

Abstract

GENERALIZED EVENT-RELATED POTENTIALS: SYSTEM MODELS FOR CONTINUOUS EEG WITH TASK VARIABLES

¹Mark E. Pflieger, ¹Richard E. Greenblatt, ²Jules P.A. Dewald

¹Source Signal Imaging, San Diego, CA USA, <http://www.sourcesignal.com>

²Rehabilitation Institute of Chicago, Chicago, IL USA

If we adopt the simplified view that the brain is a deterministic system having EEG as output, with isolated task events as input, then we can use average event-related potentials (ERPs) to approximate impulse response functions. In an actual experiment, however, task events are not isolated. Rather, they interact via brain memory systems, and their associated electrophysiological responses often overlap in time. Consequently, an average ERP may obscure the response dynamics. Moreover, an average does not capture physiological interactions between events that may be of interest in a cognitive experiment. As an alternative to averaging, we demonstrate the feasibility of using linear/nonlinear system identification methods to characterize quasi-deterministic event-related dynamics. The method requires continuous EEG data with task variables, and it produces estimates of both linear responses and nonlinear interactions, which characterize a Volterra system model. Each task variable has an associated linear impulse response waveform, that is, a temporally deconvolved ERP. Matrix-like kernels represent nonlinear interactions for each pair of task variables. In the context of the general linear model, Friston et al. (1998, *Magn Reson Med* 39:41-52; 2000, *NeuroImage* 12:466-77) have applied a similar approach to fMRI time series. Thus, dynamic system parameter estimation provides a common framework for processing both ERP and event-related fMRI experiments. We illustrate the method using data from a visual event-related potential experiment.

Presented at Cognitive Neuroscience Society, New York City, 25 March 2001



Overview

The premise of this presentation is that we can generalize the useful concept of an event-related potential or ERP (Vaughan) by explicitly modeling the linear and nonlinear dynamics of the ERP-producing subsystems of the brain. An ERP implicitly models the linear part of the dynamics. The notion of a “task-related Volterra series” is one way to extend the linear theory to encompass simple nonlinearities—that is, bilinear (or higher-order multilinear) interactions within or between task variables, as reflected in the EEG.

We start with the theory of task-related Volterra series, which we characterize by so-called “kernels”. An ERP approximates the first order kernel, or impulse response function, which is the linear part of the task-related dynamics. Thus, second order kernels are natural extensions of the ERP concept. After discussing how to estimate the kernels of a task-related Volterra series, we illustrate the method using data from a simple event-related visual paradigm.



Task-related Volterra Series

A Volterra series (Schetzen, 1980; Bendat, 1998) of L task variables $x_i(t)$ can model the task-related responses and interactions of a measured time series $y(t)$, such as the BOLD fMRI signal at a voxel (Friston *et al.*, 1998) or a M/EEG signal, as follows:

$$y(t) = \varepsilon(t) + \sum_{d=0}^{P_\alpha} \alpha_d t^d + \sum_{i=1}^L \int_{-\infty}^{\infty} h_i^{(1)}(\tau) x_i(t - \tau) d\tau + \sum_{i_1=1}^L \sum_{i_2=i_1}^L \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h_{i_1 i_2}^{(2)}(\tau_1, \tau_2) x_{i_1}(t - \tau_1) x_{i_2}(t - \tau_2) d\tau_1 d\tau_2 + \dots \quad [1]$$

Here $\varepsilon(t)$ arises from a random background process, which is primarily task-unrelated brain activity. We have included low-order powers of time with coefficients α_d in order to account for P_α slow drifts plus a constant, which we regard as uninteresting. The experimentally driven responses and interactions of interest are modeled by the Volterra kernels $h_i^{(1)}$ (order 1) and $h_{i_1 i_2}^{(2)}$ (order 2). Task variables linearly convolve with their first order kernels, and product pairs of task variables bilinearly convolve with their second order kernels. Higher order Volterra kernels n -linearly convolve with product n -tuples of task variables; however, the number of unknown parameters that we must estimate increases sharply with the order.



Linear Responses and Bilinear Interactions

A **first order kernel** $h_i^{(1)}(\tau)$ represents the linear response associated with task variable x_i at time lag τ . We assume that responses are *transient*, i.e., the kernel is identically zero except for a finite interval of time lags $[T_{\min}^{(i)}, T_{\max}^{(i)}]$, which we call the *effective range* for the i th task variable. If the task variable is associated with a stimulus, then we assume that the brain response *follows* the stimulus, which implies that a sensory kernel is identically zero for all non-positive lags. If the task variable is associated with a behavior, then we assume that the brain “response” *precedes* the behavior, which implies that a motor kernel is identically zero for all non-negative lags. For tasks that involve tight coupling between a stimulus and a behavior, we can consider sensorimotor kernels that are nonzero for both positive and negative lags. Whenever possible, however, it is preferable to preserve causality by teasing apart the sensory and motor responses.

A **second order kernel** $h_{i_1 i_2}^{(2)}(\tau_1, \tau_2)$ represents the bilinear interaction between task variables x_{i_1} and x_{i_2} at respective time lags τ_1 and τ_2 . Kernels with $i_1 = i_2$ represent quadratic self-interactions. We assume that a second order kernel can be nonzero only when *both* time lags are within the effective ranges of their associated task variables.



Approximation via Temporal Basis Functions

For each task variable x_i , we can specify a series of basis functions $b_1^{(i)}, \dots, b_{P_i}^{(i)}$ that enable us to approximate the first order kernel as a weighted sum (Schetzen, 1980; Friston *et al.*, 1998):

$$h_i^{(1)}(\tau) \approx \sum_{j=1}^{P_i} \beta_j^{(i)} b_j^{(i)}(\tau) \quad [2]$$

(We discuss below a method for optimizing these basis functions for particular task variables.) Using the same basis functions, we similarly expand the second order kernels as

$$h_{i_1 i_2}^{(2)}(\tau_1, \tau_2) \approx \sum_{j_1=1}^{P_{i_1}} \sum_{j_2=1}^{P_{i_2}} \gamma_{j_1 j_2}^{(i_1, i_2)} b_{j_1}^{(i_1)}(\tau_1) b_{j_2}^{(i_2)}(\tau_2) \quad [3]$$

Noting that the β and γ coefficients are not functions of time lags, we next substitute approximations [2] and [3] into equation [1] with an eye towards joint estimation of the α (uninteresting drift), β (interesting response), and γ (interesting interaction) parameters:



$$\begin{aligned}
y(t) \approx \mathcal{E}(t) + \sum_{d=0}^{P_\alpha} \alpha_d t^d + \sum_{i=1}^L \sum_{j=1}^{P_i} \beta_j^{(i)} \int_{T_{\min}^{(i)}}^{T_{\max}^{(i)}} b_j^{(i)}(\tau) x_i(t - \tau) d\tau + \\
\sum_{i_1=1}^L \sum_{i_2=i_1}^L \sum_{j_1=1}^{P_{i_1}} \sum_{j_2=1}^{P_{i_2}} \gamma_{j_1 j_2}^{(i_1, i_2)} \int_{T_{\min}^{(i_1)}}^{T_{\max}^{(i_1)}} b_{j_1}^{(i_1)}(\tau_1) x_{i_1}(t - \tau_1) d\tau_1 \int_{T_{\min}^{(i_2)}}^{T_{\max}^{(i_2)}} b_{j_2}^{(i_2)}(\tau_2) x_{i_2}(t - \tau_2) d\tau_2
\end{aligned} \tag{4}$$

