1998

[1]**This work was supported in part by a the National Science Foundation (BCS-9308028), and an unrestricted research grant from Research to Prevent Blindness, Inc. N.Y.**

I. Background

Steady-state visual evoked potentials (VEPs) are used in a variety of clinical applications. These include estimation of optic nerve function, estimation of visual acuity in infants, young children and adults unable to provide reliable verbal responses, detection of hysteria and malingering, assessment of amblyopia and strabismus, diagnosis of cortical blindness, assessment of delayed neurological maturation, and assessment of abnormal optic tract decussation in albinism [1, 2]. There are some instances, particularly in pediatric VEP measurements, where it is imperative that acquisition times be minimized.

Signal averaging is the most commonly used method for estimating the brain evoked potential (EP) [3, 4, 5]. The EP is measured by presenting a series of short duration sensory stimuli, and recording the EEG immediately after the presentation of the stimulus for M samples. Let $x_k = [x[kM] \;\; x[kM+1] \;\; \cdots \;\; x[(k+1)M-1]]$, $k \geq 0$, called a "trial", represent the measurement of the scalp signal, $x[n]$ taken after the k^{th} stimulus. Then the conventional average is given by

$$\overline{x}_N = \frac{1}{N} \sum_{k=1}^{N} x_k \tag{1}$$

Assume that the k^{th} trial is modeled as $x_k = s + z_k$ where z_k is the additive noise due to spontaneous EEG. If one further assumes that the noise component z_k is uncorrelated and zero-mean across trials, is uncorrelated with the signal, and that the power of the signal and noise terms, σ_s^2 and σ_z^2, respectively, is constant across trials then it can be shown that the signal-to-noise ratio (SNR) of $\overline{x}_N$ is given by

$$\frac{N\sigma_s^2}{\sigma_z^2} \tag{2}$$

There is evidence that the above assumptions are not entirely valid [6], nevertheless, the model and simplifying assumptions provide some useful insights into the convergence of signal averaged EP data.

Over the past several decades there have been numerous attempts at improving on the

conventional average. These include the Woody average wherein individual trials are time shifted to compensate for latency shifts assumed to occur uniformly over the entire trial [7]. Latency corrected averaging has also been proposed to compensate for non-uniform latency jitter occurring among the various components of the EP within a single trial [8, 9, 10]. Other approaches based on Wiener filtering [11, 12, 13, 14] and time-varying Wiener filtering [15, 16] have also been proposed . The Wiener filter approach is based on a minimum mean-squared error criterion. Other estimators based on parametric methods have also appeared [17, 18, 19, 20, 21, 22, 23]. In order to accommodate the nonstationary nature of the evoked potential, adaptive filters have been used [24, 25, 26, 27, 28, 29, 30]. Wavelet analysis has also been applied to the evoked potential [31, 32]. All of the above methods dealt with estimation or analysis of transient evoked responses. Relatively few papers have appeared on estimation of steady-state VEP's [33, 34]. We make a distinction here between estimation of steady-state VEP waveforms and detection of VEP waveforms. The body of literature for detection is more substantial and will not be considered here. A recent survey of VEP detection algorithms is found in [35]. The method in [33] used analog comb filters to estimate the steady-state VEP while the method in [34] used an adaptive line enhancer. Unfortunately, it is difficult to quantify the performance of either of these two estimation methods. In this paper a new approach to signal averaging of steady-state VEP's is described which pre-processes each trial in order to increase it's SNR prior to computing the average. This has the effect of increasing the convergence speed of the ensemble average. Moreover, we shall develop expressions which enable us to evaluate the performance of the new method relative to the conventional average as the average is being computed.

II. Subspace Averaging

We assume the following signal model for the steady-state VEP, derived via counterphase modulated contrast gratings

$$x[n] = \sum_{k=1}^{r} \alpha_k cos(\omega_k n - \phi_k) + z[n] \tag{3}$$

The signal is a sum of d sinusoids having frequencies $\omega_m = m2\pi/M, m = 1, \ldots, d$ and constant amplitudes and phases $\alpha_m, \phi_m, m = 1, \ldots, d$, respectively. The integer M is the inter-stimulus interval as well as the length of each single trial and T is the sampling interval. Hence, the frequencies of the sinusoids consist of the contrast reversal frequency and its $d-1$ subsequent harmonics. The signal $z[n]$ constitutes the background EEG. This model of the steady-state VEP is well-established in the literature [3, 25]. Its justification is based on the fact that for steady-state VEPs, the visual stimulus is a periodic signal, and hence contains only sinusoidal frequency compenents at the fundamental frequency (ω_m) and its harmonics ($\omega_2, \ldots, \omega_d$). If we assume that the visual system is time invariant, the output of the visual system (the VEP signal component) will therefore also be periodic with sinusoidal components at the same frequencies (of course, the amplitudes and phases will differ). The frequencies contained in the VEP signal component will be the same regardless of whether the visual system is linear or nonlinear provided there is no nonlinear mixing (multiplication) between signal and noise, and the visual system is time invariant over the relatively short time intervals corresponding to the length of a single-trial. The stationarity of the EEG for periods of several hundred ms has been reported in the literature [36]. We are using single-trial epochs of about 500 ms, hence, the stationarity assumption is justified. On the nonlinear mixing, there is no clear evidence in the literature that we know of which suggests that there exists a nonlinear interaction between noise and signal components in the evoked potential, hence the additive noise assumption is made by virtually all evoked potential estimation algorithms. Figure 1 shows an estimate of the power spectral density of a steady-state VEP obtained using a counterphase modulated contrast grating showing the presence

of several harmonics of the contrast reversal frequency of 7.5 Hz.

Signal subspace processing has been widely applied to problems in array processing arising in telecommunications and radar [37, 38, 39]. Since the theory and practice of signal subspace processing is well understood, we give only an abbreviated description of it here. A more detailed treatments of this topic can be found in [37]. The basic idea behind subspace processing is that if the signal measurement vector is known to exist within a low dimensional subspace of Euclidean vector space, called a *signal subspace*, the SNR of a measurement vector containing the signal can be increased by computing the orthogonal projection of the measurement vector onto the signal subspace. By orthogonal projection, we mean the approximation of an $M \times 1$ vector x by another vector, $\hat{x}$, formed by taking a linear combination of linearly independent vectors, $s_1, s_2, \ldots, s_d$,

$$\hat{x} = \sum_{k=1}^{d} \alpha_k s_k \tag{4}$$

such that the squared Euclidean norm, $\|\hat{x} - x\|^2$ is minimized. We say that $\hat{x}$ is the *orthogonal projection* of x onto the subspace spanned by the vectors $s_1, s_2, \ldots, s_d$. Moreover, the orthogonal projection is given by [40]

$$\hat{y} = P_S y \tag{5}$$

where $P_S = S\left(S^T S\right)^{-1} S^T$ is called a projection matrix and $S = \left[\begin{array}{cccc} s_1 & s_2 & \cdots & s_d \end{array}\right]$. To see why the orthogonal projection increases the SNR, we must assume that $x = s + z$, where s is an $M \times 1$ signal vector existing in a low-dimensional subspace and z is an $M \times 1$ white noise vector. If s lies in a low-dimensional subspace, then we can express s in terms of a basis for that subspace, $s = Sw$, where S is an $M \times d$ matrix whose columns are the signal subspace basis vectors and w is a $d \times 1$ random vector having the property $E\left[ww^T\right] = P$, where P is positive definite. We show below that projecting the vector x onto the subspace spanned by the columns of S will increase the SNR or x. The problem then is to determine S from the

data vector x. This can be done by noting that the autocorrelation matrix of x is given by

$$R_{xx} \equiv E\left[xx^T\right] = SPS^T + \sigma_z^2 I_M \tag{6}$$

and since adding a diagonal matrix to a symmetric matrix does not change the eigenvectors of that matrix, the eigenvectors of R_{xx} corresponding to its maximum d eigenvalues are the same as the eigenvectors corresponding to the d nonzero eigenvalues of SPS^T. Moreover, it can be shown that these d eigenvectors span the column space of S, i.e. the signal subspace. In practice, the signal subspace then can be estimated by computing the sample autocorrelation matrix $\hat{R}_{xx}$ and then finding the eigenvectors of $\hat{R}_{xx}$ corresponding the to its d maximum eigenvalues. Fortunately, for the sinusoidal signal model in (3), the signal subspace is known, it is spanned by the vectors [37]

$$\begin{aligned} s_m &= \left[\begin{array}{ccccc} 1 & cos(\omega_m) & cos(2\omega_m) & \cdots & cos((M-1)\omega_m) \end{array}\right]^T, \quad m = 1, 3, \ldots, 2r-1 \\ &= \left[\begin{array}{ccccc} 1 & sin(\omega_{m-1}) & sin(2\omega_{m-1}) & \cdots & sin((M-1)\omega_{m-1}) \end{array}\right]^T, \quad m = 2, 4, \ldots, 2r \end{aligned} \tag{7}$$

Hence, the signal subspace dimension is $d = 2r$. As mentioned earlier, for counterphase modulated grating stimuli, ω_1 is equal to the contrast (polarity) reversal frequency of the stimulus [3].

