Available online at www.sciencedirect.com
SCIENCE DIRECT®
Mathematical Biosciences 196 (2005) 1–13
Mathematical Biosciences
www.elsevier.com/locate/mbs

General methodology for nonlinear modeling of neural systems with Poisson point-process inputs

V.Z. Marmarelis *, T.W. Berger

University of Southern California, Biomedical Engineering, 3650 McClintock Avenue, Olin Hall 500, Los Angeles, CA 90089-1415, United States

Received 10 March 2003; received in revised form 24 March 2005; accepted 1 April 2005
Available online 15 June 2005

Abstract

This paper presents a general methodological framework for the practical modeling of neural systems with point-process inputs (sequences of action potentials or, more broadly, identical events) based on the Volterra and Wiener theories of functional expansions and system identification. The paper clarifies the distinctions between Volterra and Wiener kernels obtained from Poisson point-process inputs. It shows that only the Wiener kernels can be estimated via cross-correlation, but must be defined as zero along the diagonals. The Volterra kernels can be estimated far more accurately (and from shorter data-records) by use of the Laguerre expansion technique adapted to point-process inputs, and they are independent of the mean rate of stimulation (unlike their P–W counterparts that depend on it). The Volterra kernels can also be estimated for broadband point-process inputs that are not Poisson. Useful applications of this modeling approach include cases where we seek to determine (model) the transfer characteristics between one neuronal axon (a point-process 'input') and another axon (a point-process 'output') or some other measure of neuronal activity (a continuous 'output', such as population activity) with which a causal link exists.

Keywords: Volterra kernels; Wiener kernels; Nonlinear modeling; Poisson inputs; Point-process inputs; Neural systems; Neuronal modeling

* Corresponding author. Tel.: +1 213 740 0841; fax: +1 213 740 0343.
E-mail address: marmarelis@hotmail.com (V.Z. Marmarelis).

0025-5564/$ - see front matter
doi:10.1016/j.mbs.2005.04.002

1. Introduction

Modeling studies of neural systems that are stimulated by temporal sequences of action potentials (spike trains) can be placed in the general mathematical framework of functional expansions, due to the intrinsic system nonlinearities [1–3,7–9,12,10,13,11,14,17–19]. When these action potentials are idealized as impulses of fixed intensity (Dirac delta functions in continuous time or Kronecker deltas in discrete time), they can be represented mathematically in a stochastic context by 'point processes': a specific class of random processes that are formed by sequences of identical impulses representing random, instantaneous, and identical events.

These modeling studies must be properly placed in a stochastic context since actual experiments of neural systems are burdened inevitably by systemic or measurement noise and other uncontrollable factors introducing variations in the experimental data that are viewed as stochastic processes. Furthermore, the digital processing of the experimental data necessitates analysis in a discrete-time framework, although the actual biological processes take place in continuous time. Proper discretization of a sequence of action potentials requires that the sampling interval (bin-width) be approximately equal to the refractory period in order to allow the representation of each continuous-time action potential by one and only one Kronecker delta in the respective bin (the intensity of the latter ought to be the integrated area under the action potential divided by the bin-width T).

Thus, in practice, each sequence of action potentials is represented by a discrete-time point-process

$$\chi(n) = A \sum_{i=1}^{K} \delta(n - n_i), \tag{1}$$

where n denotes the discrete-time index ($t = nT$), A is the intensity of the Kronecker delta and n_i is the time-index of the ith event (i.e., discretized timing of the ith action potential). Note that this process has K events over the available data record of N bins, i.e., the mean rate of this point-process is (K/N). We seek to address the problem of neural system modeling from input–output data, where the system output $y(n)$ may be continuous or a point-process but the system input $\chi(n)$ is always a point-process described as in Eq. (1).

The potential utility of this modeling approach to the study of actual neural systems is found in those cases where we seek to determine (model) the transfer characteristics between one neuronal axon (a point-process 'input') and another axon (a point-process 'output') or some other measure of neuronal activity (a continuous 'output', such as population activity) with which a causal link exists (or must be confirmed). Another set of potential applications with point-process inputs are neural prostheses and in vitro studies (like the original studies in the hippocampal slices by Berger, Sclabassi and their associates that provided the initial motivation for this type of modeling work).

