A Volterra filter for neuronal spike detection

Mamadou MBOUP

Université Paris Descartes
UFR de Mathématiques et Informatique - CRIP5
& EPI ALIEN INRIA
45, rue des Saints-Pères - 75270 Paris cedex 06
e-mail : Mamadou.Mboup@mi.parisdescartes.fr
Phone: +33 1 44 55 35 25, Fax: +33 1 44 55 35 35

Abstract

The spike detection problem is cast into a delay estimation. Using elementary operational calculus, we obtain an explicit characterization of the spike locations, in terms of short time window iterated integrals of the noisy signal. From this characterization, we derive a joint spike detection and localization system where the decision function is implemented as the output of a digital Volterra filter. Simulation results using experimental data shows that the method compares favorably with one of the most successful one in the litterature.

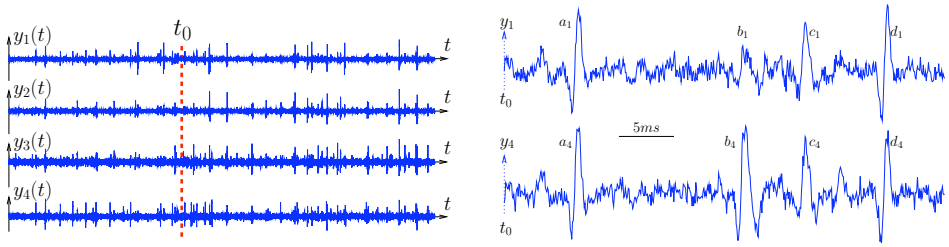
Keywords: Neuronal spike detection, operational calculus, detection and estimation, numerical integration, Volterra filter.

1 Introduction

The equilibrum of electrical charges across the axonal membrane of a neuron induces a difference of potentials known as the resting potential. When this equilibrum is broken, either in response to an extracellular stimulus or even spontaneously, an electrical discharge, called *action potential* (AP) is emitted. It is a known fact that neurons use these electrical discharges *viz* spike trains to communicate between them [1].

1.1 Signal description and problem formulation

Figure 1(a) shows 1s data delivered by four recording[1] electrodes placed in the nervous olfactory system of a locust [2]. The data are bandpass filtered between 300 Hz and 5 kHz. More details on these records are shown in fig-



(a) One second records of AP, bandpass filtered in [300Hz, 5kHz]

(b) Records of AP: zoom on a segment of $40ms$ duration

Figure 1: Four electrodes recording

ure 1(b), representing $y_1(t)$ and $y_4(t)$ zoomed in the interval $[t_0, t_0 + 40ms]$. Clearly, the signal from each recording site may be represented in the form

$$y(t) = \sum_n \sum_k a_{n,k} p_n(t - t_{n,k}) + \varpi(t), \tag{1}$$

where the delays $t_{n,k}, k = 1, 2, \ldots$ are the firing instants of neuron number n, and $p_n(t)$ is the corresponding AP waveform. A comparison of the data from different sites shows that the amplitudes $a_{n,k}$ of the recorded action potentials vary markedly with the distance of the neuron n from the recording site (compare b_1 and b_4 in figure 1(b)). In general, such attenuations also vary in time. Therefore, it is likely that much of the background additive noise $\varpi(t)$ is composed of action potentials from remote neurons [3]. The zoom in figure 1(b) also exhibits different AP waveforms, corresponding to different neurons [4], [5] (compare a_1–a_4, b_4 and d_1–d_4).

One of the most important and challenging problem in neuroscience is the decoding of the neural information conveyed by the spike trains [1], [6]. Reliable detection of the AP and accurate estimation of the spiking instants $t_{n,k}$ constitute the first mandatory step towards such neural information processing. Distinguishing which neuron fires at a given spiking instant is another prerequisite (see [4], [7] and also [8] for a tutorial). However,

[1] This experiment was realized by Ch. Pouza (see `www.biomedicale.univ-paris5.fr/physcerv/physiologie_cerebrale.htm`).

only the detection and estimation problem is investigated here. Note that the recordings are not always as clean as in the presented experiment. The problem becomes difficult in such situations (low signal to noise ratio).

