Biol. Cybern. 72, 527–532 (1995)

Biological Cybernetics

Nonlinear autoregressive analysis of the 3/s ictal electroencephalogram: implications for underlying dynamics

Nicholas D. Schiff, Jonathan D. Victor, Annemarie Canel

Department of Neurology and Neuroscience, The New York Hospital-Cornell Medical Center, 525 East 68 Street, New York City, NY 10021, USA

Received: 6 April 1994/Accepted in revised form: 18 November 1994

Abstract. In a previous study, nonlinear autoregressive (NLAR) models applied to ictal electroencephalogram (EEG) recordings in six patients revealed nonlinear signal interactions that correlated with seizure type and clinical diagnosis. Here we interpret these models from a theoretical viewpoint. Extended models with multiple nonlinear terms are employed to demonstrate the independence of nonlinear dynamical interactions identified in the 'NLAR fingerprint' of patients with 3/s seizure discharges. Analysis of the role of periodicity in the EEG signal reveals that the fingerprints reflect the dynamics not only of the periodic discharge itself, but also of the fluctuations of each cycle about an average waveform. A stability analysis is used to make qualitative inferences concerning the network properties of the ictal generators. Finally, the NLAR fingerprint is analyzed in the context of Volterra-Weiner theory.

1 Introduction

We introduced the nonlinear autoregressive (NLAR) 'fingerprint' as a pattern of nonlinear signal interactions that characterizes an ictal electroencephalogram (EEG) discharge (Schiff et al. 1991, 1994). This NLAR fingerprint was demonstrated to correlate with the clinical diagnosis and type of seizure discharge. Here we present a theoretical analysis and interpretation of the NLAR fingerprint within the context of nonlinear systems theory. A detailed analysis of NLAR models with multiple nonlinear terms allows the statistical significance of separate nonlinear interactions in the fingerprint to be examined, which later provides a basis to interpret the NLAR fingerprint in the context of Volterra-Wiener theory. The role of periodicity in the EEG waveforms is addressed through analysis of 'periodicized' data. Furthermore, the nonlinear difference equation form of the NLAR models is used to make qualitative inferences concerning the network properties of the generators of ictal activity.

As with linear autoregressive (LAR) and other models, description of the activity of populations of cortical neurons as recorded in the EEG does not specify the microdynamics of the underlying cellular elements (Wilson and Cowan 1972). Rather, the aim is to provide a window on the overall circuit properties and feedback interactions underlying ictal activity. The specific patterns of nonlinear interactions evident in the NLAR fingerprints are interpreted in the context of simple nonlinear feedback systems treated in Volterra-Wiener theory.

2 Methods

EEG records from six seizure patients were studied. The procedure for obtaining artifact-free samples of ictal discharges and clinical information is presented in Schiff et al. (1994). We model the EEG using a NLAR equation in which a sample y_n of the EEG signal is represented as a sum of a random term, x_n, a linear combination of values of the EEG at r prior times, and a quadratic term which allows EEG values at two prior times to interact. A model with all such possible terms takes the form

$$y_n = x_n - a_0 - \sum_{i=1}^{r} a_i y_{n-i} - \sum_{j=1}^{r} \sum_{k=1}^{r} c_{j,k} y_{n-j} y_{n-k} \qquad (1)$$

This model and our approach to fitting the model parameters a_i and $c_{j,k}$ are discussed in detail in Schiff et al. (1994). Therein, we focused on models of the form (1) in which only one term $c_{j,k}$ is included. The residual variance for the one-term NLAR model necessarily depends on the choice of lags (j,k) for the single included nonlinear term, and the dependence was called the NLAR fingerprint. Here we expand our focus to include models of the form (1) with up to three nonlinear terms.