2. The modeling problem

In the general formulation of the modeling problem, we seek an explicit mathematical description of the causal functional F that maps the input past (and present) upon the present value of the output,

$$y(n) = F[\chi(n'), n' \leqslant n]. \tag{2}$$

For a stationary system with finite memory, the functional F may be represented (or approximated) by the discrete-time Volterra series [8,20]

$$y(n) = k_0 + \sum_{m=0}^{M} k_1(m)\chi(n-m) + \sum_{m_1=0}^{M}\sum_{m_2=0}^{M} k_2(m_1,m_2)\chi(n-m_1)\chi(n-m_2) + \cdots, \tag{3}$$

where $\{k_0, k_1, k_2, \ldots\}$ represent the Volterra kernels of the system that fully describe the functional F, and M denotes the finite system memory. Thus, the modeling problem becomes one of estimating as many kernels as deemed practical or necessary for satisfactory output prediction from experimental input–output data $\chi(n)$ and $y(n)$.

For the point-process input of Eq. (1), the system output is given by Eq. (3) as

$$y(n) = k_0 + A\sum_{i=1}^{K} k_1(n-n_i) + A^2\sum_{i=1}^{K}\sum_{j=1}^{K} k_2(n-n_i, n-n_j) + \cdots, \tag{4}$$

where the high-order terms (i.e., second-order and above) are non-zero only for $|n_i - n_j| \leqslant M$, and (i,j) represent all pairs of event indices. Clearly, the convergence of this functional power series depends, in general, on the specific system characteristics that distinguish low-order from high-order systems. In practice, rapid convergence is desirable, as it allows truncation of the Volterra series to a few terms for satisfactory model accuracy and yields relatively concise models.

The search for a kernel estimation method that minimizes the model prediction error (as measured by the output prediction mean-square error) leads us to the construction of an orthogonal hierarchy (series) of functionals using a variant of the Gram–Schmidt orthogonalization procedure, in order to decouple the various kernels and secure maximum reduction of the approximation error at each successive model order – analogous to the Wiener series in the case of continuous-input systems.

Critical for this orthogonalization procedure is the selection of the proper input that tests the system as densely as possible over the space of all possible inputs. Thus, for stochastic inputs, ergodicity is required as well as appropriate autocorrelation properties of all orders up to twice the highest-order functional (kernel) of a given system [7]. For systems with continuous inputs, such an appropriate test input is Gaussian white noise [21] or quasi-white approximations suitable for experimental investigations [7]. For systems with point-process inputs, the proper input is the Poisson process, defined in the discrete-time context as a sequence of independent events (spikes) with fixed probability of occurrence λ at each time bin T [4,5,15]. Thus, for a discrete-time Poisson input $\chi(n)$, we have the following moments:

$$E[\chi^r(n)] = \lambda A^r \tag{5}$$

and autocorrelation functions

$$E[\chi(n_1)\chi(n_2)\ldots\chi(n_i)] = \lambda^j A^i, \tag{6}$$

for j distinct time indices $(n_1, \ldots, n_i)$, due to the statistical independence of the values of the Poisson process in each time bin where $E[\cdot]$ denotes the expected value or ensemble-average. These statistical properties are critical for the development of the orthogonal functional series, which we term the Poisson–Wiener (P–W) series. The development of the P–W series is greatly simplified if we use the de-meaned input

$$z(n) = \chi(n) - \lambda A. \tag{7}$$

Then the P–W orthogonal functionals $\{Q_j\}$, which involve a new set of characteristic P–W kernels $\{p_j\}$ and are constructed by a Gram–Schmidt orthogonalization procedure, take the form

$$Q_0 = p_0, \tag{8}$$

$$Q_1[z(n); p_1] = \sum_{m=0}^{M} p_1(m) z(n-m), \tag{9}$$

$$Q_2[z(n); p_2] = \sum_{m_1=0}^{M} \sum_{m_2=0}^{M} p_2(m_1, m_2) z(n-m_1) z(n-m_2) - \frac{\mu_3}{\mu_2} \sum_{m=0}^{M} p_2(m,m) z(n-m) - \mu_2 \sum_{m=0}^{M} p_2(m,m), \quad \text{etc.} \tag{10}$$

where the orthogonality of the functionals is defined by $E[Q_i(n)Q_j(n)] = 0$, for $i \neq j$, and

$$y(n) = \sum_{i=0}^{\infty} Q_i[z(n); p_i]. \tag{11}$$

Note that these orthogonal functionals depend on the statistical central moments,

$$\mu_1 \triangleq E[z(n)] = 0, \quad \mu_2 \triangleq E[z^2(n)] = \lambda(1-\lambda)A^2, \quad \mu_3 \triangleq E[z^3(n)] = \lambda(1-\lambda)(1-2\lambda)A^3, \text{ etc.}$$

Most importantly, we note the following key relation:

$$\mu_4 \triangleq E[z^4(n)] = \mu_2^2 + \frac{\mu_3^2}{\mu_2}, \tag{12}$$

that attains critical importance in the sequel.