The subspace averaging algorithm is then given by:

1. Form the signal subspace projection matrix $P_S = SS^T$, where S is given by

$$\left[\begin{array}{cccccc} 1 & 1 & \cdots & 1 & 1 \\ cos(\omega_1) & sin(\omega_1) & \cdots & cos(\omega_r) & sin(\omega_r) \\ cos(2\omega_1) & sin(2\omega_1) & \cdots & cos((2\omega_r) & sin(2\omega_r) \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ cos((M-1)\omega_1) & sin((M-1)\omega_1) & \cdots & cos((M-1)\omega_r) & sin((M-1)\omega_r) \end{array}\right] \tag{8}$$

2. Acquire the steady-state VEP time series $x[n], \quad n = 1, \ldots, N$.

3. Pre-whiten $x[n]$ with a suitable linear prediction-error filter, the steps for carrying this out are:

   (a) Form the $p \times L$ sample autocorrelation sequence (via autocorreltion method)

$$\hat{r}_{xx}[k] = \frac{1}{N} \sum_{n=1}^{N-k} x[n]x[n+k], \quad k = 0, \ldots, p$$

   (b) Apply the Levinson algorithm to $\hat{r}_{xx}[k]$ to obtain the autoregressive parameter estimates $a[1], a[2], \ldots, a[p]$ [37].

   (c) Prewhiten $x[n]$ by filtering it with the FIR filter

$$w[n] = \begin{bmatrix} 1 & a[1] & a[2] & \cdots & a[p] \end{bmatrix}$$

$$x \leftarrow w[n] \star x[n]$$

4. for $k = 1, 2, \ldots$

   (a) Update the conventional ensemble average as

$$\overline{x}_k = \overline{x}_{k-1} + x_k \tag{9}$$

   where

$$x_k = \begin{bmatrix} x[kM] & x[kM+1] & \cdots & x[(k+1)M-1] \end{bmatrix}, k \geq 1 \tag{10}$$

   corresponds to the k^{th} single trial and $\overline{x}_0$ has been initialized to a zero vector.

   (b) Compute the subspace average as

$$\overline{x}_k^s = P_S \overline{x}_k \tag{11}$$

5. *end for*

6. If desired, the effects of prewhitening can be compensated for by multiplying the inverse

filter,

$$\frac{1}{A\left(e^{j\omega}\right)}$$

with the Fourier transform of $\overline{x}_k^s$, where $A\left(e^{j\omega}\right)$ is the Fourier transform of the inverse filter. Since linear prediction error filters are minimum phase, the inverse operation is stable.

The effect of prewhitening on steady-state VEP data can be seen in the power spectral density estimates before and after prewhitening in Figure 2. As is evident, prewhitening seems to have virtually no effect on the SNR of the data. A good prewhitener results by applying a linear prediction-error filter based on an autoregressive (AR) model for $z[n]$ as described in [35]. Another example of prewhitening in evoked potential signal processing is found in [23]. Prewhitening is used for several reasons, (1) it enables the use of a performance measure, $\rho(k)$, described below, which assumes that the additive noise is white. (2) Prewhitening makes it easier to validate the mathematical model used as a basis for this algorithm since one can simply look for sinusoidal frequency components in a flat noise background in data spectra. (3) We have been interested un using the VEP for measurement of visual acuity, which requires detecion of low-level VEP's. Hence, prewhitening is very useful for implementing signal detection algorithms [41].

A related method called the Adaptive Fourier Linear Combiner (AFLC) has previously been described in [25]. The AFLC differs from this method in its use of a Least Mean Square (LMS) algorithm for adjusting the coefficients of the linear combiner. In subspace averaging, the coefficients are, in effect, the quantity $\left(S^T S\right)^{-1} S^T x_i$, which corresponds to a least squares approximation of x_i by the signal subspace, hence the subspace projection will not exhibit the excess mean-squared error typical of LMS adaptive filter algorithms [30]. The idea of projecting onto a signal subspace for signal estimation is not new, it was evidently first proposed by Tufts and Kumaresan in [39]. The subspace projection approach described here also enables a simple performance analysis based on signal to noise ratio as discussed in the following section.

III. Performance Analysis

Next we develop expressions for the SNR of the subspace average and compare it to the SNR of the conventional average given above. We first assume that the signal and noise power in the single trial x_k is constant over all k. Since $x_k = s + z_k$ where s and z_k are the respective signal and noise components in x_k, the signal power after projecting a single trial onto the signal subspace is given by

$$\sigma_{ss}^2 = E\left[s^T P_S s\right] = s^T s_k = M\sigma_s^2 \tag{12}$$

where we have used the fact that projection matrices are idempotent ($P_S P_S = P_S$). The noise power, after projecting onto the signal subspace, is

$$\sigma_{sz}^2 = E\left[z_k^T P_S z_k\right] = \sigma_z^2 tr\left\{P_S\right\} = d\sigma_z^2 \tag{13}$$

where $tr\{\}$ denotes the matrix trace operation and $z[n]$ is assumed to have been pre-whitened. The fact that the trace of a projection matrix is equal to the dimension of the subspace onto which it projects was also used [40]. Hence, the SNR of the projected single trial is

$$\frac{\sigma_{ss}^2}{\sigma_{sz}^2} = \frac{M}{d}\frac{\sigma_s^2}{\sigma_z^2} \tag{14}$$

where $\frac{\sigma_s^2}{\sigma_z^2}$ is the SNR of the single trial. The SNR of a single trial after projecting onto the signal subspace is increased to the extent that $M/d > 1$. Typically, $M \gg d$ so one can anticipate a substantial increase in the single trial SNR. The SNR of the subspace average is then

$$N\frac{M}{d}\frac{\sigma_s^2}{\sigma_z^2} \tag{15}$$

where N is the number of trials in the average. At first glance, this would appear to be an unlikely result, given that the SNR can evidently be increased by simply increasing the length M of each trial. The reason for this is, that as M is increased, the noise component,

z_k tends to become more and more "orthogonal" to the basis vectors in the signal subspace. Hence less noise energy projects onto the signal subspace as M increases.

We note that for *any* vector s lying in the signal subspace, $P_S s = s$. The implication here is that if the signal amplitudes and phases change from trial-to-trial, the projected measurement vector still exhibits an increase in SNR, hence the assumption that the signal and noise powers are constant across trials was only necessary for purposes of developing formulas for performance analysis, however if these assumptions are violated, the subspace average will still have a higher SNR.

IV. A Performance Measure For Ensemble Averaging

Since the signal component is unavailable during ensemble averaging, it would be useful to have some measure of the quality of the ensemble average which can be computed while the average is being acquired. We propose the following error measure for comparing ensemble averages,

$$\tilde{\rho}(k) = \frac{\|\overline{x}_k - \overline{x}_{k-1}\|^2}{\|\overline{x}_k\|^2} \tag{16}$$

where $\overline{x}_k$ is the ensemble average of k trials as defined in (1), and $\| \ \|$ is the Euclidean vector norm (or vector 2-norm) [42]. If we assume that $s^T s = M\sigma_s^2$, $z_i^T z_j \approx 0, i \neq j$, $z_k^T z_k \approx M\sigma_z^2$, and $z_k^T s \approx 0$, then

$$\tilde{\rho}(k) = \frac{\sigma_z^2}{k(k-1)\left(\sigma_s^2 + \sigma_z^2/k\right)} \tag{17}$$

Now clearly the $k(k-1)$ factor in the denominator will cause the error measure to decrease with each single trial, and is therefore redundant, so we define the normalized error measure

$$\rho(k) = k(k-1)\tilde{\rho}(k) \tag{18}$$

which can be expressed as

$$\rho(k) = \frac{\sigma_z^2}{\sigma_s^2 + \sigma_z^2/k} = \frac{1}{\mu + 1/k} \tag{19}$$

where $\mu = \sigma_s^2/\sigma_z^2$ is the SNR of a single trial. Interestingly, $\rho(k)$ converges to the "noise-to-signal ratio" (NSR),

$$\lim_{k\to\infty} \rho(k) = \frac{\sigma_z^2}{\sigma_s^2} \tag{20}$$

This convergence is dependent on the SNR of the single trials, higher SNR's give faster convergence than lower SNR's. The significance of the measure $\rho(k)$ is that it provides a method of comparing the performance of different ensemble averaging methods by providing an estimate of the effective NSR of the single trials. Figure 3 shows a graph of $\rho(k)$ for several different SNR's. It can be seen that convergence to σ_z^2/σ_s^2 is faster for the higher SNR's. For SNR's below -20 dB, it can be seen that $\rho(k)$ has yet to converge after 200 trials. Note that for the -20 dB SNR, after 200 trials $\rho(k)$ has yet to even come close to its asymptotic value of 100. Figure 3 can effectively be used to calibrate corresponding curves for real data to estimate single-trial VEP SNR's. Also, since $\rho(k)$ converges to the noise to signal power ratio, for and SNR of 0, $\rho(k)$ will never converge to a fixed value.