By convolving task variables with basis functions, we derive new time series

$$z_j^{(i)}(t) \equiv \int_{T_{\min}^{(i)}}^{T_{\max}^{(i)}} b_j^{(i)}(\tau) x_i(t - \tau) d\tau \tag{5}$$

Finally, we simplify equation [4] by substituting definition [5]:

$$y(t) \approx \mathcal{E}(t) + \sum_{d=0}^{P_\alpha} \alpha_d t^d + \sum_{i=1}^L \sum_{j=1}^{P_i} \beta_j^{(i)} z_j^{(i)}(t) + \sum_{i_1=1}^L \sum_{i_2=i_1}^L \sum_{j_1=1}^{P_{i_1}} \sum_{j_2=1}^{P_{i_2}} \gamma_{j_1 j_2}^{(i_1, i_2)} z_{j_1}^{(i_1)}(t) z_{j_2}^{(i_2)}(t) \tag{6}$$

Because we obtain the z functions by convolving known task variables with predetermined basis functions, we thus can jointly estimate, via generalized least squares, the α (drift), β (linear response), and γ (nonlinear interaction) parameters. Equations [2] and [3] reconstruct the first and second order Volterra kernels from these parameters. We note that the self-interaction kernels are symmetric.

⇓

Gauss-Markov estimation of Volterra kernel parameters

The total number P of estimated parameters for the second order Volterra model is

$$P = 1 + P_\alpha + \sum_{i=1}^L P_i + \frac{1}{2} \sum_{i=1}^L P_i(P_i - 1) + \sum_{i_1=1}^{L-1} \sum_{i_2=i_1+1}^L P_{i_1} P_{i_2} \quad [7]$$

This total includes a constant, the α trend coefficients, the β response parameters, the symmetric self-interaction γ parameters, and the cross-interaction γ parameters, from which we form a vector \mathbf{w} of length P

$$\mathbf{w} \equiv [\alpha_0 \cdots \alpha_{P_\alpha}, \beta_1^{(1)} \cdots \beta_{P_L}^{(L)}, \gamma_{11}^{(1,1)} \cdots \gamma_{LL}^{(P_L, P_L)}]^\top \quad [8]$$

Corresponding with the column vector \mathbf{w} , we form a row vector $\mathbf{a}(t)$ of length P that reflects the experimental situation at time t (in terms of basis-convolved task variables):

$$\mathbf{a}(t) \equiv [t^0 \cdots t^{P_\alpha}, z_1^{(1)}(t) \cdots z_{P_L}^{(L)}(t), z_1^{(1)}(t)z_1^{(1)}(t) \cdots z_{P_L}^{(L)}(t)z_{P_L}^{(L)}(t)] \quad [9]$$

For an experimental run consisting of T time samples, the full $T \times P$ design matrix \mathbf{A} is

$$\mathbf{A} \equiv [\mathbf{a}(t_1) \cdots \mathbf{a}(t_T)]^\top \quad [10]$$



Collecting the measured signals for all time samples yields a data vector \mathbf{y} of length T :

$$\mathbf{y} \equiv [y(t_1) \cdots y(t_T)]^T \quad [11]$$

Likewise, we form an error vector $\boldsymbol{\varepsilon}$ of length T :

$$\boldsymbol{\varepsilon} \equiv [\boldsymbol{\varepsilon}(t_1) \cdots \boldsymbol{\varepsilon}(t_T)]^T \quad [12]$$

We may rewrite equation [6] in matrix form as

$$\mathbf{y} \approx \mathbf{A}\mathbf{w} + \boldsymbol{\varepsilon} \quad [13]$$

The time series $\boldsymbol{\varepsilon}$ has a serial correlation structure that derives from “background” neural activity, other physiological activity, motion effects, and/or scanner/system noise (Zarahn *et al.*, 1997). This, in turn, implies that the frequency spectrum is not flat (i.e., $\boldsymbol{\varepsilon}$ is not white noise). If the error process is approximately gaussian and stationary, we can solve [13] via ordinary least squares estimation if we first prewhiten the data. In the frequency domain, we divide by the amplitude at each frequency, which flattens (“whitens”) the spectrum. The time domain equivalent of prewhitening can be formally written as equation [14], where $\boldsymbol{\Sigma}$ is the $T \times T$ variance-covariance matrix for the error time series. Stationarity of the background activity implies that each autocovariance $\Sigma_{t_i t_j}$ depends only on the lag $\tau = |t_i - t_j|$. If we

extrapolate autocovariance $\tau = |t_i - t_j|$ s beyond a specified lag τ_{\max} using the principle of maximum entropy, then we can use autoregressive (AR) methods to design a prewhitening filter in the time domain by inverting an AR model of the error process, which involves only autocovariances with lags less than τ_{\max} (Friston, Josephs *et al.*, 2000; Appendix A). Applying the prewhitening filter $\Sigma^{-1/2}$ to each term of equation [13], we obtain

$$\bar{\mathbf{y}} \equiv \Sigma^{-1/2} \mathbf{y}, \quad \bar{\mathbf{A}} \equiv \Sigma^{-1/2} \mathbf{A}, \quad \bar{\boldsymbol{\varepsilon}} \equiv \Sigma^{-1/2} \boldsymbol{\varepsilon} \quad [14]$$

Under the assumptions, $\bar{\boldsymbol{\varepsilon}}$ is identically and independently distributed, in which case the Gauss-Markov estimate for the parameters is

$$\hat{\mathbf{w}} = \bar{\mathbf{A}}^+ \bar{\mathbf{y}} \quad [15]$$

Here the “+” superscript represents the pseudoinverse operation.

In practice, we typically must use the same data to estimate the autocovariances of the background brain activity with lags less than τ_{\max} . A “feasible” approach (e.g., Dale *et al.*, 2000) is to obtain a series of estimates, $\hat{\mathbf{w}}_0, \hat{\mathbf{w}}_1, \dots$, where $\hat{\mathbf{w}}_0$ is computed using $\Sigma \approx \mathbf{I}$ (or another initial estimate); $\hat{\mathbf{w}}_1$ is computed using the approximation $\Sigma \approx \mathbf{S}_1$, where we form \mathbf{S}_1 by calculating lag covariances from the residual vector $\mathbf{y} - \mathbf{A}\hat{\mathbf{w}}_0$; and so on. This series converges to mutually consistent estimates of \mathbf{w} and Σ , given