1.2 Local piecewise polynomial model

Since we do not require to distinguish between the spiking activities of the neurons, the model in equation (1) may be simplified in the form

$$y(t) = \sum_{i \geqslant 0} H(t - t_i)\, f_i(t - t_i) + \varpi(t), \tag{2}$$

where each $f_i(t)$ is a smooth segment. A change from f_{i-1} to f_i represents the occurrence of a transient (a spike), at time t_i. Each segment f_i is multiplied by the unit step (Heaviside) $H(\cdot)$, for causality purpose. We set $t_0 = 0$. Based on the observation $y(t)$, we want to detect the transient phenomena and estimate their locations t_i. This is a common description of the so-called *change-point* detection and estimation problem [9], [10]. It is casted here into the estimation of the delays t_i, with a local piecewise polynomial signal representation.

From now on, $x(t)$ will denote the unobserved noise-free signal: $x(t) = y(t) - \varpi(t)$. Let $T > 0$ be given. To any $\tau \geqslant 0$, we associate the interval $I_\tau^T = [\tau, \tau + T)$ and we assume that there is at most one AP in each such interval. In the sequel, we will define

$$x_\tau(t) \stackrel{\triangle}{=} H(t)x(t + \tau), \quad t \in [0, T), \tag{3}$$

for the restriction of the signal x in I_τ^T and we redefine the possible spiking instant, say t_τ, relatively to I_τ^T with: $t_\tau = 0$ if $x_\tau(t)$ is smooth and $0 < t_\tau < T$ otherwise. The presence of an AP is thus interpreted as a non smoothness of $x_\tau(t)$. The N^{th} order (generalized) derivative [11] of $x_\tau(t)$ then satisfies

$$\frac{d^N x_\tau}{dt^N} = [x_\tau^{(N)}](t) + \sum_{k=0}^{N-1} \left(\mu_k^0\, \delta(t)^{(k)} + \mu_k\, \delta(t - t_\tau)^{(k)} \right) \tag{4}$$

where the superscript (k) denotes order k differentiation and where $[f]$ stands for the regular part of f. We have set $\mu_{N-k-1}^0 \stackrel{\triangle}{=} x_\tau^{(k)}(0+) - x_\tau^{(k)}(0-) = x_\tau^{(k)}(0+)$ and μ_{N-k-1} is defined similarly with the origin replaced by the point t_τ.

Equation (4) is the cornerstone of the approach we are presenting. This consists in a joint detection/estimation, based on the following observations:

1. I_τ^T is free of AP if, and only if, $\mu_0 = \cdots = \mu_{N-1} = 0$.
2. If one (and only one) AP is present in I_τ^T, then one can identify the spiking instant t_τ directly from equation (4).

A key step towards the estimation of t_τ is to express (4) in the operational domain (via the Laplace transform). Indeed, terms like $\delta(t-r)^{(k)}$ are much easier to handle using operational calculus [12]. To proceed, we henceforth assume a piecewise polynomial model for $x_\tau(t)$. However, we mention at once that the detection/estimation method presented in the next section does not rely on curve fitting: identifying the piecewise polynomials is not required.

2 Spike detection and estimation

First, we derive from (4) a linear estimator for t_τ. This will yield an explicit expression which degenerates when the corresponding interval I_τ^T is devoid of an AP. Testing the consistency of that expression will thus afford a spike detection procedure. The result of this detection process, in turn, will allow a precise location of the spiking instant, if any.

2.1 Change-point estimation

Several change-point estimators may be devised from equation (4), depending on the differentiation order N and on the assumed maximum degree of the segments of $x_\tau(t)$. The following estimator is based on a piecewise affine model. Such model is sufficient to roughly match the impulsive shape of an AP. Let $N = 2$ so that the regular component $[\ddot{x}_\tau(t)]$ vanishes. Once translated into the operational domain [12], (4) becomes

$$s^2\hat{x}_\tau(s) - sx_\tau(0) - \dot{x}_\tau(0) = (\mu_0 + \mu_1 s)e^{-t_\tau s}. \tag{5}$$

By elementary manipulations of (5) we will annihilate the unknown and irrelevant coefficients μ_i, and initial conditions $x_\tau^{(k)}(0)$. This will lead to a linear estimator for t_τ. Once translated back in the time domain, we will get an explicit expression depending on integral terms as $\int_0^T p(\lambda)x_\tau(\lambda)d\lambda$. The next algorithm describes these manipulations which stem from the estimation theory due to [13] (see also [14]).