V. Simulations

The steady-state VEP was simulated with a 21,600-point time series as

$$x[n] = \sum_{m=1}^{8} \alpha_k cos(\omega_m(n-1)T - \phi_m) + z[n] \tag{21}$$

with $\omega_1 = 15\pi, \omega_2 = 30\pi, \omega_3 = 45\pi, \omega_4 = 60\pi$, $\phi_1 = 0, \phi_2 = 0.3\pi, \phi_3 = 0.7\pi, \phi_4 = \pi$, and $\alpha_1 = 1, \alpha_2 = 0.9, \alpha_3 = 0.85, \alpha_4 = 0.4$. The sampling interval T was selected as $1/202.5$ in order to give a trial length of $M = 108$ samples. The noise $z[n]$ was white Gaussian noise having a variance of 15 which produced an SNR of approximately -10 dB. In all, thirty 21,600 point runs were generated, each run corresponding to 200 trials. A second set of 30 runs at an SNR of -19 dB was also generated. The signal in each trial is depicted in Figure 4. Each single trial contains two stimulus presentations, in order to have a long enough trial length (M) to more readily uncorrelate noise and signal components as discussed above; moreover,

it should be observed that even though the number of trials has been halved (from 400 to 200), M has been doubled thereby doubling the SNR of the subspace average. We also note that for contrast gratings, typically used for measuring steady-state VEP's, a single stimulus presentation corresponds to two contrast (i.e. polarity) reversals.

Both the conventional and subspace averages were computed for each 200 trial run. For each run, the signal and noise powers were computed as a function of trial number and subsequently averaged over the 30 independent runs for each SNR. The signal and noise power estimates vs. trial number was then used to form an estimate of the subspace (conventional) SNR versus trial number, which is shown in Figure 5, along with the theoretical SNR (see (2), (14)). Figure (6) shows the function $\rho(k)$ as a function of trial number for the subspace and conventional average for a single run, along with the theoretical value of $\rho(k)$ given in (19). The experimental $\rho(k)$ tracks the theoretical value quite closely. It can also be seen that the $\rho(k)$ for the -19 dB run does not converge for the conventional average after 200 trials suggesting a low single-trial SNR. Finally, Figure 7 shows the subspace and conventional averages obtained after a single 200-trial run. The subspace average is seen to give a much closer representation of the actual signal in Fig 4 than the conventional average.

VI. Steady-State VEP Experiments

Counterphase modulated steady-state VEP's were acquired by presenting a black and white vertical square wave gratings with contrast set to 92%. The stimulus was created on a video monitor using a high resolution graphics board (Omnicomp, Texan ET, 1280 by 1024 pixels, 60 Hz, non-interleaved). The subjects viewed the video monitor binocularly from a distance of 3 m in a darkened room. The video screen was masked to reveal a 5.5 degree circular field. A small fixation dot was placed at the center of the display. The luminance of the display was approximately 30 Foot Lamberts. Counter-phase contrast reversal (7.5 contrast reversals per sec, square wave modulation) was used as the visual stimulus. Four 22,000 sample runs where obtained at stimulus spatial frequencies of 4, 10, 24,and 40 c/d, the first

three (4, 10, 24 c/d) where chosen to produce VEP's with relative high, moderate, and low SNR's, respectively. The 40 c/d stimulus appears to the subject as a homogeneous field which produces no VEP (spatial frequency threshold is about 30 c/d) so the 40 c/d data was used as a noise-only reference. In the high SNR data, four harmonics where clearly observed in the frequency domain, so d was set equal to 4 and the harmonics used for the subspace average were the same as those in the simulations.

For each fixed spatial frequency run, the VEP data was prewhitened with an order $L = 20$ linear prediction error filter determined from the Autocorrelation Method [37]. The prewhitened data was then used to compute both the conventional and the subspace average for a total of 200 non-overlapping trials. The performance measure, $\rho(k)$ was also computed for both subspace and conventional averages. Figure 8 depicts estimated power spectral densities for the VEP data at each of the four spatial frequencies. The PSD was estimated via Welch's averaged periodogram method using the Matlab function "psd". It can be seen that the signal component consists of approximately four harmonics at 7.5, 15, 22.5, and 30 Hz, as predicted by the VEP model (3), furthermore, the harmonic amplitudes decrease with increasing stimulus spatial frequency thereby reducing the VEP SNR. Figure 9 shows $\rho(k)$ computed from conventional and subspace averages at the four spatial frequencies. It can be seen that each $\rho(k)$ is fairly stable, though there is greater variation than in the simulations suggesting that the SNR of the VEP data changes from trial to trial. Nevertheless, the subspace average shows a significantly lower value of $\rho(k)$ for the three spatial frequencies containing signal (4, 10, and 24 c/d), while there is virtually no difference in $\rho(k)$ for the noise-only spatial frequency (40 c/d). The difference in $\rho(k)$ between the subspace and conventional averages for the signal-bearing data seems to be by about a factor of 10 which is close to the theoretically predicted SNR gain of $M/d = 108/8 = 13.5$ provided by the subspace average. Figure 10 shows a comparison between the ensemble averages computed via the subspace average and the conventional average. Even though the true signal morphology is unknown, it seems evident that the subspace averages have the higher SNR.

VII. Summary

A new algorithm for averaging steady-state VEP's was described. This algorithm increases the SNR of the VEP by projecting the VEP data into a signal subspace. The resulting subspace average was seen to produce higher SNR VEP's relative to the conventional average by a factor of M/d where M is the single trial length and d is the number of harmonics used to model the VEP. A new measure of VEP SNR which can be computed while the average is being acquired was also described. The measure, $\rho(k)$, was shown to converge to the noise to signal power ratio of the single trial VEP.

References

[1] J. R. Heckenlively and G. B. Arden, *Principles and Practice of Clinical Electrophysiology of Vision*. Mosby, 1991.

[2] G. A. Fishman and S. Sokol, *Electrophysiological Tesing in Disorders of the Retina, Optic Nerve, and Visual Pathway*. American Academy of Ophthalmolgy, 1990.

[3] D. Regan, *Human Brain Electrophysiology*. Elsevier, 1989.

[4] A. S. Gevins, "Analysis of electromagnetic signals of the human brain: milestones, obstacles, and goals," *IEEE Trans. Biomed. Eng.*, vol. BME-31, pp. 833–850, Dec. 1984.

[5] J. I. Aunon, C. D. McGillem, and D. G. Childers, "Signal processing in evoked potential research: Averaging and modeling," *CRC Critical Reviews in Bioengineering*, vol. 5, pp. 323–367, July 1981.

[6] C. D. McGillem, J. I. Aunon, and K. B. Yu, "Signal and noise in evoked potentials," *IEEE Transactions on Biomedical Engineering*, vol. 32, pp. 1012–1016, Dec. 1985.

[7] C. D. Woody, "Characterization of an adaptive filter for the characterization of variable latency neuroelectric signals," *Medical and Biological Engineering*, vol. 5, pp. 539–553, 1967.

[8] C. D. McGillem, J. I. Aunon, and C. A. Pomalaza, "Improved waveform estimation procedures for event-related potentials," *IEEE Transactions on Biomedical Engineering*, vol. BME-32, pp. 371–379, 1985.

[9] X. Yu, Y. Zhang, and Z. He, "Peak component latency-corrected average method for evoked potential waveform estimation," *IEEE Transactions on Biomedical Engineering*, vol. 41, pp. 1072–1082, Nov. 1994.

[10] L. Gupta, D. L. Molfese, R. Tammana, and P. G. Simos, "nonlinear alignment and averaging for estimating the evoked potential," *IEEE Transactions on Biomedical Engineering*, vol. 43, pp. 348–356, April 1996.

[11] D. O. Walter, "*A Posteriori* Wiener filtering of average evoked responses," *Electroencephalography and Clinical Neurophysiology*, pp. 61–70, suppl. 27 1969.

[12] T. Nogawa, K. Katayama, Y. Yabata, T. Kawahara, and T. Ihsijo, "Visual evoked potentials estimated by Wiener filtering," *Electroencephalography and Clinical Neurophysiology*, vol. 35, pp. 375–378, Oct. 1973.

[13] K. B. Yu and C. D. McGillem, "Optimum filters for estimating evoked potential waveforms," *IEEE Transactions on Biomedical Engineering*, vol. BME-30, pp. 730–737, Nov. 1983.

[14] M. Furst and A. Blau, "Optimal *a posteriori* time domain filter for average evoked potentials," *IEEE Transactions on Biomedical Engineering*, vol. 38, pp. 827–833, Sept. 1991.

[15] J. P. C. de Weerd, "*A posteriori* time-varying filtering of averaged evoked potentials. i. introduction and conceptual basis," *Biological Cybernetics*, vol. 41, pp. 211–222, 1981.

[16] J. P. C. de Weerd, "*A posteriori* time-varying filtering of averaged evoked potentials. ii. mathematical and computational basis," *Biological Cybernetics*, vol. 41, pp. 223–234, 1981.

[17] M. V. Spreckelsen and B. Bromm, "Estimation of single evoked cerebral potentials by means of parametric modeling and Kalman filtering," *IEEE Transactions on Biomedical Engineering*, vol. BME-35, pp. 691–700, Sept. 1988.