By tracing Wiener's steps in the continuous case with Gaussian white-noise input, we may obtain the unknown P–W kernels $\{p_i\}$ by evaluating the co-variances between the output $y(n)$ and known orthogonal 'instrumental' P–W functionals of the input $z(n)$ – i.e., evaluating the 'orthogonal projections' of the output signal upon each of these 'instrumental' orthogonal functionals that form an orthogonal 'coordinate system' in the functional space of the system. If these 'instrumental' functionals are chosen to be simple shift-operators, then this approach is an adaptation of the cross-correlation technique [6] used extensively for continuous-input systems, which was first proposed by Krausz [4] for point-process input systems in a continuous-time formulation and studied in discrete-time by Scaringe [16].

The present formulation yields the properly weighted cross-correlation formulae and shows that the diagonal values of the P–W kernels must be defined as zero for the P–W series – a result that was correctly but heuristically derived by Krausz. The latter is proven for the second-order kernel by the identity of Eq. (12), as shown below.

For the zeroth-order kernel (term), we must evaluate the expected value of the output signal, i.e., the co-variance between the output and a unity constant (zeroth-order instrumental functional),

$$p_0 = E[y(n)], \tag{13}$$

i.e., p_0 represents the average output value or the 'orthogonal projection' of the output on a unity constant.

Likewise, for the evaluation of the first-order kernel, we consider the known first-order 'instrumental functional': $I_1(n;m) = z(n-m)$, which has the structure of a first-order P–W functional with an impulsive kernel $\delta(\ell - m)$ (i.e., a Kronecker delta located at m), and is orthogonal for each m to all functionals $Q_i(n)$ for $i \neq 1$. Thus

$$E[y(n)z(n-m)] = E[Q_1(n)z(n-m)] = \mu_2 \cdot p_1(m), \tag{14}$$

since $E[z(n-k)z(n-m)] = \mu_2\delta(k-m)$, and therefore

$$p_1(m) = \frac{1}{\mu_2}E[y(n)z(n-m)]. \tag{15}$$

For the evaluation of the second-order kernel, we consider the 'instrumental functional'

$$I_2(n;m_1,m_2) = \left[z(n-m_1)z(n-m_2) - \frac{\mu_3}{\mu_2}z(n-m_1)\delta(m_1-m_2) - \mu_2\delta(m_1-m_2)\right],$$

which has the structure of a second-order P–W functional with a kernel $\delta(\ell_1 - m_1)\delta(\ell_2 - m_2)$. The co-variance between $y(n)$ and $I_2(n;m_1,m_2)$ eliminates all Q_i functionals for $i \neq 2$ and yields (after considerable analytical manipulations),

$$\begin{aligned} E[y(n)I_2(n;m_1,m_2)] &= E[Q_2(n)I_2(n;m_1,m_2)] \\ &= 2\mu_2^2 p_2(m_1,m_2) + \delta(m_1-m_2)\left(\mu_4 - 3\mu_2^2 - \frac{\mu_3^2}{\mu_2}\right)p_2(m_1,m_2). \end{aligned} \tag{16}$$

The key (and surprising) realization is that, because of the identity of Eq. (12) relating μ_2, μ_3, and μ_4 for any Poisson process, the co-variance in Eq. (16) for $m_1 = m_2$ is zero! In other words, the 'orthogonal projection' of the output signal $y(n)$ upon the signal: $[z^2(n-m) - (1-2\lambda)A\cdot z(n-m) - \lambda(1-\lambda)A^2]$, is zero for all m (i.e., they are orthogonal). Therefore, the diagonal values of the second-order P–W kernel $p_2(m,m)$ must be defined as zero. Thus

$$p_2(m_1,m_2) = \begin{cases} \dfrac{1}{2\mu_2^2}E[y(n)z(n-m_1)z(n-m_2)], & \text{for } m_1 \neq m_2, \\ 0, & \text{for } m_1 = m_2. \end{cases} \tag{17}$$

In general, the rth-order P–W kernel is obtained by the cross-correlation formula

$$p_r(m_1,\ldots,m_r) = \begin{cases} \dfrac{1}{r!\mu_2^r}E[y(n)z(n-m_1)\ldots z(n-m_r)], & \text{for distinct } (m_1,\ldots,m_r), \\ 0, & \text{otherwise,} \end{cases} \tag{18}$$

where $\mu_2 = \lambda(1-\lambda)A^2$, is the second moment of $z(n)$.