⇓

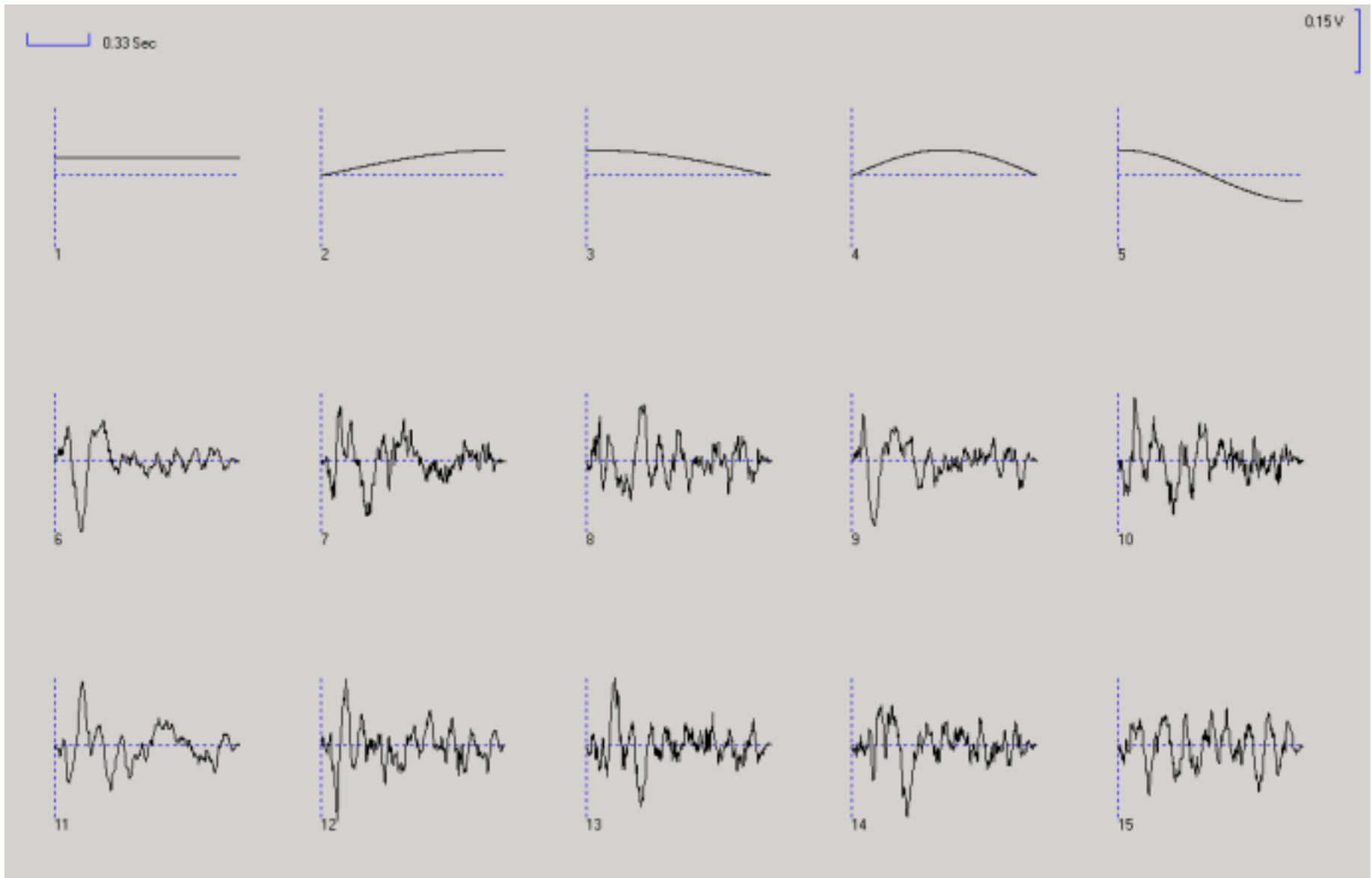
the data and the modeling assumptions. Friston, Josephs et al. (2000) have recently discussed caveats regarding the effect of prewhitening filters on statistical bias.

Although we have just indicated the theory of using temporal covariances in the “residual” (unmodeled) brain activity, the results shown here are simple least squares estimates.

Where do the basis functions come from?

One approach is to select basis functions on an a priori basis, e.g., Laguerre polynomials (Schetzen, 1980) or Gamma distributions (Friston, 1998). Here we have adopted an empirical approach that so far seems to perform well in practice. The idea is to start with a “calibration” set of event-related data, where the events are relatively isolated in time (unlike the actual task of interest). We form event-related epochs in the usual fashion from multi-channel EEG data. Then we form the sum of cross-products and cross-product of sums matrices as described in Pflieger and Nakada (2000, p. 158)—although we do this here in a *temporal* rather than a spatial fashion. From these, we form a temporal SNR matrix, which is, essentially, a ratio of the “foreground” to “background” event-related activity. Principal component analysis of the SNR matrix yields a ranking of components that are “tailored” to the event that will be used in the experiment. Finally, we obtain the basis functions via the inverse SNR transform of the principal components.





Subsequent analysis used the basis functions shown above. These consist of 5 a priori trends “not of interest” (which we estimate and discard) and 10 waveforms of interest extracted from calibration data. Calibration data, in this case, was a 100 ms visual checkerboard onset that repeated ~5 s. In the actual experiment, the stimulus had shorter, random ISIs.





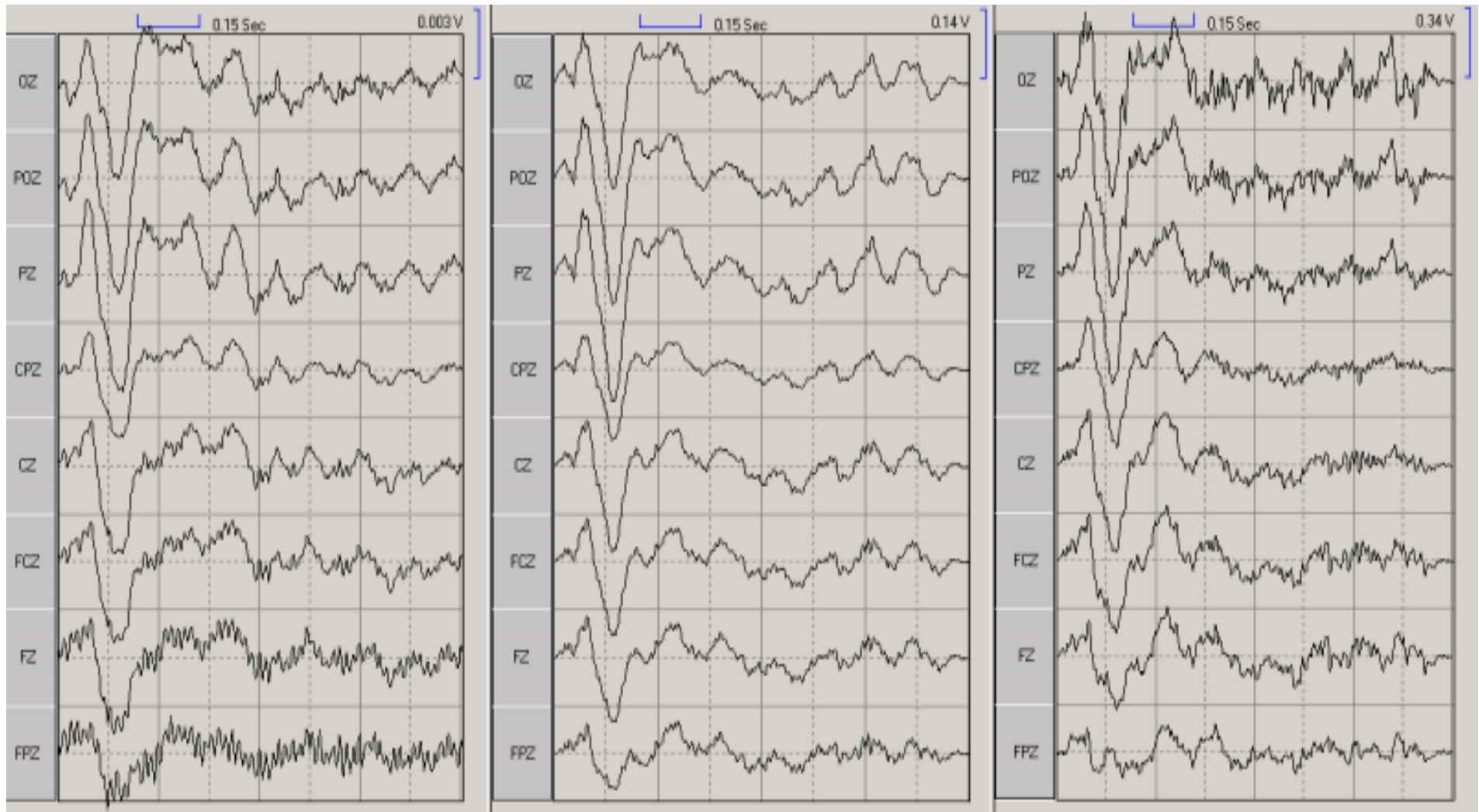
Continuous EEG (subset of 124 channels, 500 Hz, bandpass DC– 100 Hz) with stimulus events (1 = pattern onset, 2 = pattern offset; 1-2 duration = 100 ms; 2-1 ISI = Poisson with mean = variance = 750 ms)