[18] S. Cerutti, G. Chiarenza, D. Liberati, P. Mascellani, and G. Pavesi, "A parametric method of identification of single-trial event-related potentials in the brain," *IEEE Transactions on Biomedical Engineering*, vol. BME-35, pp. 701–711, 1988.

[19] H. J. Heinze, H. Kunkel, and W. Massing, "Selective filtering of single evoked potentials by high performance ARMA methods," in *Recent Advances in EEG and EMG Data Processing* (N. Y. et al, ed.), pp. 29–46, Amsterdam: North Holland, 1981.

[20] H. Al-Nashi, "Maximum likelihood method for estimating eeg evoked potentials," *IEEE Transactions on Biomedical Engineering*, vol. BME-33, pp. 1087–1095, Dec. 1986.

[21] D. H. Lange and G. F. Inbar, "A robust parametric estimator for single-trial movement related brain potentials," *IEEE Transactions on Biomedical Engineering*, vol. 43, pp. 341–347, April 1996.

[22] M. Hansson and G. Salomonsson, "Estimation of single event-related potentials utilizing the Prony method," *IEEE Transactions on Biomedical Engineering*, vol. 43, pp. 973–981, Oct. 1996.

[23] D. H. Lange, H. Pratt, and G. F. Inbar, "Modeling and estimation of single evoked brain potential components," *IEEE Transactions on Biomedical Engineering*, vol. 44, pp. 791–799, Sept 1997.

[24] N. V. Thakor, "Adaptive filtering of evoked potentials," *IEEE Transactions on Biomedical Engineering*, vol. BME-34, pp. 6–12, 1987.

[25] C. A. Vaz and N. V. Thakor, "Adaptive fourier estimation of time-varying evoked potentials," *IEEE Transactions on Biomedical Engineering*, vol. BME-36, pp. 448–455, 1989.

[26] P. Laguna, R. Jane, O. Meste, P. Poon, P. Caminal, H. Rix, and N. Thakor, "Adaptive filter for event-related bioelectric signals using an impulse correlated reference input," *IEEE Transactions on Biomedical Engineering*, vol. BME-39, pp. 1032–1044, October 1992.

[27] O. Svensson, "Tracking of changes in latency and amplitude of the evoked potential by

using adaptive LMS filters and exponential averagers," *IEEE Transactions on Biomedical Engineering*, vol. 40, pp. 1074–1079, Oct. 1993.

[28] N. V. Thakor, X. Guo, C. Vaz, P. Laguna, R. Hane, P. Caminal, H. Rix, and D. F. Hanley, "Orthonormal (Fourier and Walsh) models of time varying evoked potentials in neurological injury," *IEEE Transactions on Biomedical Engineering*, vol. 40, pp. 213–221, March 1993.

[29] X. Yu and Z. He, "Time-varying adaptive filters for evoked potential estimation," *IEEE Transactions on Biomedical Engineering*, vol. 41, pp. 1062–1071, Nov. 1994.

[30] A. K. Barros and N. Ohnishi, "Mse behavior of biomedical event-related filters," *IEEE Transactions on Biomedical Engineering*, vol. 44, pp. 848–855, Sept. 1997.

[31] A. Ademoglu, E. Micheli-Tzanakou, and Y. Istefanopulos, "Analysis of pattern reversal visual evoked potentials (PRVEP's) by spline wavelets," *IEEE Transactions on Biomedical Engineering*, vol. 44, pp. 881–890, Sept. 1997.

[32] N. V. Thakor, G. Xin-rong, S. Yi-Chun, and D. F. Hanley, "Multiresolution wavelet analysis of evoked potentials," *IEEE Transactions on Biomedical Engineering*, vol. 40, pp. 1085–1094, Nov. 1993.

[33] T. F. Collura, "Real-time filtering for the estimation of steady-state visual evoked potentials," *IEEE Transactions on Biomedical Engineering*, vol. BME-37, pp. 650–652, June 1990.

[34] C. E. Davila, A. Abaye, and A. Khotanzad, "Estimation of single-sweep steady state visual evoked potentials by adaptive line enhancement," *IEEE Transactions on Biomedical Engineering*, vol. BME-41, pp. 197–200, Feb. 1994.

[35] C. E. Davila, R. Srebro, and I. Ghaleb, "Optimal detection of visual evoked potentials," *IEEE Transactions on Biomedical Engineering*, vol. 45, pp. 80–83, June 1998.

[36] J. A. McEwen and G. B. Anderson, "Modeling the stationarity and gaussianity of spontaneous electroencephalographic activity," *IEEE Trans. Biomed. Eng.*, vol. BME-22, pp. 361–369, Sept. 1975.

[37] S. M. Kay, *Modern Spectral Estimation*. Prentice Hall, 1988.

[38] D. W. Tufts and R. Kumaresan, "Estimation of frequencies of multiple sinusoids: Making linear prediction perform like maximum likelihood," *Proc. IEEE*, vol. 70, pp. 975–989, Sept. 1982.

[39] D. W. Tufts and R. Kumaresan, "Data adaptive signal estimation by singular value decomposition of the data matrix," *Proceedings if the IEEE*, vol. 70, pp. 684–685, 1982.

[40] G. Strang, *Linear Algebra and Its Applications*. Academic Press, 1980.

[41] H. L. V. Trees, *Detection, Estimation, and Modulation Theory*. Wiley, 1968.

[42] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Johns Hopkins University Press, 1989.

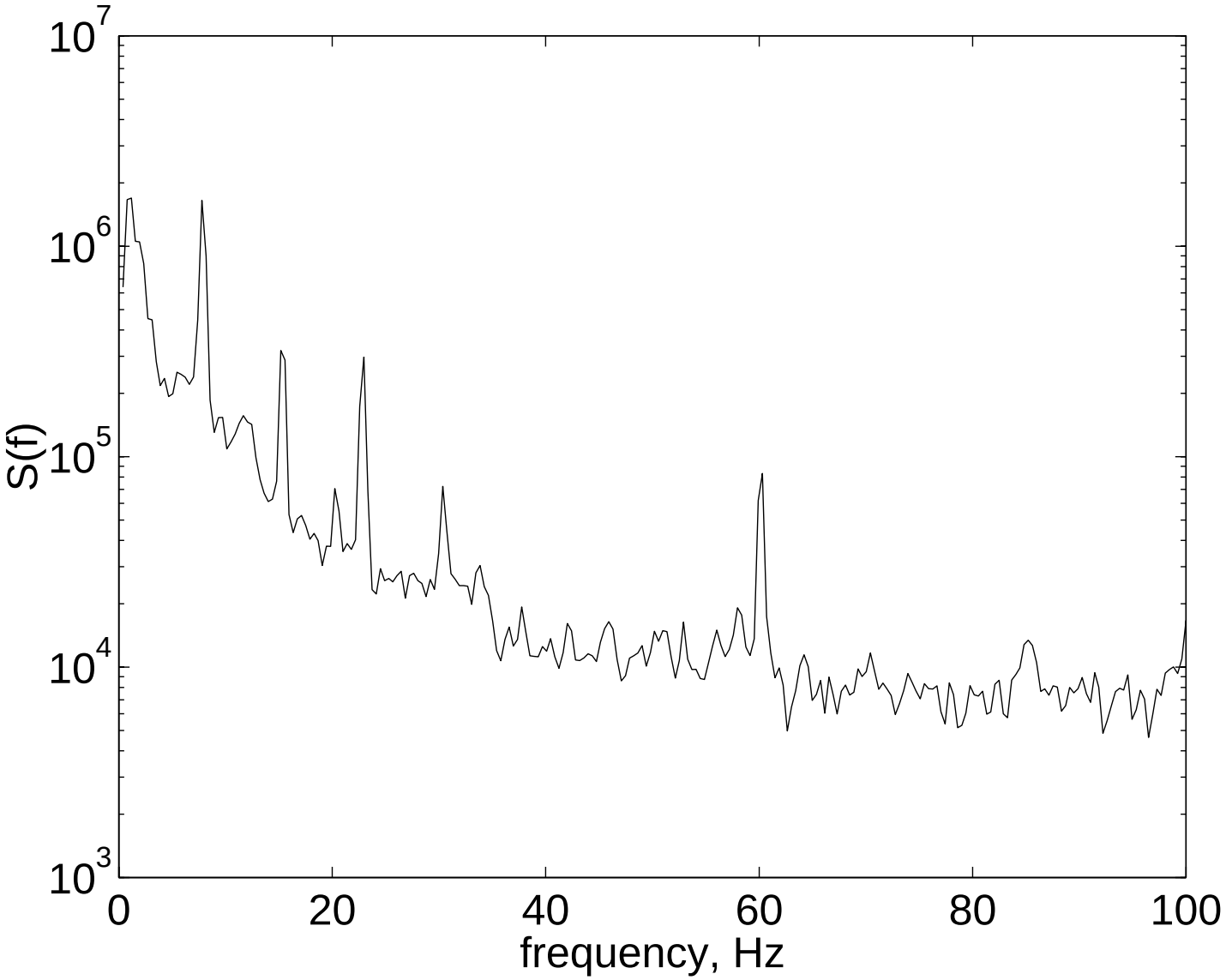

Figure 1: Power spectral density estimate of a steady-state VEP showing the contrast reversal frequency (7.5 Hz) and several harmonics (15 Hz, 22.5 Hz, and 30 Hz).