The key definition that the diagonal values of the P–W kernels be zero leads to considerable simplification of the form of the P–W functionals – e.g., eliminating two terms of Q_2 in Eq. (10) since $p_2(m,m) \equiv 0$. Thus, the P–W series takes the simpler form

$$\begin{aligned} y(n) = p_0 &+ \sum_m p_1(m)z(n-m) + \sum_{m_1}\sum_{m_2} p_2(m_1,m_2)z(n-m_1)z(n-m_2) \\ &+ \sum_{m_1}\sum_{m_2}\sum_{m_3} p_3(m_1,m_2,m_3)z(n-m_1)z(n-m_2)z(n-m_3) + \cdots, \end{aligned} \tag{19}$$

identical to the Volterra series form for the de-meaned input $z(n)$, with the important distinction that the diagonal values of the P–W kernels be zero by definition. This result was anticipated by Krausz [4], although his estimation formulae are different than the ones provided below in Eq. (20).

For ergodic and stationary processes, the ensemble averaging can be replaced by time-averaging over infinite data-records. Since, in practice, we only have the benefit of finite data-records, the aforementioned time-averaging is limited to a finite domain of time and results in 'estimates' of the precise kernels. The accuracy of these estimates is, of course, a matter of critical practical concern; especially as it relates to key experimental parameters such as data record length, bandwidth, etc. The P–W kernel estimation formulae can be adapted to the specific input form of Eq. (1) by utilizing the properties of the Kronecker delta

$$\begin{aligned}
\hat{p}_0 &= \frac{1}{N}\sum_{n=1}^{N} y(n) = \frac{KA}{N} = \lambda A,\\
\hat{p}_1(m) &= \frac{A}{N\mu_2}\sum_{i=1}^{K} y(n_i+m) - \frac{\lambda A}{\mu_2}\hat{p}_0,\\
\hat{p}_2(m_1,m_2) &= \begin{cases} \dfrac{A^2}{2N\mu_2^2}\left\{\displaystyle\sum_{i_1=1}^{K}\sum_{i_2=1}^{K}\frac{1}{2}[y(n_{i_1}+m_1)+y(n_{i_2}+m_2)]\right. & \\ \quad\times\delta[(n_{i_1}-n_{i_2})-(m_2-m_1)] - \lambda\displaystyle\sum_{i=1}^{K}[y(n_i+m_1) & \\ \left.\quad + y(n_i+m_2)] + \lambda^2 N\hat{p}_0\right\}, & \text{for } m_1 \neq m_2,\\ 0, & \text{for } m_1 = m_2, \end{cases}\\
&\text{etc.}
\end{aligned} \tag{20}$$

reducing the multiplication operations of cross-correlation to additions. Note that the parameter λ used in the expressions above is estimated in practice as, $\lambda = K/N$.

The P–W series of Eq. (19) can be expressed in terms of the original point-process input $\chi(n)$ by use of Eq. (7). The resulting expression for the system output in terms of the P–W kernels and the original input can be used to derive the Volterra kernel approximation

$$\hat{k}_0 = \sum_{r=0}^{\infty}(-\lambda A)^r \sum_{m_1}\cdots\sum_{m_r} p_r(m_1,\ldots,m_r), \tag{21}$$

$$\hat{k}_1(m) = \sum_{r=1}^{\infty} r(-\lambda A)^{r-1} \sum_{l_1}\cdots\sum_{l_{r-1}} p_r(m,l_1,l_2,\ldots,l_{r-1}), \tag{22}$$

$$\hat{k}_2(m_1,m_2) = \sum_{r=2}^{\infty}\frac{r!}{2!(r-2)!}(-\lambda A)^{r-2} \sum_{l_1}\cdots\sum_{l_{r-2}} p_r(m_1,m_2,l_1,\ldots,l_{r-2}) \tag{23}$$

etc.

Evidently, the diagonal values of these Volterra kernel approximations are also zero for systems with Poisson point-process inputs, reflecting the fact that a point-process input (having fixed spike

values) is intrinsically unable to probe (and therefore estimate) the kernel diagonal values. The kernels of Eqs. (21)–(23) are not identical to the Volterra kernels of the system, even away from the diagonals where they are defined to be zero. We will refer to these kernels as the Poisson–Volterra (P–V) kernels in order to make the important distinction with both the Volterra and the P–W kernels. This distinction applies also to the corresponding functional series. For instance, the first-order P–V kernel incorporates the main diagonals of all higher order Volterra kernels, in addition to the first-order Volterra kernel, i.e.,

$$\hat{k}_1(m) = k_1(m) + Ak_2(m,m) + A^2k_3(m,m,m) + \cdots. \tag{24}$$

Likewise, the second-order P–V kernel incorporates the secondary diagonals of all higher order Volterra kernels, i.e.,

$$\hat{k}_2(m_1,m_2) = \{k_2(m_1,m_2) + 3Ak_3(m_1,m_2,m_2) + \cdots\}[1-\delta(m_1-m_2)]. \tag{25}$$

The resulting P–V series attains the following meaning: the zeroth-order term is the average value of the output; the first-order term accounts for the responses to individual input impulses; the second-order term accounts for interactions between pairs of input impulses; the third-order term accounts for interactions among triplets of input impulses, etc. One could not hope for a more orderly and elegant mathematical model of the hierarchy of nonlinear interactions of point-process inputs.