(a) Averaged

(b) Deconvolved: Linear

(c) Deconvolved: Nonlinear



Total time equals 1 second, with the visual pattern onset at the beginning. (a) shows the event-related average, which includes randomly overlapping events; (b) shows the temporally deconvolved waveforms (linear kernel) using a purely linear model, which assumes no temporal interactions between events; (c) shows the linear part of a nonlinear model that permits bilinear interactions between events. Although there is an unresolved absolute scaling issue for (b) and (c), relative scaling is correct.

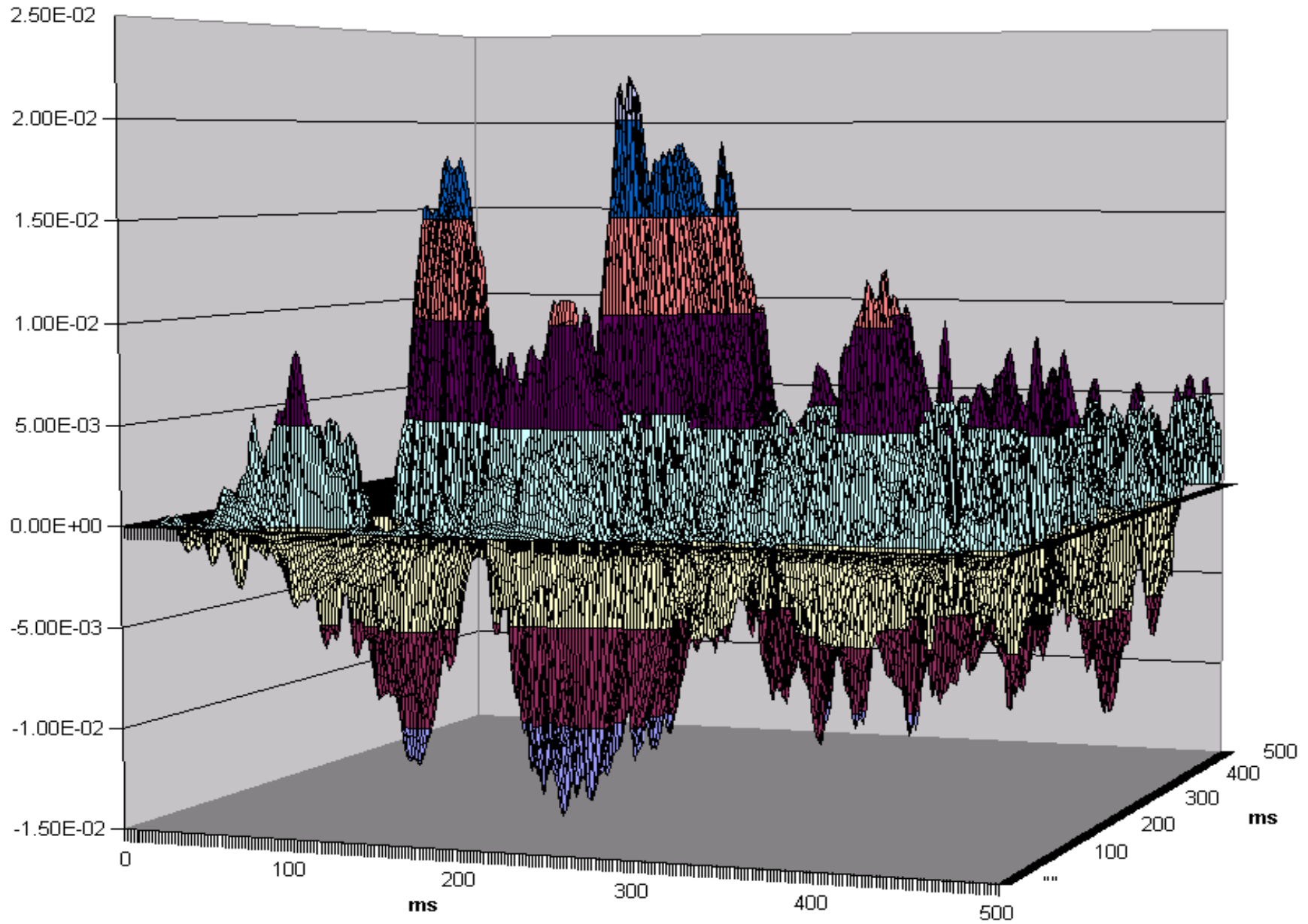


Some Observations

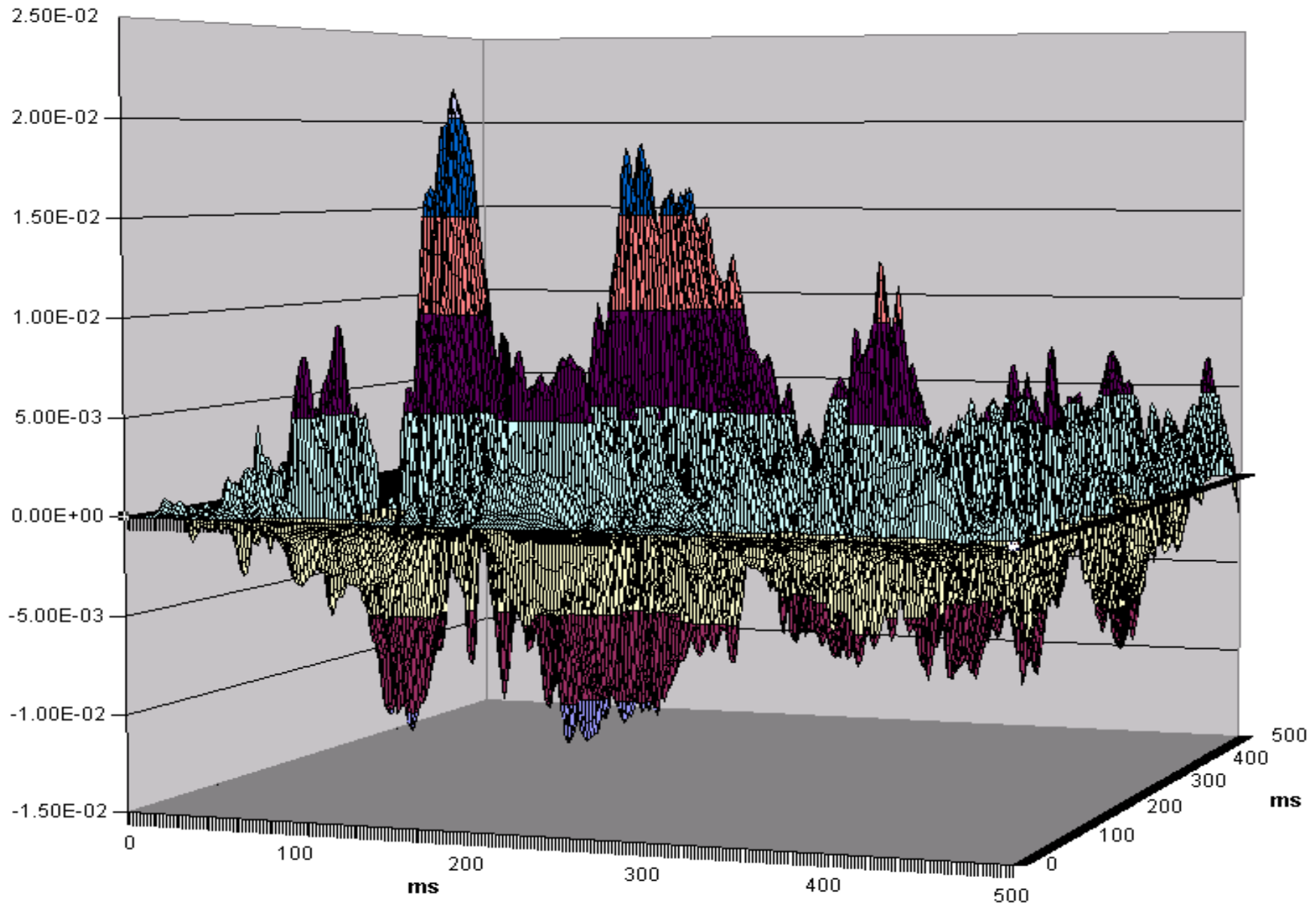
1. This confirms one thing we knew already, namely, that an event-related average approximates the impulse response (i.e., linear kernel) of the quasi-deterministic brain subsystem that generates EEG evoked potentials.
2. **Main Peaks:** Compared with the average (a), the main peaks of the linear deconvolution (b) are generally better defined. Compared with (b), the main peaks of the nonlinear deconvolution (c) appear even more “sharpened” and possibly even “fractal”.
3. Channels OZ, POZ, and PZ have a broad peak between ~240 ms and ~360 ms (with sub-peaks). In (a) and (b), this is roughly a “sagging plateau”. However, in (c) there is a marked incline to the plane.
4. There is a peak at about 400 ms in the average (a) (see channel PZ) that is greatly attenuated in the linear deconvolution (b) and apparently absent in the nonlinear deconvolution (c).
5. Noise that is apparent in the average at FZ and FPZ is largely absent in (b) and (c). This is probably because we chose the basis functions for their high signal-to-noise ratios. Thus, there seems to be at least some built-in resistance to noise in the deconvolution procedures, even though there was no prewhitening here.
6. Some late (~750 ms) peaks are enhanced in (b) and (c) at channels OZ, POZ, and PZ, which are attenuated in (a).