Correspondence to: N. D. Schiff

3 Results

3.1 Independence of fingerprint features

Nonlinear fingerprints derived from ictal discharges display multiple local minima and ridges [i.e., loci of lags (j,k) which correspond to nonlinear terms (1) which provide a substantial reduction in variance]. Each of these large reductions in variance corresponds to a quadratic interaction of signal values at prior times, which contributes to the current value. Since there are strong first-order correlations between serial values of the ictal EEG, a quadratic interaction at one pair of lags might also be manifest at a second pair of lags. This would result in a single, nonlinear interaction generating two or more minima within a fingerprint. The first question we address is whether these observed multiple minima represent independent dynamical features. We constructed NLAR models with two or more nonlinear terms, with each term drawn from a separate local minimum. As in Schiff et al. (1994), we used the Akaike criterion (AIC) to estimate the significance of the reduction in variance accompanying the addition of a nonlinear term. The AIC is similarly:

$$\mathrm{AIC} = N \log V + 2J \qquad (2)$$

where N represents the number of data points, J is the number of model terms, and V is the residual variance. Further rationale for the use of the AIC may be found in Victor and Canel (1992). We then compared the AIC obtained from this NLAR model with multiple nonlinear terms to the AIC obtained from NLAR models containing only one of the nonlinear terms. If the selected minima represented the same nonlinear interaction, then the addition of multiple terms would not improve the model, i.e., the AIC should not decrease. Conversely, if these minima represented statistically independent influences, then the AIC would be expected to decrease.

Figure 1 shows the results of this analysis applied to the second data segment from patient 1. We considered pairwise and three-way combinations of the four minima that were most consistent across patients: lags (1, 9), (5, 8), (8, 17), and (15, 15). As seen in Fig. 1, all pairwise combinations of non-linear terms resulted in a decrease in the AIC. Furthermore, models with three nonlinear terms had a lower AIC than models with only two terms. This argues for a statistical independence of the interactions.

The AIC is only an approximate measure of significance for NLAR models with two or more terms (Victor and Canel 1992). Therefore, as an internal control for this procedure, we also chose two nonlinear terms from the same feature of the fingerprint. Inclusion of lags (5, 8) and (6, 9), both within the off-diagonal ridge, resulted in a rise in the AIC. This indicates that the interactions at these lags are not statistically independent and confirms the intuition that this ridge is a single feature.

Analysis of NLAR models with multiple nonlinear terms derived from the minima listed in Table 2 in Schiff et al. (1994) revealed a similar finding in patients 2–5. That is, minima which were well-separated on the fingerprints resulted in reductions in the AIC and thus reflect statistically independent interactions.

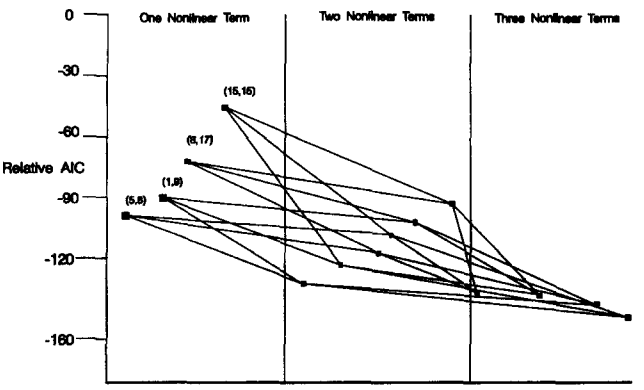

Fig. 1. Akaike criterion (AIC) values for nonlinear autoregressive (NLAR) models composed of one, two, and three non-linear terms. In all cases, combinations of two or three nonlinear terms from distinct minima resulted in a model whose AIC was less than that of models containing a single quadratic term

3.2 Role of near-periodicity in the data

The highly stereotyped and nearly periodic nature of the 3/s spike wave discharge in some patients raises the question of whether the fingerprint structure merely reflects a basic periodic waveform. We investigated this by applying NLAR analysis to a time series derived from an ictal record by averaging the signal with respect to its repeat period. To implement the averaging process, the duration of the repeat period was determined by dividing the time between the first and last spikes by the number of intervening periods.

Figure 2A shows the NLAR fingerprint of such 'periodicized' data derived from the ictal record analyzed in detail in Schiff et al. (1994) (for fingerprints of periodicized data, we do not indicate significance levels, because the assumptions underlying the AIC are not valid). The fingerprint derived from strictly periodic data has lost all of the features of the fingerprint (Fig. 3A) derived from the original record. The minimum at lag (8, 17), a feature common to the fingerprints of the 3/s discharges in all five patients, is absent.