We must emphasize that, in actual applications, the obtained model should be ideally put in the Volterra form, since the Volterra kernels are independent of the specific parameters (λ, A) of the Poisson input – unlike the estimated P–W kernels which depend on λ and A, or the P–V kernels which depend on A. However, the Volterra kernels cannot be obtained in practical applications, unless we can vary the impulse strength A experimentally in order to probe the diagonal values of the high-order kernels. Thus, in many practical applications, we may only be able to obtain the P–V kernels (directly or indirectly from the P–W kernels) which yield a system characterization (model) that does not depend on the specific parameter λ of the Poisson point-process input. Note that for a finite-order system, the highest order P–V and P–W kernels are identical.

Eqs. (21)–(23) can be used to reconstruct the P–V kernels of a system using a complete set of P–W kernel measurements. They also suggest a practical means for estimating the nonlinear order of the required model by varying the input-specific parameter (λ) and observing the resulting effects on the obtained P–W kernels. Since the latter are power series (or polynomial) expressions of (λ) with coefficients dependent on the respective P–V kernels, an indication of the significant high-order terms can be obtained (recall that the P–V kernels are independent of λ).

Having established the proper definitions and meaning for the P–V and P–W kernels, we now turn to the key issue in actual applications: the accurate and efficient estimation of P–W and P–V kernels from input–output data.

3. Kernel estimation

The cross-correlation technique (CCT) has been the first to be used for kernel estimation in actual applications [2,6,4,5,7,17–19]. Although it has been used extensively, it has been shown

to require long data-records for reasonable accuracy of the kernel estimates – a serious burden in actual experimental studies where the preparation can be kept stable only for limited time. A far better estimation method is proposed here, that is based on Laguerre expansions of the kernels and least-squares fitting procedures. This method was originally developed and applied to systems with continuous inputs, where it was shown to yield far superior kernel estimates (relative to the CCT) even for shorter data records [12].

It is critical to note that the CCT yields the P–W kernel estimates, however the Laguerre expansion techniques may yield either the Volterra or the P–V kernel estimates depending on whether the Laguerre expansion includes or not the diagonal values, respectively. In the former case, the kernel estimates at the diagonal points are interpolations (using the respective set of Laguerre functions) between the estimates at the nearest off-diagonal points.

This Laguerre expansion technique (LET) employs the orthonormal basis of Laguerre functions $\{b_j(m)\}$ to expand the system kernels, and then uses least-squares fitting to estimate the requisite expansion coefficients. Thus, we may consider the Laguerre expansions of the Volterra kernels

$$k_1(m) = \sum_{j=0}^{L} c_1(j) b_j(m), \tag{26}$$

$$k_2(m_1, m_2) = \sum_{j_1=0}^{L} \sum_{j_2=0}^{L} c_2(j_1, j_2) b_{j_1}(m_1) b_{j_2}(m_2) \tag{27}$$

etc.

Substitution of these kernel expansions into the Volterra series of Eq. (3) results in the modified Volterra model

$$y(n) = k_0 + \sum_{j=0}^{L} c_1(j) v_j(n) + \sum_{j_1=0}^{L} \sum_{j_2=0}^{L} c_2(j_1, j_2) v_{j_1}(n) v_{j_2}(n) + \cdots, \tag{28}$$

where

$$v_j(n) = \sum_{m=0}^{M} b_j(m) \chi(n-m). \tag{29}$$

Note that, if we seek the estimation of the P–V kernels, then the kernel expansion should employ not the regular Laguerre functions previously used for Volterra kernel expansion, but the 'associated Laguerre functions' of first order which allow the representation of the zero-valued diagonals of the high-order P–V kernels. This is the main innovation with respect to the use of LET for P–V kernel estimation.

The expansion coefficients c_1, c_2, etc. of Eq. (28) are estimated by least-squares fitting of the input–output data (note that the signals $\{v_j(n)\}$ are known since they are the convolutions of the input $\chi(n)$ with the Laguerre functions). The efficacy of this approach relies on the fact that many real system kernels can be represented adequately with a relatively small number of Laguerre functions, i.e., $L \ll M$ [12]. An illustrative example is provided below, using computer-simulated data.