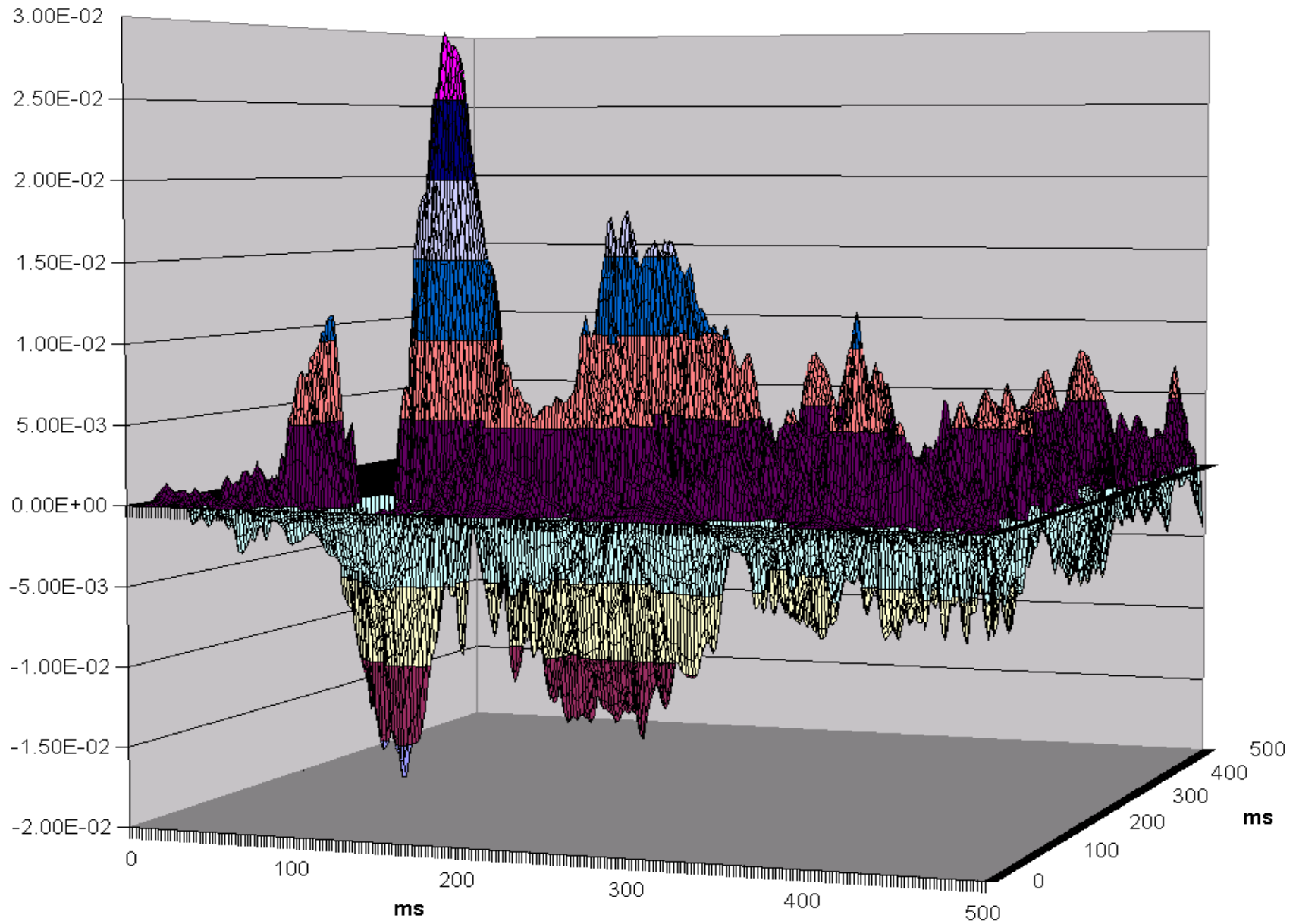
Self-Interaction Kernel Estimated at OZ



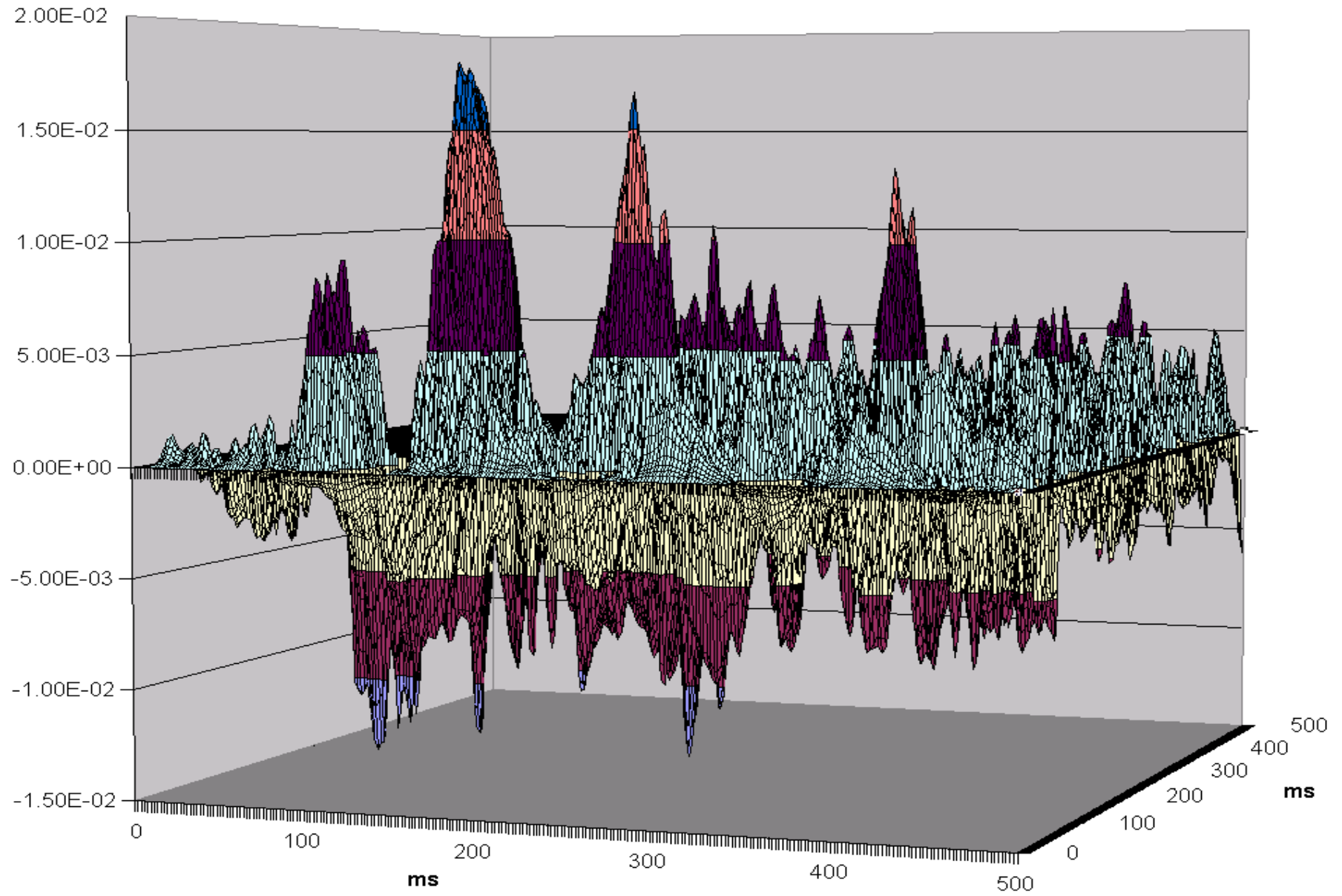
Self-Interaction Kernel Estimated at POZ