A similar analysis of periodicized data from patients 3–5 from Schiff et al. (1994) also yielded obvious distortions of the fingerprint. For example, Fig. 3B shows the fingerprint of periodicized data from patient 4. Again, the minimum at lag (8, 17), a main feature of the fingerprint of the raw data, is absent. The data from patient 2 could not be analyzed in this fashion because the fragmented nature of the ictal discharge prevented determination of an accurate period to serve as the basis for averaging. However, the observation that the NLAR fingerprint derived from this fragmented record (Fig. 4A of Schiff et al. 1994) is nearly identical to that derived from the continuous discharge of patient 1 (Fig. 3A) provides further evidence that the NLAR features do not merely represent periodicity in the data.

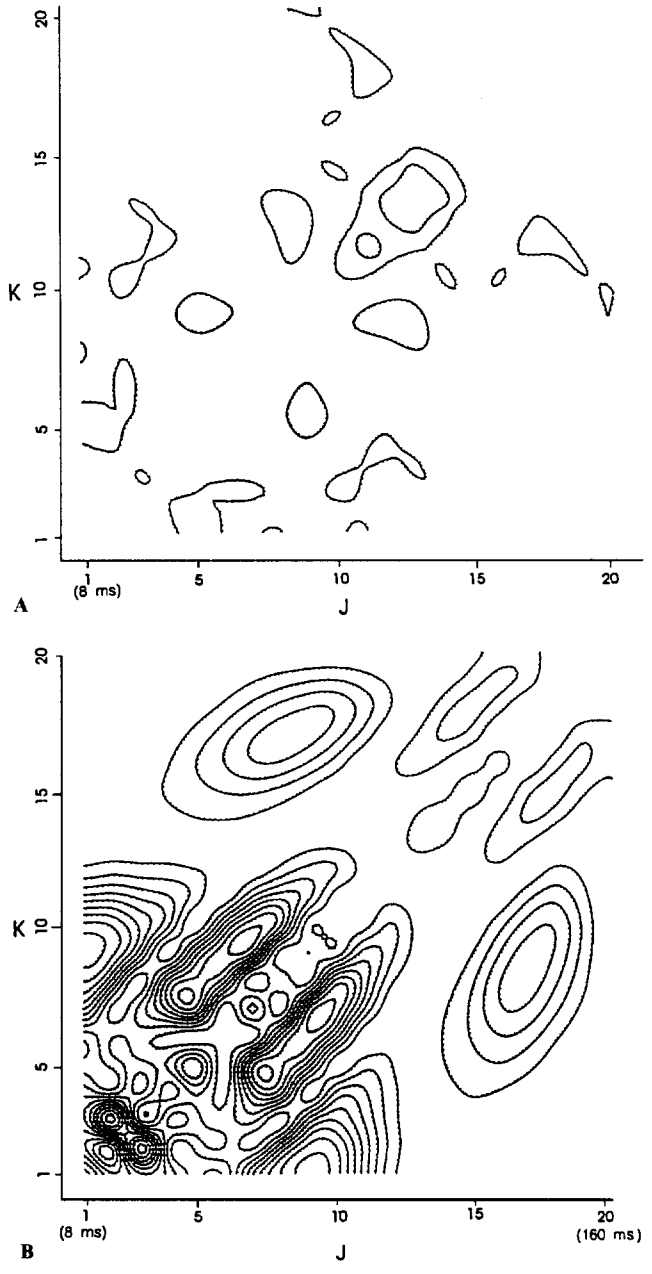

Fig. 2. **A** NLAR fingerprint derived from an ictal discharge recorded from patient 1 (Fig. 3A), after averaging the data with respect to its mean repeat period. Each *contour line* represents 0.25% of the variance of the original data. **B** NLAR fingerprint derived from an ictal discharge recorded from patient 1 (Fig. 3A), after averaging the data with respect to a repeat period which stretched uniformly from the beginning to the end of the record