4. Illustrative example

To demonstrate the efficacy of LET and its superior performance relative to CCT, we consider a second-order system defined by the first-order and second-order Volterra kernels shown in Fig. 1(a) ($M = 50$). This system is simulated for a Poisson process input with $\lambda = 0.1$ over $N = 2048$ data-points of a training dataset ($A = 1$, $T = 1$).The first-order and second-order P–W kernel estimates using CCT on this training dataset are shown in Fig. 1(b) and demonstrate the fact that the CCT estimates are rather poor. Note that the LET estimates (for $L = 10$) are identical to the actual P–V kernels in this case!

A segment of the input–output data that were used for kernel estimation (termed 'the training dataset') and the corresponding P–W model prediction using the CCT kernel estimates of Fig. 1(b) are shown in Fig. 2, and two segments for a different input–output dataset (termed 'the testing dataset') is shown in Fig. 3. As expected, the CCT model prediction is rather poor in both cases (relative to the LET model prediction which is identical to the actual output data) and becomes poorer for the testing dataset (out-of-sample prediction).

Of particular interest for point-process input systems is the fact (discussed in Section 2) that the diagonal values of the second-order Volterra kernel cannot be estimated (the P–W kernel diagonal values are zero by definition) and that the first-order P–W kernel is distinct from its P–V counterpart according to the expressions of Eqs. (21)–(23). For this example (see Eq. (22))

$$\hat{k}_1(m) = p_1(m) - 2\lambda A \sum_{m'=0}^{M} p_2(m, m') \tag{30}$$

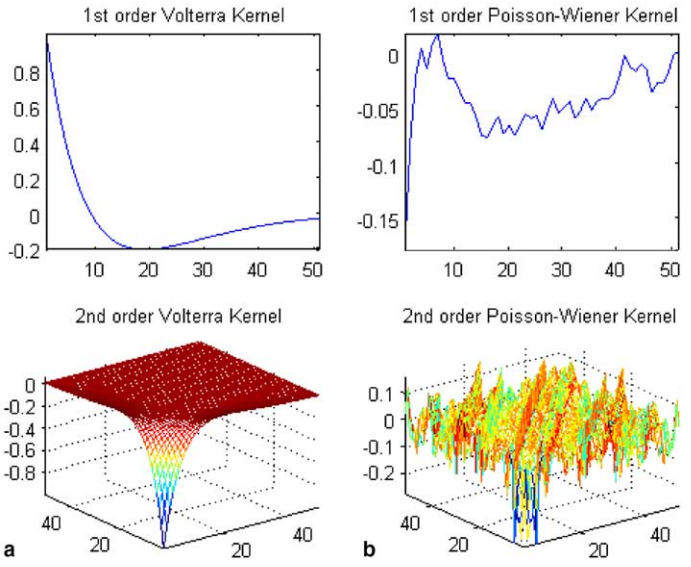

Fig. 1. (a) The first-order (top) and second-order (bottom) Volterra kernels of the simulated system (left column); and (b) the P–W kernel estimates obtained via CCT for an input–output dataset of 2048 data-points (right column). The poor quality of the P–W kernel estimates is evident, as well as the 'diagonal problem' (see text). The Volterra kernel estimates using LET are identical to the exact kernels shown in (a).

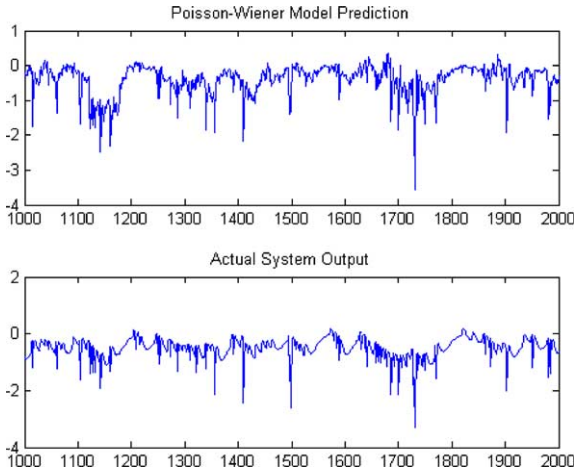

Fig. 2. Segments of the P–W model prediction using the CCT kernel estimates of Fig. 1(b) and the corresponding output data for the training dataset. Note that the model prediction for the LET kernel estimates is identical to the output data.