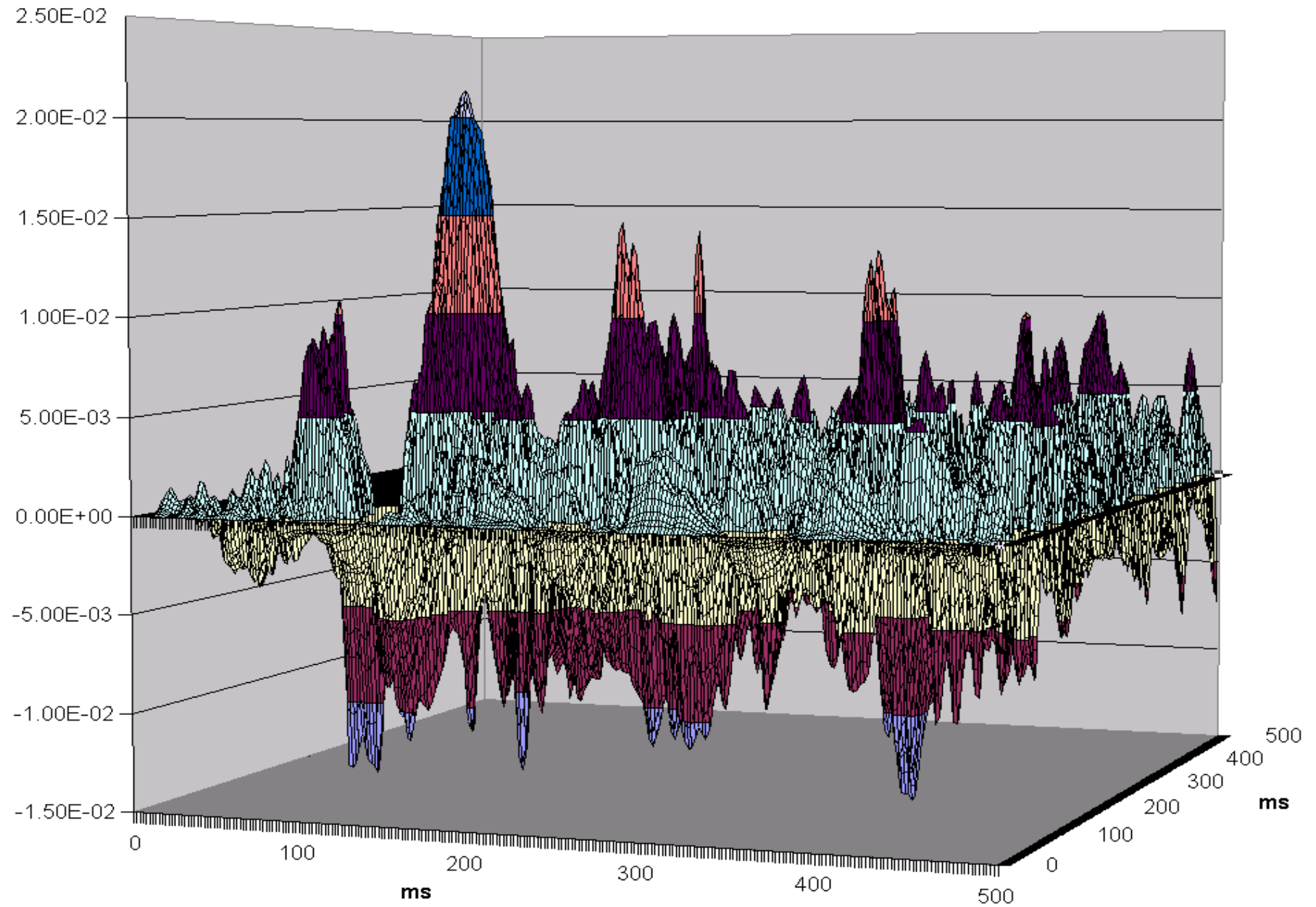
Self-Interaction Kernel Estimated at PZ



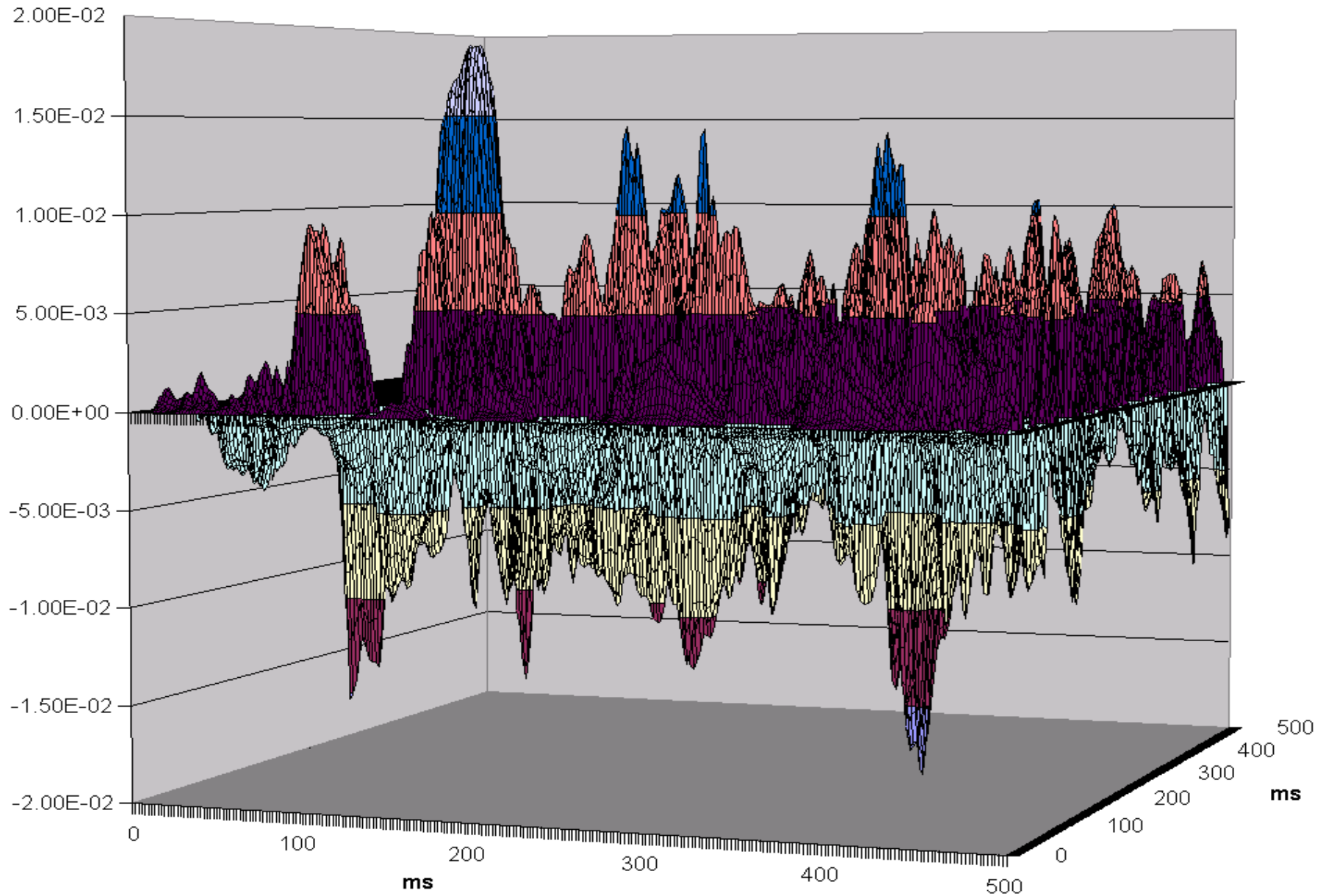
Self-Interaction Kernel Estimated at CPZ