Step 1 - First, we get rid of all irrelevant unknowns. For this, observe that the right member of (5), namely $\hat{u}(s) \stackrel{\triangle}{=} (\mu_1 + \mu_0 s)e^{-t_\tau s}$, satisfies the

differential equation $\hat{u}''(s) + 2t_\tau \hat{u}'(s) + t_\tau^2 \hat{u}(s) = 0$. To simplify the notations, we will write $\hat{x}_\tau$ in place of $\hat{x}_\tau(s)$ and likewise for $\hat{u}(s)$. Noting that $\hat{u}'' = \frac{d^2}{ds^2}(s^2\hat{x}_\tau)$, we have for $i \geqslant 2$,

$$\frac{d^{i+2}}{ds^{i+2}}(s^2\hat{x}_\tau) + 2t_r \frac{d^{i+1}}{ds^{i+1}}(s^2\hat{x}_\tau) + t_r^2 \frac{d^i}{ds^i}(s^2\hat{x}_\tau) = 0. \tag{6}$$

Step 2 - By the very classical rules of operational calculus, multiplication by s corresponds to time derivation: $s\hat{x}_\tau \bullet\!\!-\!\!\circ \frac{d}{dt}x_\tau(t)$. Now, numerical differentiation is difficult and ill-conditioned. Divide both members of (6) by a s^ν, for an integer $\nu > 2$. Only negative powers of s, and hence integral operators, will intervene in the resulting equation. Replace the unobserved signal x_τ by its noisy observation counterpart y_τ. Gathering the resulting equations for $i = \kappa + 2, \kappa \geqslant 0$ and $i + 1$, we obtain the linear system

$$\underbrace{\begin{bmatrix} \hat{v}_\kappa(s,\tau) & \hat{v}_{\kappa+1}(s,\tau) \\ \hat{v}_{\kappa+1}(s,\tau) & \hat{v}_{\kappa+2}(s,\tau) \end{bmatrix}}_{\widehat{P}_\kappa(s,\tau)} \underbrace{\begin{bmatrix} \widetilde{t_\tau^2} \\ \widetilde{2t_\tau} \end{bmatrix}}_{\widetilde{\boldsymbol{\Theta}}_\tau} = \underbrace{-\begin{bmatrix} \hat{v}_{\kappa+2}(s,\tau) \\ \hat{v}_{\kappa+3}(s,\tau) \end{bmatrix}}_{\widehat{Q}_\kappa(s,\tau)} \tag{7}$$

where $\hat{v}_\kappa(s,\tau) = \dfrac{1}{s^\nu}\dfrac{d^{\kappa+2}}{ds^{\kappa+2}}\left(s^2\hat{y}_\tau\right)$. The solution $\widetilde{\boldsymbol{\Theta}}_\tau$ of (7) is an estimator of $\boldsymbol{\Theta}_\tau = [t_\tau^2 \;\; 2t_\tau]^T$ because x_τ has been replaced by the noisy observation $\hat{y}_\tau$.

Step 3 - The numerical estimates are now obtain by expressing the system (7) back in the time domain. For this, let us first recall that differentiation with respect to s corresponds to multiplication by $-t$: $\frac{d}{ds}\hat{x}_\tau \bullet\!\!-\!\!\circ -tx_\tau(t)$, and division by s^m, $m > 0$ maps to the m^{th}-iterated time integration:

$$(m-1)!s^{-m}\hat{x} \bullet\!\!-\!\!\circ \int_0^t (t-\alpha)^{m-1}x(\alpha)d\alpha.$$