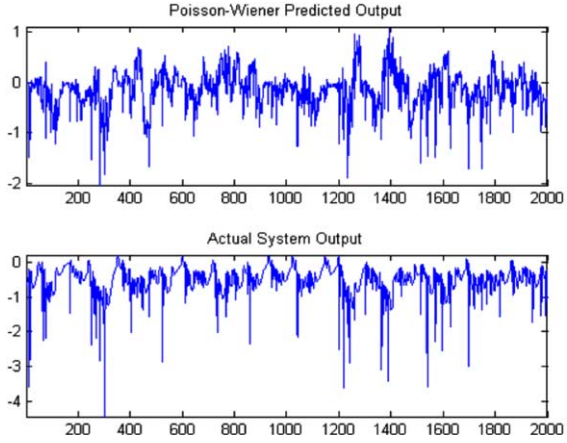

Fig. 3. Segments of the P–W model prediction using the CCT kernel estimates of Fig. 1(b) and the corresponding output data for the testing dataset (out-of-sample prediction). Note that the model prediction for the LET kernel estimates is again identical to the output data.

and

$$p_1(m) = \hat{k}_1(m) + 2\lambda A \sum_{m'=0}^{M} \hat{k}_2(m, m'). \tag{31}$$

Since $p_2 \equiv \hat{k}_2$ for all second-order systems. Note also that for a second-order system

$$\hat{k}_2(m, m') = [1 - \delta(m - m')]k_2(m, m'). \tag{32}$$

LET yields precise estimates of the P–W kernels or the P–V kernels (depending on whether we use the de-meaned or the actual Poisson input, respectively) using a short input–output data-record, while CCT requires much longer data-records to approach comparable performance in the estimation of the P–W kernels (P–V kernels cannot be directly estimated via CCT). This is a critical advantage in practice. Another important practical advantage is the fact that LET is rather robust in the presence of noise in the output data [12].

As indicated above, the P–V kernels away from the diagonals can be reconstructed from a complete set of P–W kernels. However, the diagonal values cannot be directly estimated by either of these methods; a limitation imposed by the point-process nature of the input. The estimation of the diagonal second-order kernel values using LET relies on interpolation based on the structure of the Laguerre functions used for the kernel expansion.

Eqs. (24) and (25) indicate that, if the experiment can be repeated for various values of A, then the diagonal kernel values of finite-order system can be recovered. Note that the estimation of the diagonal values of the second-order kernels (or higher) is also possible if the input impulses have variable amplitude. For instance, if the Poisson point-process input is amplitude-modulated by a Gaussian white noise process, then data collected with this modified input can be analyzed using the original LET algorithm to yield precise estimates of the Volterra kernels, including the diagonal values.

5. Conclusions

This paper clarifies the distinctions between Volterra and Wiener kernels obtained from Poisson point-process inputs, termed Poisson–Volterra (P–V) and Poisson–Wiener (P–W) kernels, respectively. It shows that the diagonal values of high-order Volterra or Wiener kernels cannot be directly estimated using Poisson point-process inputs and must be defined as zero in the P–W and P–V kernels. Only the P–W kernels can be estimated via cross-correlation. The P–V kernels can be reconstructed from a complete set of P–W kernel estimates (see Eqs. (21)–(23)). The latter can be estimated far more accurately (and from shorter data-records) by use of the Laguerre expansion technique (LET) adapted to point-process inputs, rather than the conventional cross-correlation technique. Since the P–V kernels are independent of the mean rate of stimulation (i.e., the Poisson parameter λ), they constitute the preferred set of kernels in practice. Their direct estimation is possible via LET when the input Poisson process is not de-meaned. It is also possible to estimate the Volterra kernels of the system via LET, when the diagonal values of the high-order kernels can be interpolated by the Laguerre set of basis functions.

An important distinction exists between the Volterra and the P–V kernels away from the diagonal values, as well (recall that the P–V kernels are defined as zero along the diagonals). Eqs. (24) and (25) describe this distinction for first-order and second-order kernels. It is a manifestation of the fact that the diagonal values of high-order kernels 'fold' onto lower order kernels for point-process inputs. For instance, the main diagonals of all high-order kernels fold onto the first-order kernel, as described by Eq. (24). Likewise, the secondary diagonals of all kernels of order higher than second fold onto the second-order kernel, as described by Eq. (25), and so on. This is a direct consequence of the impulsive nature of a point-process input. The diagonal values of the kernels can be estimated directly only if the amplitude of the impulses is variable (e.g., random

modulation) or, indirectly, by repeating the experiment for various values of A and combining the results according to Eqs. (24) and (25) for finite-order systems.

Changes in the Poisson parameter λ (i.e., mean rate of stimulation) lead to different P–W kernel estimates (see Eqs. (21)–(23)) but do not alter the P–V kernel estimates which, however, depend on the amplitude A of the input impulses (see Eqs. (24) and (25)). Note that the P–V kernels can also be estimated via LET for non-Poisson point-process inputs, so long as the latter do not have a strict deterministic structure (e.g., periodic) and retain broadband characteristics (e.g., natural or spontaneous operation of the system).