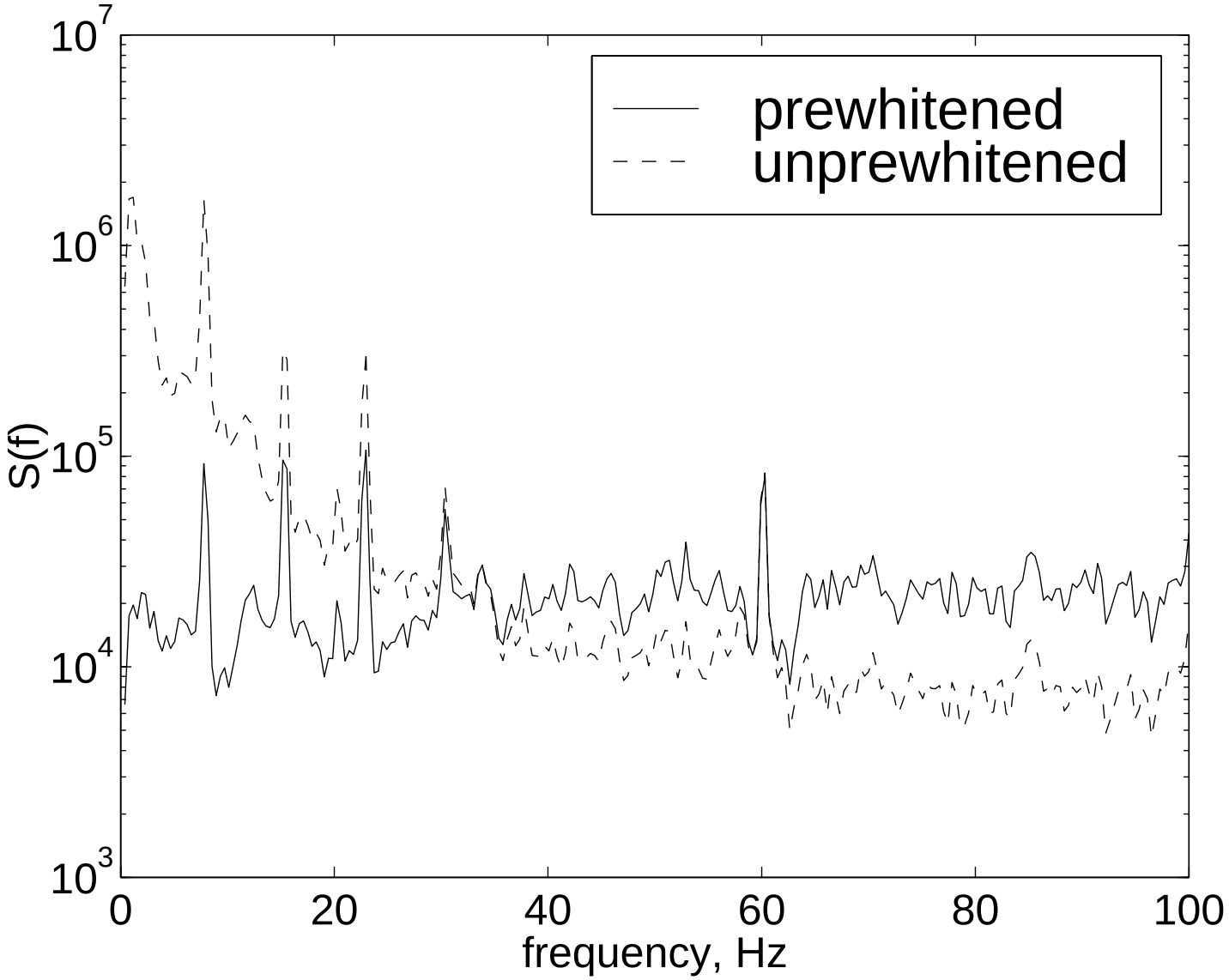

Figure 2: Power spectral density estimate of a steady-state before and after prewhitening, the data SNR remains relatively unchanged by prewhitening.

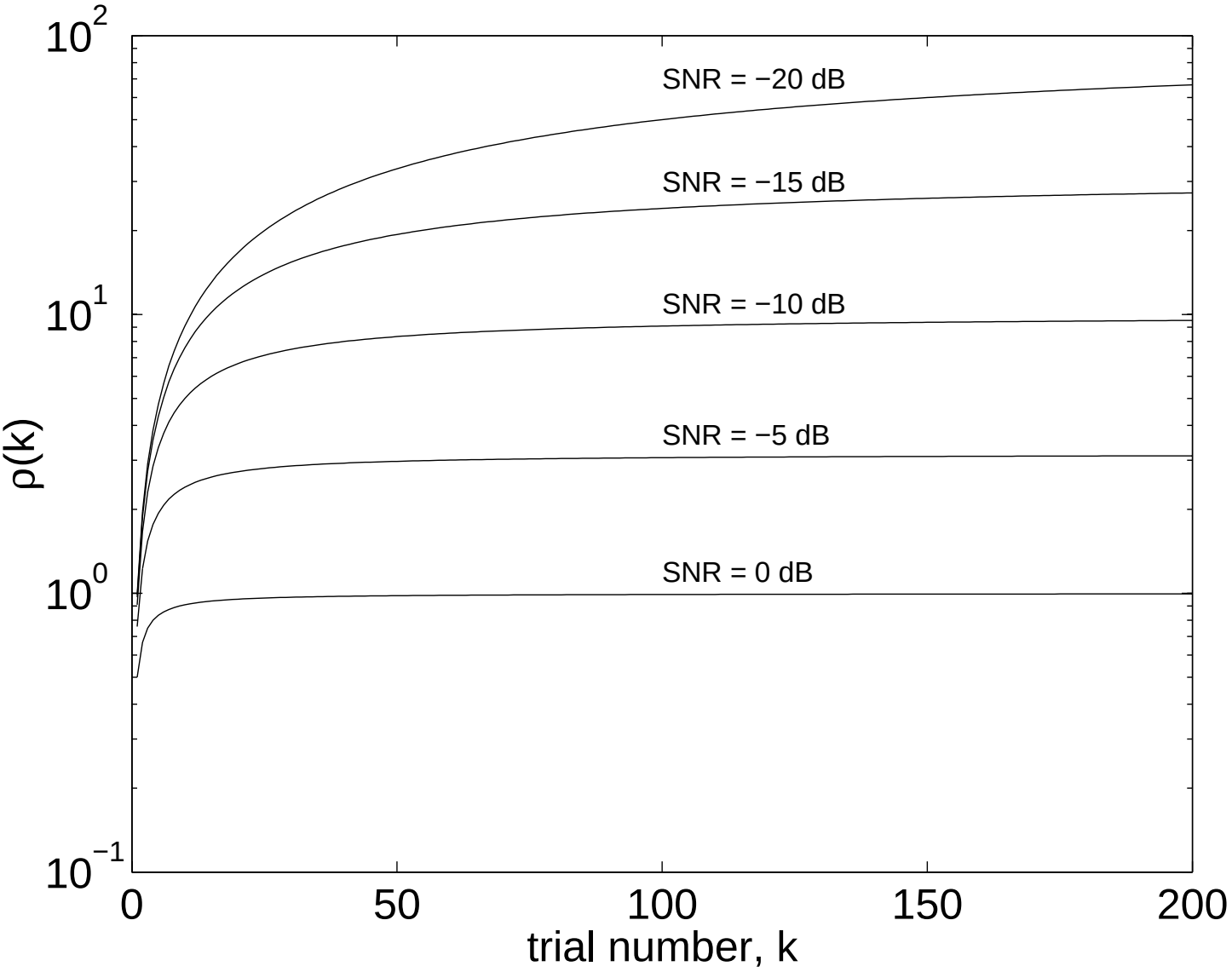

Figure 3: $\rho(k)$ converges to the single-trail noise-to-signal ratio, σ_z^2/σ_s^2.