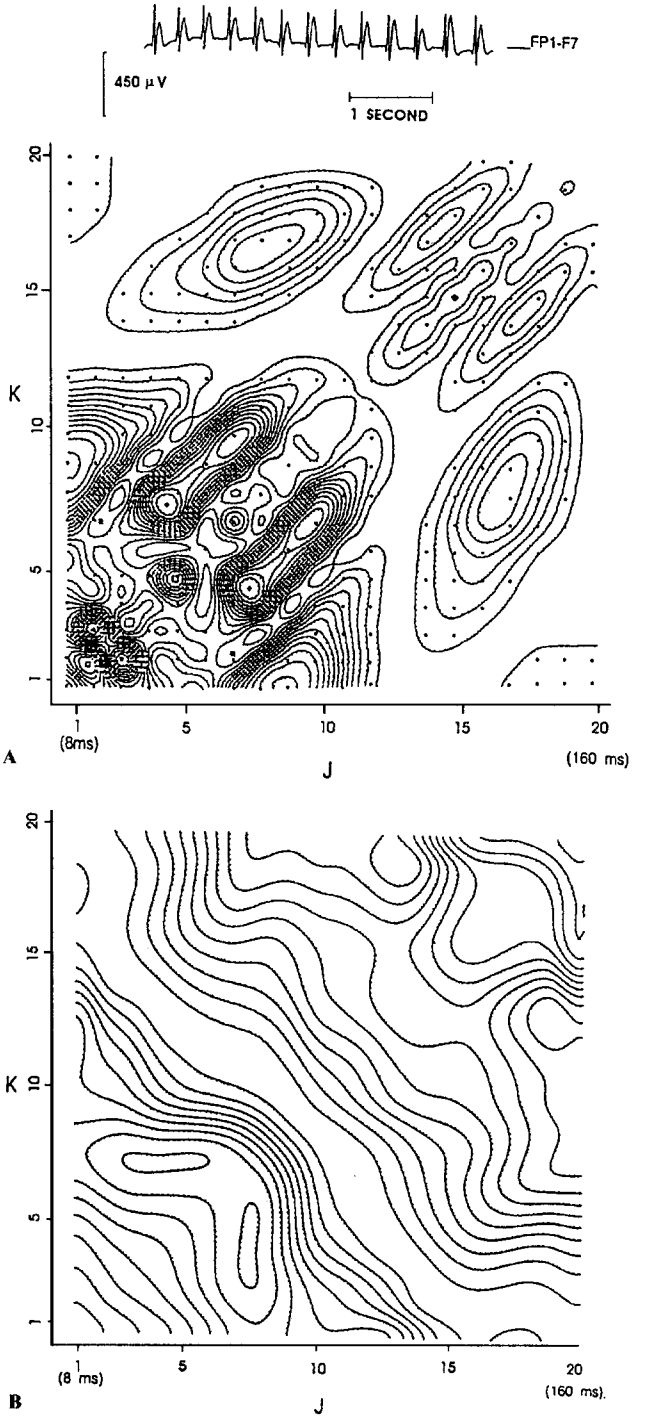

Fig. 3. **A** NLAR fingerprint derived from patient 1 of Schiff et al. (1994) reproduced. Each time unit (j or k) represents 8 ms. *Tickmarks* point downhill, and each *contour line* represents 0.25% of the variance. Statistical significance is indicated by placing a *small dot* at the coordinates (j,k) for which the reduction in variance satisfies $N\Delta V_{j,k} > 4$. **B** NLAR fingerprint derived from an ictal discharge recorded from Patient 4 of Schiff et al. (1994), after averaging the data with respect to its mean repeat period. Each *contour line* represents 0.01% of the variance of the original data

Inspection of the ictal discharge from patient 1 (Fig. 3A) shows that the period of the discharge slows down with time. Thus, averaging with respect to an overall repeat period as described above might combine together noncorresponding parts of the wave. To reduce this distortion, we also averaged the record of Fig. 3A with respect to a period whose duration stretched uniformly from the beginning to the end of the record. Figure 2B shows the NLAR fingerprint derived from this record. The qualitative features of the fingerprint have indeed been recovered, although the minima are only about 2/3 as deep as in the original NLAR fingerprint

(Fig. 3A). This indicates that the fingerprint depends not only on the average 3/s waveform, but also on cycle-to-cycle variations and fluctuations about this average. The EEG features that shape the fingerprint cannot be reproduced by simply a gradual stretching of an average waveform. Rather, the fingerprint represents, in part, more subtle deterministic period-to-period variations in the discharge.