With these rules in mind and using integration by parts, we express the time domain analog of $\hat{v}_\kappa(s,\cdot)$ as:

$$v_\kappa(t,\tau) = \frac{(-1)^\kappa}{(\nu-1)!}\int_0^t \frac{d^2}{d\lambda^2}\left\{\lambda^{\kappa+2}(t-\lambda)^{\nu-1}\right\} y_\tau(\lambda)d\lambda. \tag{8}$$

The analog of (7) is thus in the form $P_\kappa(t,\tau)\widetilde{\boldsymbol{\Theta}}_\tau(t) = Q_\kappa(t,\tau)$ so that we have an estimate $\widetilde{\boldsymbol{\Theta}}_\tau(t)$ for each estimation time $t \leqslant T$. To be consistent with the signal model in (3), we now on fix this estimation time to the width of I_τ^T: $t = T$. We can therefore simplify the notations above by dropping the argument t, and write *e.g.* $v_\kappa(\tau)$ instead of $v_\kappa(T,\tau)$.

2.2 Joint detection-estimation

Recall that in equation (5) we have $\mu_0 = \mu_1 = 0$ when the interval I_τ^T is devoid of an action potential. In such situation, $t_\tau = 0$ and then the estimator $P_\kappa(\tau)\widetilde{\mathbf{\Theta}}_\tau = Q_\kappa(\tau)$ degenerates into $P_\kappa(\tau) = 0$ and $Q_\kappa(\tau) = 0$ (up to the output noise level). Otherwise, an AP is detected in I_τ^T. Thus, if an AP happens at time $t^\star$, then it will be detected in all intervals I_τ^T for which $t^\star - T \leqslant \tau \leqslant t^\star$. This observation shows how the detection process also leads to the estimation of the spiking instants.

It now remains to determine a decision function $\mathcal{J}$ along with an associated threshold $\gamma_\mathcal{J}$, and devise a detection system,

$$\mathcal{J}(\tau) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \gamma_\mathcal{J} \tag{9}$$

to test the degeneration of the system $P_\tau\widetilde{\mathbf{\Theta}}_\tau = Q_\tau$ (null hypothesis $\mathcal{H}_0$), against the alternative $\mathcal{H}_1$: *one AP occurs in* I_τ^T. In the following, we consider the family of functions $\mathcal{J}_\kappa$:

$$\mathcal{J}_\kappa(\tau) = [v_{\kappa+1}(\tau)]^2 - v_\kappa(\tau)v_{\kappa+2}(\tau) = -\det P_\kappa(\tau), \quad \kappa \geqslant 0. \tag{10}$$

This also coincides with the discriminant of the quadratic equation formed by the first line of the system $P_\kappa(\tau)\mathbf{\Theta}_\tau = Q_\kappa(\tau)$. The condition $\mathcal{J}_\kappa(\tau) \geqslant 0$ is thus necessary to have real solutions while $\mathcal{J}_\kappa(\tau) = -\det P_\tau$ must be different from zero. The noise and the local mismodelling apart, we then have $\mathcal{J}_\kappa(\tau) > 0$ if, and only if, an AP happens in I_τ^T. Due to the noise presence, the condition $\mathcal{J}(\tau) > 0$ must be replaced by $\mathcal{J}(\tau) > \gamma_\mathcal{J}$. As the statistical properties of the noise are unknown, the threshold $\gamma_\mathcal{J}$ will be determined by experience.

2.3 Discrete nonlinear filter

All the preceding developments rely on a continuous-time setting. Although this is in accordance with the true nature of the spike train, only discrete observation is assumed in the implementation step. In this connection, we show that each $v_\kappa(\cdot)$, $\kappa \geqslant 0$ may be implemented as the output of a discrete finite impulse response (FIR) filter, with the sample observation y as input. Consequently, for each $\kappa \geqslant 0$, the corresponding decision function $\mathcal{J}_{\kappa,n}$, $n \in \mathbb{N}$, will read as the ouput of a discrete quadratic filter.