It is evident from this analysis that it is preferable in practice to estimate directly the P–V or the Volterra kernels using LET for any Poisson input parameter λ. Most importantly, kernel estimation accuracy is much higher using LET rather than cross-correlation. Since the P–V kernels are independent of λ but depend on A, repeated experiments with different (but fixed) values of A can yield insight into the diagonal kernel values based on Eqs. (24) and (25).

Acknowledgment

This work was supported by Grant No. 01861 of the National Institutes of Health (NIBIB) awarded to the Biomedical Simulations Resource at the University of Southern California.

References

[1] T.W. Berger, T.P. Harty, G. Barrionuevo, R.J. Sclabassi, Modeling of neuronal networks through experimental decomposition, in: V.Z. Marmarelis (Ed.), Advanced Methods of Physiological Systems Modeling, vol. 2, Plenum, New York, 1989, p. 113.
[2] T.W. Berger, G. Barrionuevo, S.P. Levitan, D.N. Krieger, R.J. Sclabassi, Nonlinear systems analysis of network properties of the hippocampal formation, in: M. Gabriel, J. Moore (Eds.), Foundations of Adaptive Networks, MIT, Massachusetts, 1990, p. 283.
[3] T.W. Berger, T.P. Harty, C. Choi, X. Xie, G. Barrionuevo, R.J. Sclabassi, Experimental basis for an input/output model of the hippocampal formation, in: V.Z. Marmarelis (Ed.), Advanced Methods of Physiological System Modeling, vol. 3, Plenum, New York, 1994, p. 29.
[4] H.I. Krausz, Identification of nonlinear systems using random impulse train inputs, Biol. Cybern. 19 (1975) 217.
[5] J.P. Kroeker, Wiener analysis of nonlinear systems using Poisson–Charlier crosscorrelation, Biol. Cybern. 27 (1977) 221.
[6] Y.W. Lee, M. Schetzen, Measurement of the Wiener kernels of a nonlinear system by crosscorrelation, Int. J. Control 2 (1965) 237.
[7] P.Z. Marmarelis, V.Z. Marmarelis, Analysis of Physiological Systems: The White-Noise Approach, Plenum, New York, 1978, Russian translation: Mir, Moscow, 1981. Chinese translation: Academy of Sciences, Beijing, 1990.
[8] V.Z. Marmarelis (Ed.), Biomedical Simulations Resources, Advanced Methods of Physiological System Modeling, vol. 1, University of Southern California, Los Angeles, California, 1987.
[9] V.Z. Marmarelis (Ed.), Advanced Methods of Physiological System Modeling, vol. II, Plenum, New York, 1989.
[10] V.Z. Marmarelis (Ed.), Advanced Methods of Physiological System Modeling, vol. III, Plenum, New York, 1994.
[11] V.Z. Marmarelis, Nonlinear Dynamic Modeling of Physiological Systems, Wiley Interscience & IEEE, New York, 2004.
[12] V.Z. Marmarelis, Identification of nonlinear biological systems using Laguerre expansions of kernels, Ann. Biomed. Eng. 21 (1993) 573.

[13] V.Z. Marmarelis, Modeling methodology for nonlinear physiological systems, Ann. Biomed. Eng. 25 (1997) 239.
[14] V.Z. Marmarelis, M.E. Orme, Modeling of neural systems by use of neuronal modes, IEEE Trans. Biomed. Eng. 40 (1993) 1149.
[15] H. Ogura, Orthogonal functionals of the Poisson process, IEEE Trans. Inform. Theory IT 18 (1972) 473.
[16] W.A. Scaringe, Cross-correlation methods for quantification of nonlinear input–output transformations of neural systems using a Poisson random test input. Thesis for Master of Science, University of Southern California, Los Angeles, California, 1995.
[17] R.J. Sclabassi, D.N. Krieger, T.W. Berger, Nonlinear systems analysis of the somatosensory system, in: V.Z. Marmarelis (Ed.), Biomedical Simulations Resource, Advanced Methods of Physiological System Modeling, vol. I, University of Southern California, Los Angeles, California, 1987, p. 104.
[18] R.J. Sclabassi, D.N. Krieger, T.W. Berger, A systems theoretic approach to the study of CNS function, Ann. Biomed. Eng. 16 (1988) 17.
[19] R.J. Sclabassi, B.R. Kosanovic, G. Barrionuevo, T.W. Berger, Computational methods of neuronal network decomposition, in: V.Z. Marmarelis (Ed.), Advanced Methods of Physiological System Modeling, vol. III, Plenum, New York, 1994, p. 55.
[20] V. Volterra, Theory of Functions and of Integral and Integro-differential Equations, Dover Publications, New York, 1930.
[21] N. Wiener, Nonlinear Problems in Random Theory, The Technology Press of MIT and Wiley, New York, 1958.