3.3 Stability analysis

The NLAR models which underlie the fingerprints are dynamical systems which evolve in discrete time. Fixed points and periodic orbits are key indicators of the behavior of such systems. In this section we examine the existence and stability of fixed points and two-cycles of the NLAR models to provide insight into the relationship of the linear and nonlinear terms. Following Feigenbaum (1983), we define a vector state variable $Y_n = (y_n, y_{n-1}, \ldots)$, and a vector equation:

$$Y_n = \vec{\mathbf{D}}(Y_{n-1}) + X_n \tag{3}$$

The stability of the autonomous version of this system ($X_n = 0$) at a fixed point $Y_{\text{fix}} = \mathbf{D}(Y_{\text{fix}})$ depends on the eigenvalues of the matrix transformation which best approximates $\mathbf{D}$ at the fixed point. If the maximum eigenvalue of this matrix has absolute value less than 1, then the fixed point is stable. If the maximum eigenvalue of this matrix has absolute value greater than 1, then the fixed point is unstable.

3.3.1 Findings. The results of this analysis applied to the data of patient 1 (ictal discharge of Fig. 3B, Schiff et al. 1994) are shown in Table 1. The LAR model has a single stable fixed point. NLAR models with one, two, or three quadratic terms all had two fixed points, one relatively near the fixed point of the LAR model (within 27 μV), and one at a great distance (106–2015 μV). The unexplained variance for these models was typically 700 μV^2, which corresponded to a standard deviation for the driving terms x_n of approximately 26 μV or less. That is, the 'near' fixed point was always within the range explored by the driving terms, and the second fixed point was never in that range. The maximum eigenvalue at the near fixed point of the NLAR model ranged from 0.962 to 1.025: stability in 10 of 14 cases, minimal instability in the remaining 4 cases. The maximum eigenvalue associated with the distant fixed point always indicated decidedly unstable behavior: it ranged from 1.087 to 5.542. A two-cycle was present for one of the 14 NLAR models shown in Table 1. This two-cycle was unstable, and positioned far away from the fixed point of the linear model.

This analysis, and a similar analysis of the models derived from the other data sets, shows that the addition of one or more quadratic terms results in (i) a small shift of the fixed point of the LAR model and (ii) a new, unstable fixed point far from the fixed point of the linear model. Because the second fixed point is generally far beyond the range explored by the time series, it is not relevant to the dynamics of the ictal discharge. Two-cycles, when present, were always unstable and positioned far from the range explored by the time series. That is, the occasional two-cycles appear to be consequences of the quadratic fit per se, and not of the intrinsic dynamics.

We emphasize that the tools and concepts used here (stability analysis, fixed points, two-cycles) are also used by those investigating chaotic dynamics, but here they play a different role: the stability analysis of the models evidently does not reflect the global stability properties of the EEG. The NLAR models represent a small adjustment to the LAR model valid for a particular range of signal size. The global behavior of the NLAR model is not necessarily related to global dynamics of the EEG.

Table 1. Stability analysis of LAR and NLAR models derived from a 3/s seizure from patient 1 of Schiff et al. (1994). Positions of fixed points in μV; variance in μV^2. Data from Fig. 3B of Schiff et al. (1994)

Nonlinear terms	Fixed points	Maximum eigenvalue	Variance
None (LAR only)	0.3	0.9757	945
(1,9)	122.4	1.2390	797
	− 4.5	1.0063	
(5,8)	− 157.4	1.1449	778
	3.9	0.9620	
(8,17)	− 155.1	1.1142	826
	3.3	1.0249	
(15,15)	− 203.3	1.0868	866
	27.1	0.9801	
(1,9) (5,8)	2014.9	5.5418	705
	0.6	0.9479	
(1,9) (8,17)	712.5	2.7178	756
	− 1.4	1.0171	
(1,9) (15,15)	251.1	1.5630	729
	25.8	0.9799	
(5,8) (8,17)	− 118.2	1.1224	736
	4.3	0.9905	
(5,8) (15,15)	− 126.6	1.1130	750
	16.8	0.9811	
(15,15) (8,17)	− 121.4	1.1001	777
	17.8	1.0218	
(1,9) (5,8) (15,15)[a]	− 606.0	1.2375	687
	15.9	0.9652	
(5,8) (8,17) (15,15)	− 108.6	1.1110	722
	14.1	0.9926	
(1,9) (8,17) (15,15)	− 792.1	1.3104	720
	19.7	0.9994	
(1,9) (5,8) (8,17)	− 489.4	1.1877	701
	0.7	0.9755	

[a] Model exhibits an unstable two-cycle (− 792.7, − 1219.1)

4 Discussion

4.1 Interpretation of the fingerprints

Before suggesting a possible means to interpret the fingerprints, we would like to exclude certain alternative interpretations. One possibility is that the fingerprints' structure was somehow related to discrete sampling of the data. This possibility is ruled out by a comparison of NLAR fingerprints obtained from records sampled at

different intervals done in Schiff et al. (1994). The lag numbers associated with positions of features in the fingerprints depended on the sampling rate, in a manner that the lag times they represented were independent of sampling rate.