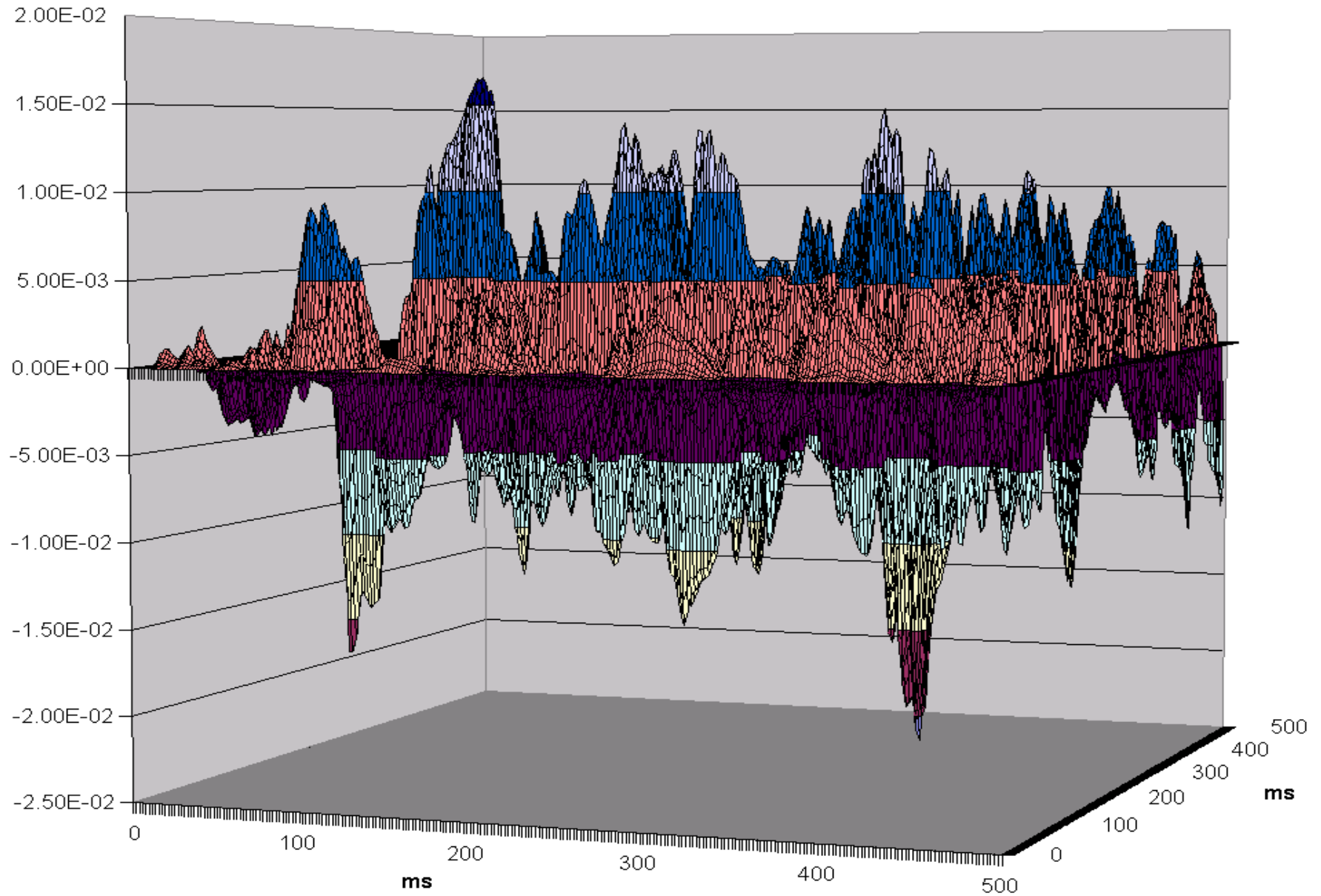
Self-Interaction Kernel Estimated at CZ



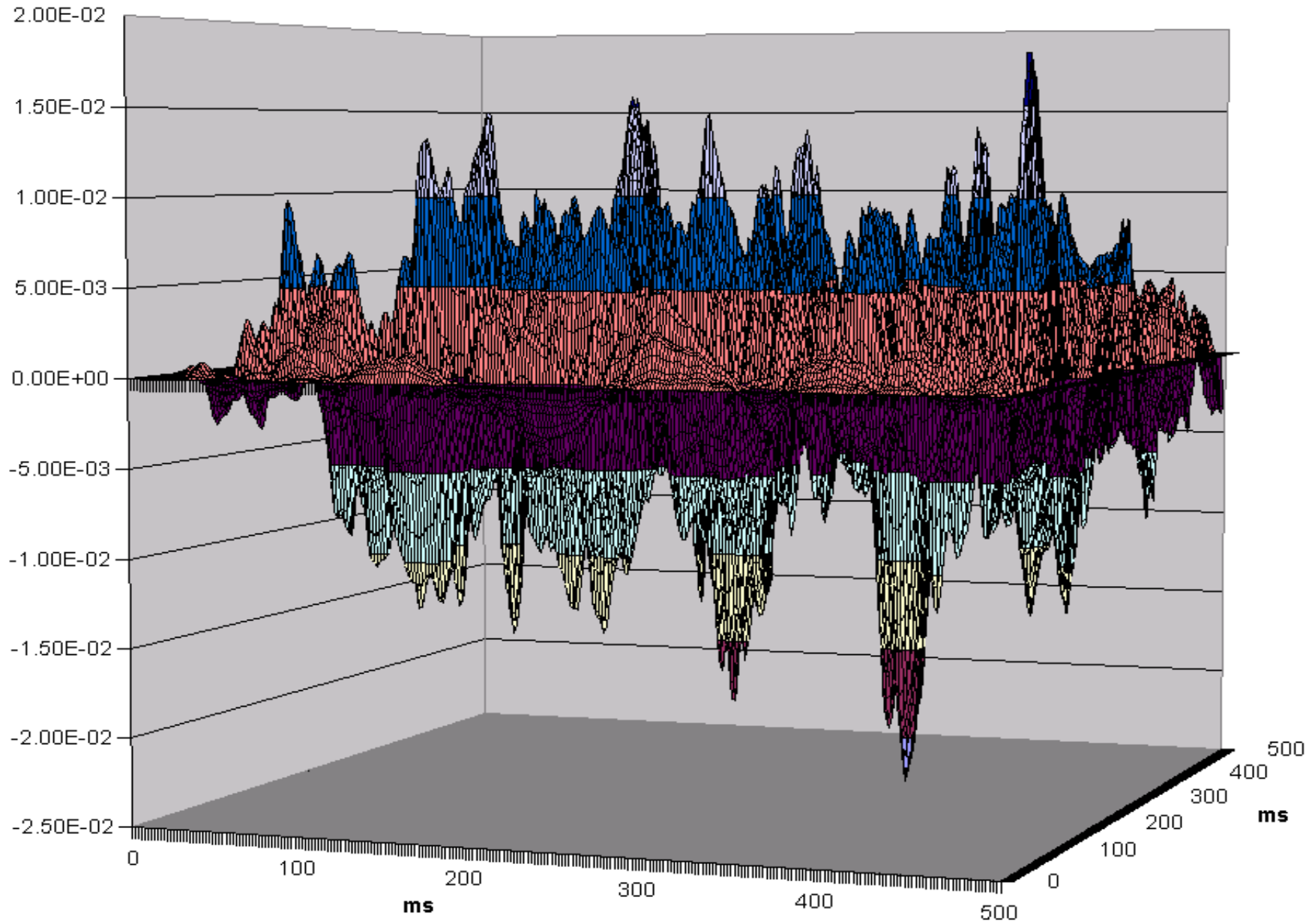
Self-Interaction Kernel Estimated at FCZ



Self-Interaction Kernel Estimated at FZ



Self-Interaction Kernel Estimated at FPZ



Of what use might this be?

1. Time will tell. However, from the above series of self-interaction kernels along the midline, we can detect some topographic patterns of peaks and troughs in the kernels. This is reminiscent of regular ERPs, with the addition of a second temporal axis.
2. A peak in an interaction kernel corresponds with an “enhancement” effect, whereas a trough corresponds with a “suppression” effect. These might be telling us something about underlying excitatory and inhibitory brain dynamics. These might even convey information of potential clinical relevance, e.g., along the lines of the P50 suppression effect.
3. Aside from the second order analyses, the temporal deconvolution procedure is an alternative way to remove temporal overlap from ERPs presented at a fast rate (Woldorff and Hillyard, 1991; Woldorff, 1993), or stimulus-related from response-related components (Hansen, 1983; Zhang, 1998). Similar methods are useful in the analysis of fMRI data (Glover, 1999).
4. Although task-related Volterra framework presented here has been applied to EEG data, Friston and colleagues have applied a similar framework to fMRI data. Thus, a common set of tools could be developed for analyzing both event-related EEG and event-related fMRI data.



References

- Bendat JS (1998): *Nonlinear System Techniques and Applications*. New York: John Wiley.