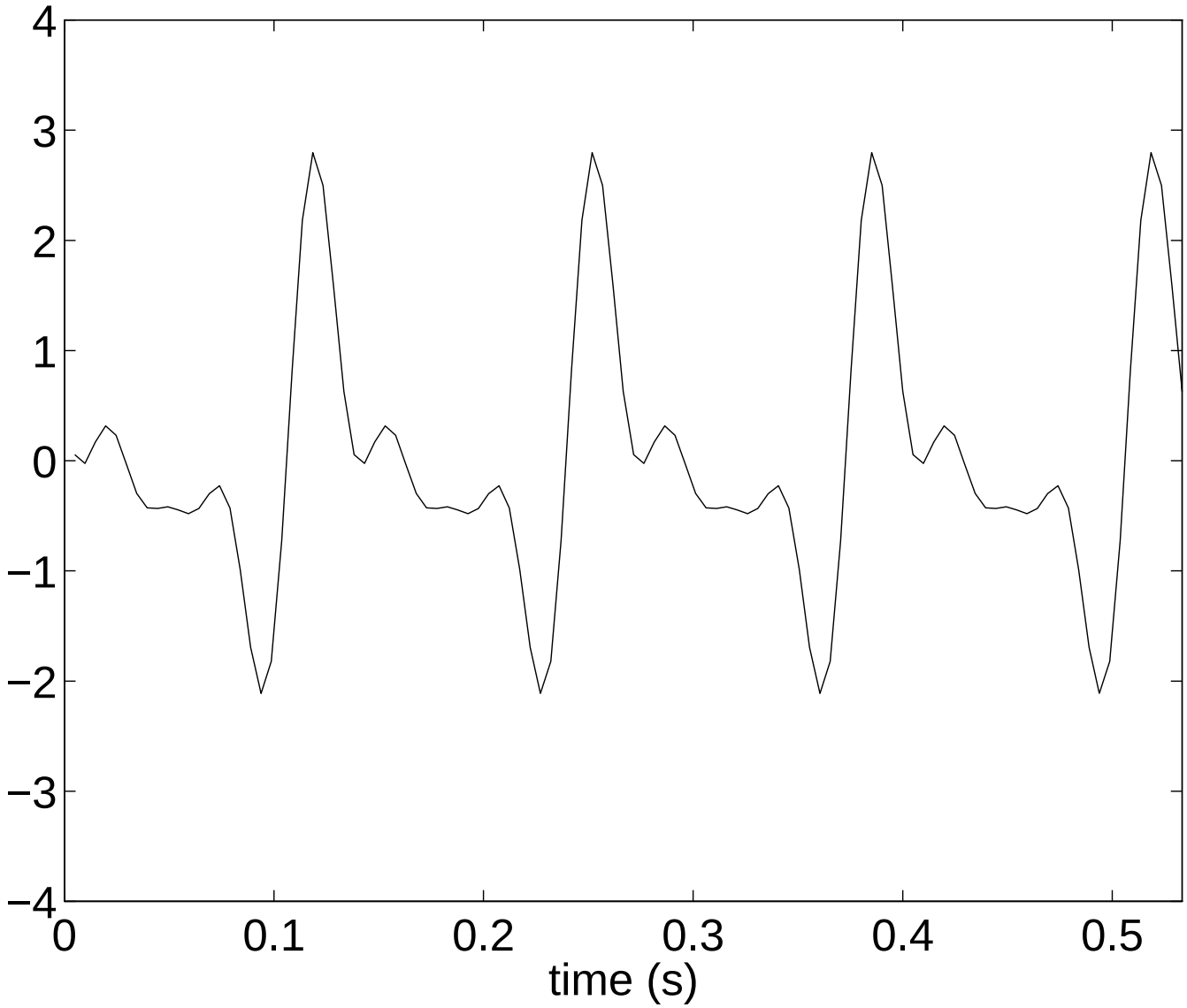

Figure 4: Single trial signal component used in simulations.

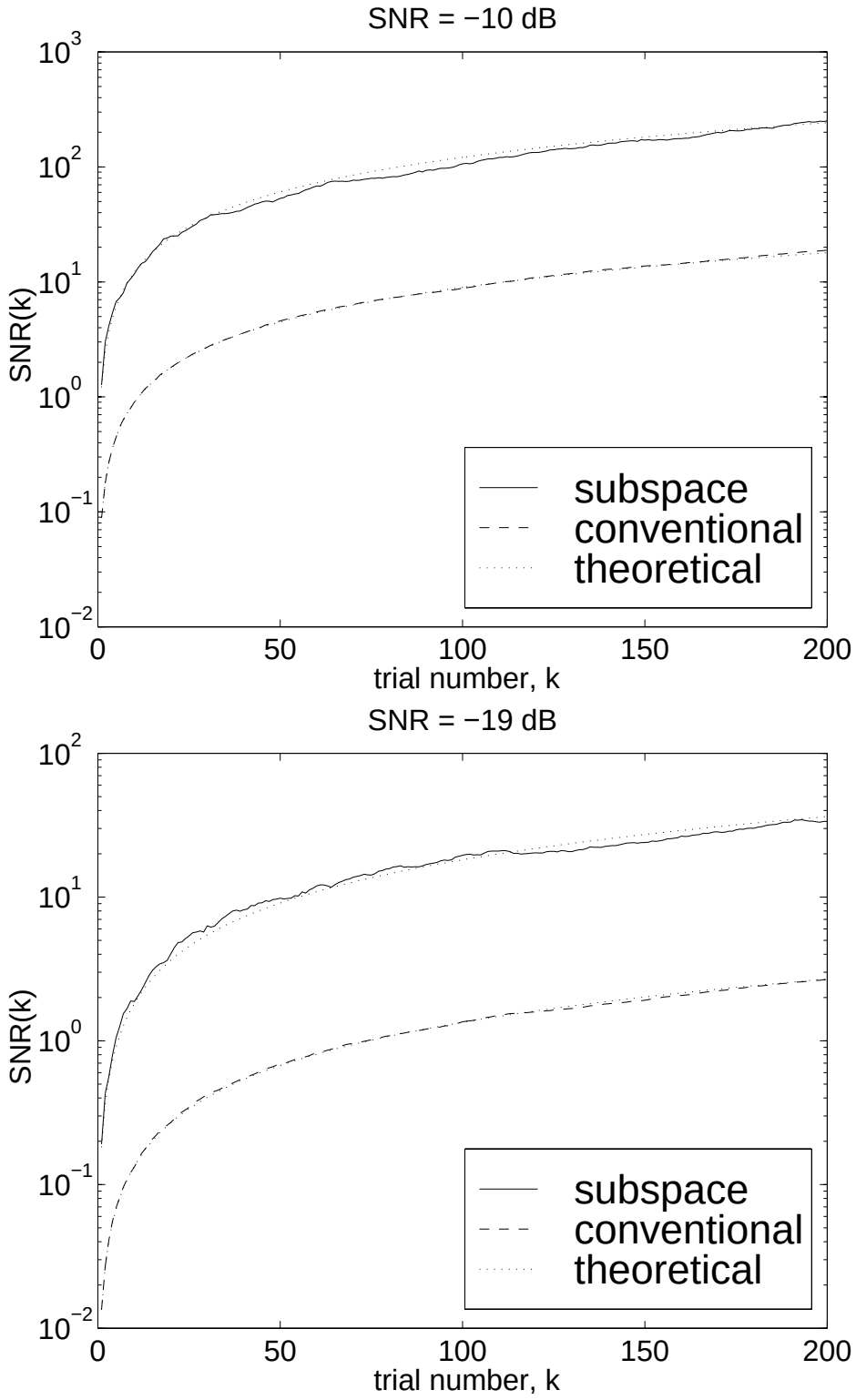

Figure 5: Comparison of estimated SNR and theoretically predicted SNR versus trial number for the two simulated data experiments (SNR of -10 dB and -19 dB).