A second possibility is that the fingerprints are a trivial consequence of the highly periodic nature of the ictal records analyzed. The evidence against this is (a) the fingerprints derived from nearly continuous 3/s records (patient 1, Fig. 2 of Schiff et al. 1994) were similar to those derived from very fragmented records (patient 2, Fig. 3A of Schiff et al. 1994); (b) the fingerprint of a highly periodic, complex, partial seizure discharge (patient 6, Fig. 4B of Schiff et al. 1994) was dramatically different from the fingerprints of all of the 3/s discharges and (c) fingerprints of 'periodicized' data (Figs. 2 and 3) differed from fingerprints of the original records.

A third possibility was that the quadratic NLAR models which underlie the fingerprints involved an improved representation of the global dynamics of the ictal EEG. This potentially very interesting possibility was ruled out by an analysis of the location and character of stable points and two-cycles (Table 1).

In view of these findings, we think it is reasonable to interpret NLAR fingerprints in a manner which generalizes the interpretation of LAR models. A LAR model is a difference equation which corresponds to the differential equation of a physical system whose output y_n is the result of applying linear feed-back to a noise input x_n. The coefficients a_i of the LAR model represent the impulse response of the linear feedback at time lag i. For a NLAR model, the NLAR terms may be considered to represent nonlinear feedback. In this interpretation, the coefficients $c_{j,k}$ represent a generalization of the impulse response, expressing the interaction of inputs at time lags (j,k). That is, the array of coefficients $c_{j,k}$ represent an approximation to the second-order Volterra kernel (Marmarelis and Marmarelis 1978) of a non-linear feedback filter. The justification for this interpretation (despite the fact that each $c_{j,k}$ was calculated individually) is that: (i) individual non-linear terms, though significant, account for a relatively small portion of the residual variance, and (ii) individual terms from different minima represent independent contributions to the model (Fig. 1). We emphasize that the contour maps do not represent second-order Volterra kernels of a 'system' whose output is the EEG. In principle, it is not possible to calculate a Volterra kernel for the input-output relation without access to the inputs to the system. To an extent, the fingerprint of nonlinear terms approximates a Volterra kernel. However, it is the kernel of a presumed nonlinear feedback loop between output and assumed white-noise input, rather than that of the input-output relationship itself.

Once the array of individually fitted $c_{j,k}$ are considered to be akin to the second-order Volterra kernel $K_2(t_1,t_2)$ of an effective feedback, standard tools of the Volterra-Wiener theory (Korenberg 1973; Marmarelis and Marmarelis 1978) may be used to interpret these maps. [As shown in Victor and Canel (1992), the fingerprints, which are maps of the reduction in variance, correspond closely to the arrays $c_{j,k}$.] There is no general procedure for deducing a system model from a kernel, but it is possible to ask whether certain simple nonlinear systems are consistent with the observed fingerprints. For a non-linear system consisting of a linear filter with impulse response $G(t)$ followed by a static nonlinearity, the second-order kernel has the form:

$$K_2(t_1,t_2) = aG(t_1)\,G(t_2) \tag{4}$$

The second-order kernel of a system consisting of a static nonlinearity followed by a linear filter with impulse response $H(t)$ has the form:

$$K_2(t_1,t_2) = \begin{cases} aH(t_1), & t_1 = t_2 \\ 0, & t_1 \neq t_2 \end{cases} \tag{5}$$

For a system consisting of a linear filter with impulse response $G(t)$, followed by a static nonlinearity, followed by a second linear filter with impulse response $H(t)$, the second-order kernel has the form:

$$\int_0^\infty aG(t_1 - t)\,G(t_2 - t)\,H(t)\,\mathrm{d}t \tag{6}$$

Equation (4) implies that the second-order kernel of linear → static nonlinear systems has prominent contours running parallel to the axes, and the largest values on the diagonal $t_1 = t_2 (j = k)$. This is not seen in any of the fingerprints of 3/s discharges. The square appearance of the fingerprint of the complex partial seizure record (Fig. 4B, Schiff et al. 1994) is suggestive of this feature, with $G(t)$ peaking at approximately 48 ms (lag 6).