- Friston KJ, Josephs O, Rees G, & Turner R (1998): Nonlinear event-related responses in fMRI. *Magn Reson Med* 39:41-52.
- Friston JK, Josephs O, Zarahn E, Holmes AP, Rouquette S, Poline J (2000): To smooth or not to smooth? Bias and efficiency in fMRI time-series analysis. *Neuroimage* 12(2): 196-208.
- Friston KJ, Mechelli A, Turner R, Price CJ (2000): Nonlinear responses in fMRI: the Balloon model, Volterra kernels, and other hemodynamics. *Neuroimage* 12(4): 466-77.
- Glover GH (1999): Deconvolution of impulse response in event-related BOLD fMRI. *Neuroimage* 9(4):416-29.
- Hansen JC (1983): Separation of overlapping waveforms having known temporal distributions. *J Neurosci Methods* 9: 127-139.
- Pflieger ME, Nakada T (2000). The spatial resolving power of high-density EEG: An assessment of limits. In: T. Nakada (ed.), *Integrated Human Brain Science: Theory, Method, Application (Music)*. Elsevier Science.
- Schetzen M (1980): *The Volterra and Wiener Theories of Nonlinear Systems*. New York: John Wiley.
- Woldorff MG, Hillyard SA (1991): Modulation of early auditory processing during selective listening to rapidly presented tones. *Electroencephalogr Clin Neurophysiol* 79(3): 170-91.
- Woldorff MG (1993): Distortion of ERP averages due to overlap from temporally adjacent ERPs: analysis and correction. *Psychophysiology* 30(1): 98-119.
- Zarahn E, Aguirre G, D'Esposito M (1997): Empirical analyses of BOLD fMRI statistics. I. Spatially unsmoothed data collected under null-hypothesis conditions. *Neuroimage* 5: 179-97.
- Zhang J (1998): Decomposing stimulus and response component waveforms in ERP. *J Neurosci Methods* 80(1): 49-63.