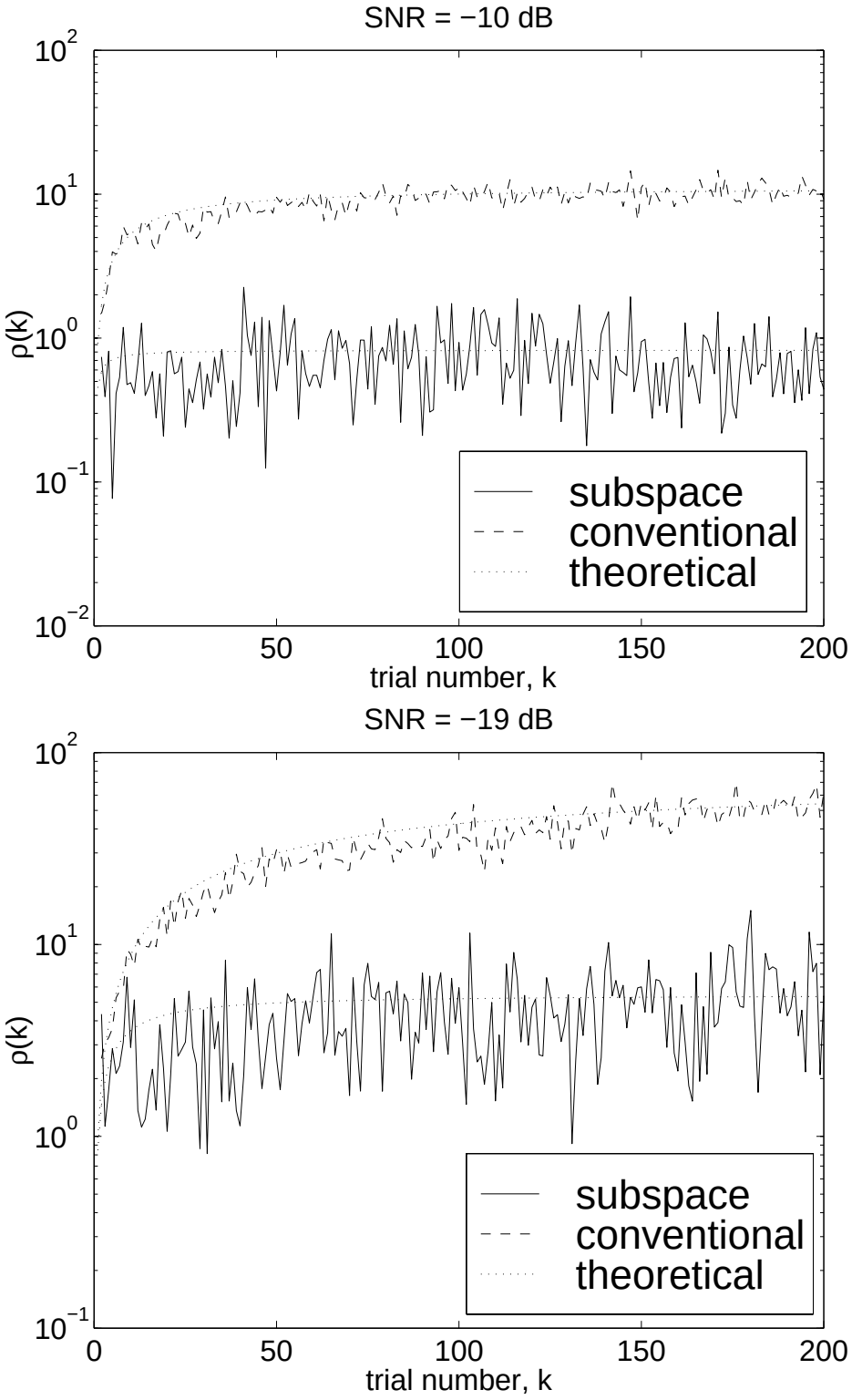

Figure 6: The function $\rho(k)$ for the two simulated data experiments (SNR of -10 dB and -19 dB) derived from a single run, along with the theoretically predicted value of $\rho(k)$

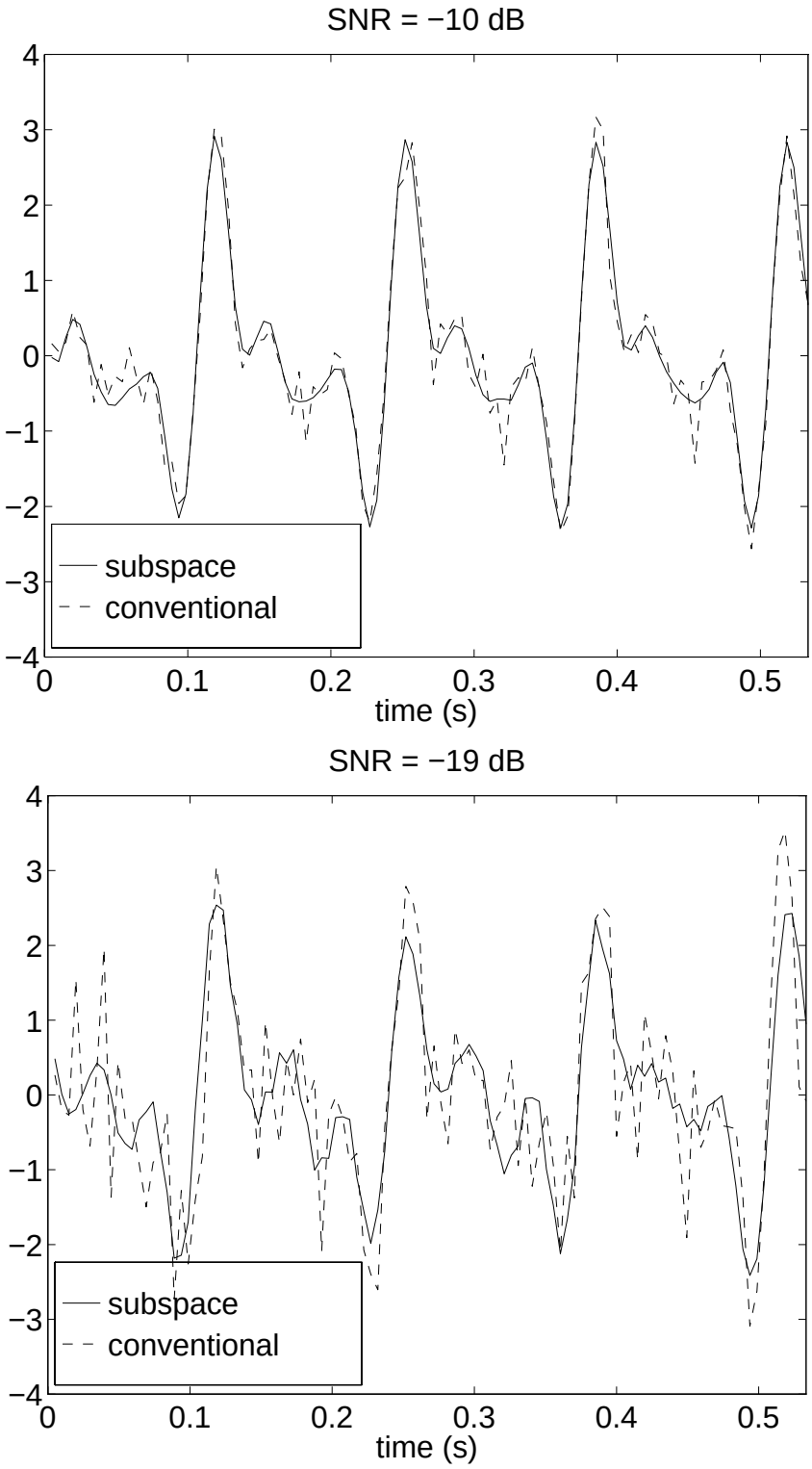

Figure 7: Conventional and subspace averages after a 200-trial run for synthetic single-trial SNR's of -10 dB and -19 dB.

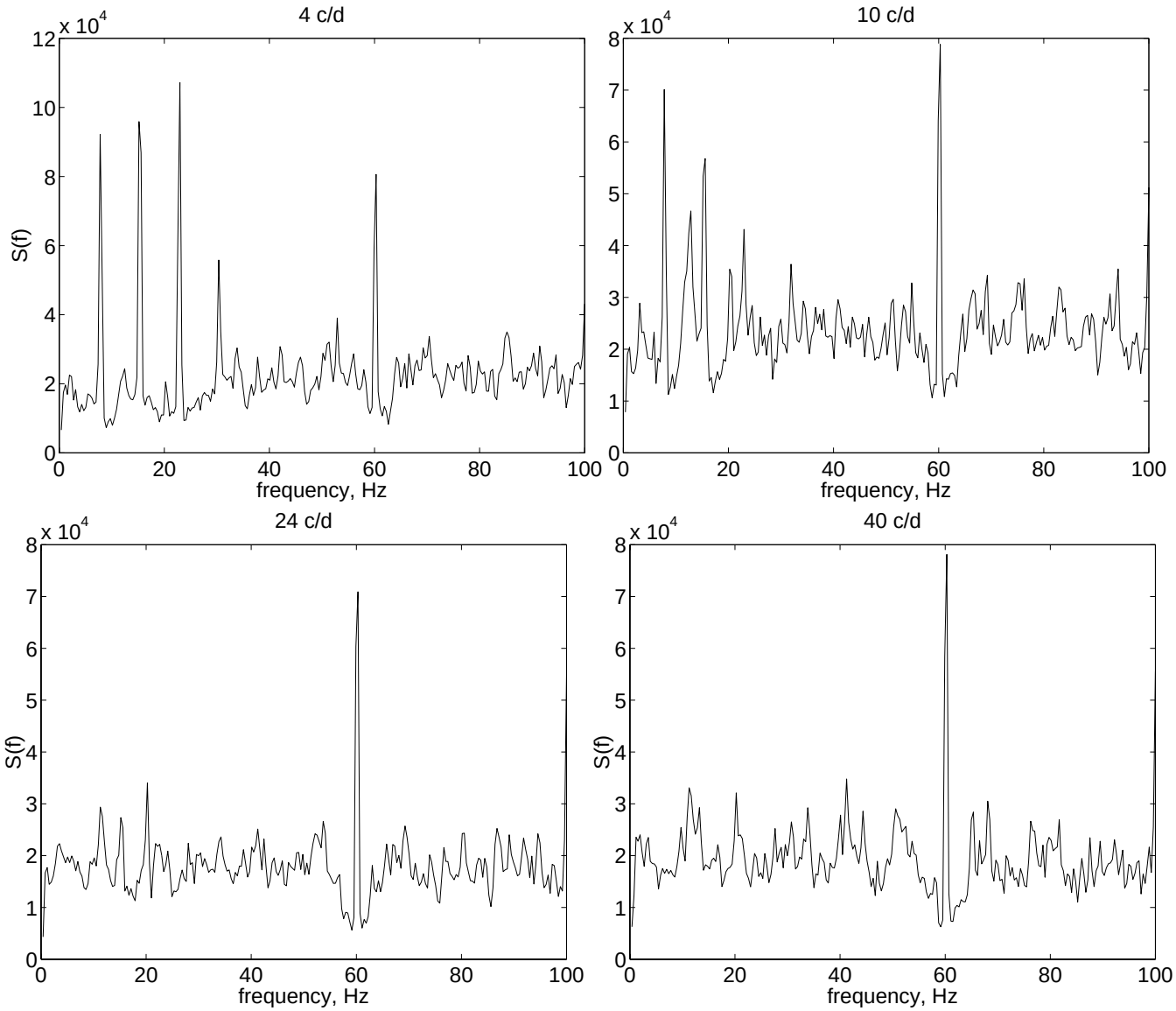

Figure 8: Estimated power spectral density obtained from actual VEP data for stimulus spatial frequencies of 4 c/d, 10 c/d, 24 c/d, and 40 c/d (noise only).