Equation (5) implies that the second-order kernel of static nonlinear → linear systems would have all of its significant values on the diagonal. This is not seen in any of the fingerprints. Finally, (6) implies that the second-order kernel would look similar to that of (4), but with peaks spread out parallel to the diagonal according to $H(t)$. Such spreading can produce an off-diagonal ridge, but only if the ridge is accompanied by an on-diagonal feature of similar size. However, the fingerprints from all three patients with isolated 3/s seizures had off-diagonal ridges (24 ms off the diagonal, ranging from lags of 40 to 60 ms) without an accompanying on-diagonal feature at corresponding lags.

The simple cascade systems described by (4)–(6) cannot account for the qualitative features of the NLAR fingerprints of 3/s seizure discharges. While it is useful to be able to exclude models, it would be more satisfying to also identify models that are consistent with the fingerprints. One possibility is that each of the fingerprint features represent nonlinear interactions which are not only statistically independent, but also biologically independent – and that a more appropriate model is a parallel combination of several such cascades (Korenberg 1987). This possibility is suggested by the observation that in patients 4 and 5 (from Schiff et al. 1994), some of the features of the fingerprint are lost. Parallel combinations of cascade systems might reproduce the observed fingerprints, but such models have so many parameters, that it is unclear whether such a fit has any explanatory value in this context.

A more satisfying mechanistic interpretation of the fingerprints would rest on reproducing the observed kernels with a specific model of nonlinear dynamics, such as that proposed by Wilson and Cowan (1972). However, these simple models cannot account for the difference between the fingerprints of the periodicized data and those of the original data. The minima of the fingerprint of an ictal discharge differ from those obtained from periodicized data, even after stretching the waveform. This implies that there are contributions to the fingerprint related to cycle-to-cycle variations of the discharge, including (but not limited to) gradual stretching in time. Models of ictal discharges based on limit cycle behavior (Wilson and Cowan 1972; Kawahara 1980) are unlikely to have this feature.

Acknowledgements. N.S. was supported by a Howard Hughes Medical Student Research Training Fellowship and a medical student fellowship from the Epilepsy Foundation of America. J.V. was supported in-part by EY7977 and EY9314 from the National Eye Institute and the Klingenstein Foundation. We thank Mary Conte for expert technical assistance.

References

Akaike H (1974) A new look at statistical model identification. IEEE Trans Auto Control AC-19:716–723

Farmer JD, Ott E, Yorke JA (1983) The dimension of chaotic attractors. Physica 7D:153–180

Feigenbaum M (1983) Universal behavior in nonlinear systems. Physica 7D:16–39

Kawahara T (1980) Coupled Van der Pol oscillators – a model of excitatory and inhibitory neural interactions. Biol Cybern 39:37–43

Korenberg MJ (1973) Identifications of biological cascades of linear and static nonlinear systems. Proc 16th Midwest Symp Circuit Theory 18:1–9

Korenberg MJ (1987) Functional expansions, parallel cascades, and nonlinear difference equations. In: Marmarelis VZ (ed) Advanced methods of physiological system modelling. University of Southern California, Los Angeles

Marmarelis PZ, Marmarelis VZ (1978) Analysis of physiological systems: the white noise approach. Plenum, New York

Schiff ND, Victor JD, Canel A, Labar DR (1991) Nonlinear autoregressive analysis of ictal EEG records (Abstract P32.40). Third IBRO World Congress of Neuroscience, Montreal, p 223

Schiff ND, Victor JD, Canel A, Labar DR (1995) Characteristic nonlinearities of 3/second ictal EEG identified by nonlinear autoregressive analysis. Biol Cybern 72:519–526

Victor JD, Canel A (1992) A relation between the Akaike criterion and reliability of parameter estimates, with application to nonlinear autoregressive modelling the ictal EEG. Ann Biomed Eng 20:167–180

Wilson HR, Cowan JD (1972) A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. Kybernetik 13:55–80