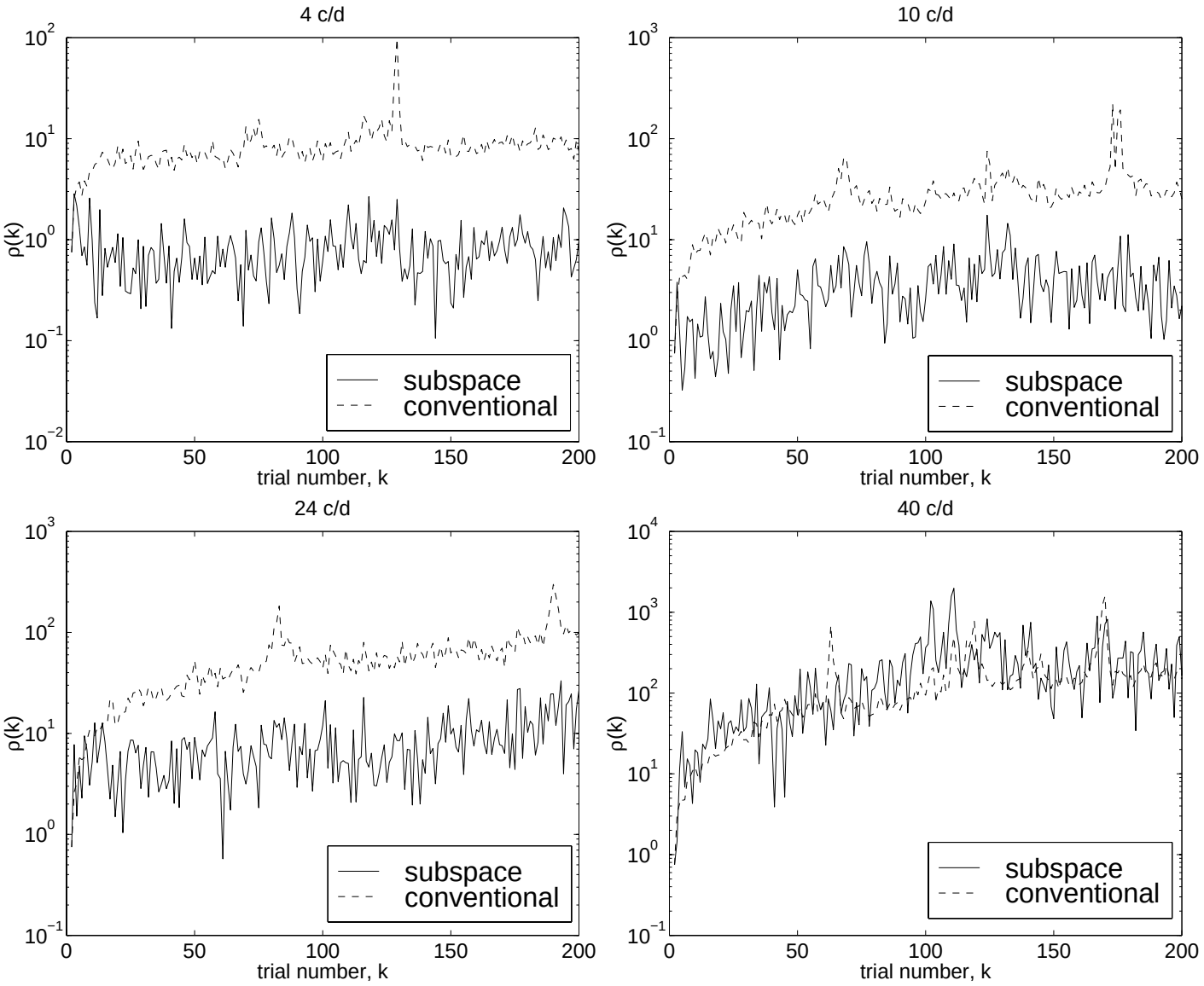

Figure 9: $\rho(k)$ computed from subspace and conventional averages for actual VEP data for stimulus spatial frequencies of 4 c/d, 10 c/d, 24 c/d, and 40 c/d (noise only), $\rho(k)$ is a measure of noise to signal ratio.

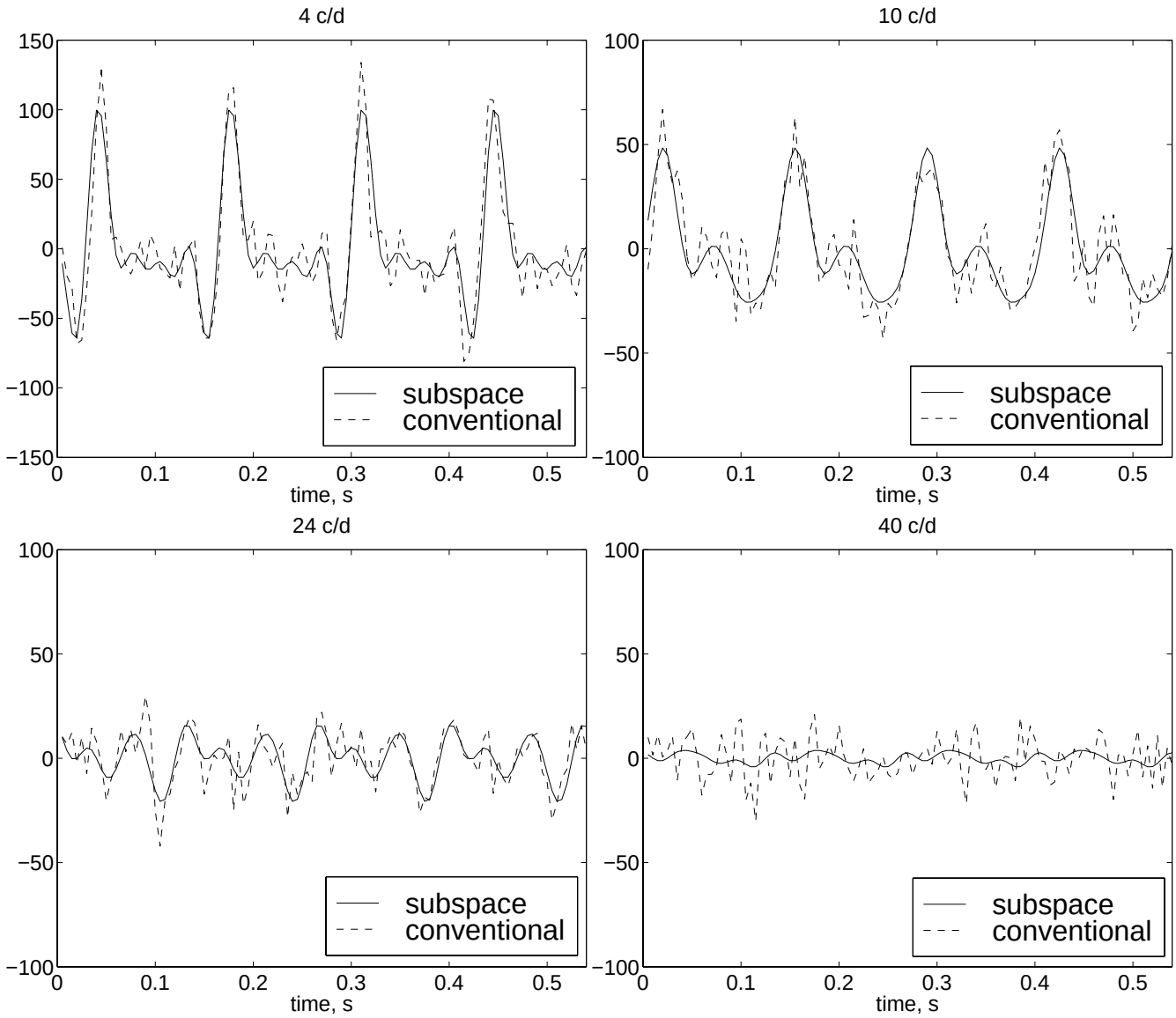

Figure 10: Subspace and conventional averages computed from actual VEP data for stimulus spatial frequencies of 4 c/d, 10 c/d, 24 c/d, and 40 c/d (noise only).