To begin, let us define

$$h_\kappa(\mu) = \begin{cases} \frac{(-1)^{\kappa+1}}{(\nu-1)!}\frac{d^2}{d\mu^2}\left\{(1-\mu)^{\kappa+2}\mu^{\nu-1}\right\}, & 0 \leqslant \mu \leqslant T \\ 0 & \text{else.} \end{cases} \tag{11}$$

Consider (8) with the change of variable $\lambda = T - \mu$ and set $\tau' = \tau + T$. Then, we may directly rewrite v_κ in (8) as

$$v_\kappa(\tau) = \int_0^T h_\kappa(\mu) y_\tau(T-\mu) d\mu = \int_0^\infty h_\kappa(\mu) y(\tau' - \mu) d\mu. \tag{12}$$

We turn now to the discrete-time setting where only time samples of the observation, $y_m = y(t_m)$, are available. We assume a uniform sampling, $t_m = mT_s, m = 0, 1, \ldots$, with sampling period T_s. Let $T = MT_s$ and $\nu \geqslant 2$ be given and compute the finite sequence $\{h_{\kappa,m} \stackrel{\triangle}{=} h_\kappa(t_m)\}_{m=0}^M$. Then, by a numerical integration method with abscissas t_m and weights W_m, $m = 0, \cdots, M$, we obtain the approximation of (12):

$$v_\kappa(\tau) \approx v_{\kappa,n} = \sum_{m=0}^M W_m h_{\kappa,m} y_{n+M-m},$$

for $\tau = nT_s$ (and $\tau' = (n+M)T_s$). In all the sequel, we use the trapezoidal method for which $W_0 = W_M = 0.5$, and $W_m = 1, m = 2, \cdots, M-1$. As $\tau = \tau_n = nT_s$ varies (sliding windows I_n^M), $v_{\kappa,n}$ reads as the output of the FIR filter, with impulse response sequence $\{g_{\kappa,m} \stackrel{\triangle}{=} W_m h_{\kappa,m}\}_{m=0}^M$. Accordingly, we obtain for each $\kappa \geqslant 0$, a discrete approximation $\mathcal{J}_{\kappa,n} \approx \mathcal{J}_\kappa(nT_s)$ of (10) as the output of the discrete Volterra filter with input y_n,

$$\mathcal{J}_{\kappa,n} = [v_{\kappa+1,n}]^2 - v_{\kappa,n} v_{\kappa+2,n} = \sum_{m,\ell=0}^M \mathcal{K}_{\ell,m} y_{n-\ell} y_{n-m} \tag{13}$$

where $\mathcal{K}_{\ell,m} = g_{\kappa+1,m} g_{\kappa+1,\ell} - g_{\kappa,m} g_{\kappa+2,\ell}$. In the following, a detection will be validated if it meets the agreement of a given number, K, of elementary decision functions $\mathcal{J}_{\kappa,n}$. Therefore, the final decision function is given by the combination

$$\mathcal{J}_n = \prod_{\kappa=0}^{K-1} \widetilde{\mathcal{J}}_{\kappa,n}, \quad n = 0, 1, \ldots, \tag{14}$$

where $\widetilde{\mathcal{J}}_{\kappa,n} = \max(0, \mathcal{J}_{\kappa,n})$ for each κ and n. This makes the detection more robust to fake spikes. The performance of the proposed method are investigated below, using Monte Carlo simulation.

3 Simulation results

We compare our joint estimation and detection method with one of the most successful spike detection method: the wavelet detection method (WDM) presented in [5].

3.1 Simulated signal

The following simulations are based on experimental data from 20s recording at 15kHz (see §1.1). Some manipulations are however necessary before we may compute any performace measure (see [5]).

Spike templates An average of 256 spikes have been extracted from each of the four recordings and aligned into 5 clusters using Lloyd's algorithm [15]. The centroids S_i of the clusters are then used as spike templates. They all have the same duration of 3.33ms and are normalized: $\|S_i\|_\infty = 1$.

Noise template Once all the spikes have been extracted from the experimental data, the remaining background noise is used as a neural noise template. We form a single long (285000 samples) observation by concatenating the data from the four recording sites. As we already discuss, a significant part of the noise is composed of remote spikes.

Noisy spike train simulation Each run is based on a signal segment of $2/3$ s (10000 samples). For each run, the spike locations are first drawn from the successes of 10000 samples Bernoulli trials with a parameter p corresponding to a given mean firing rate FR. The experiment is repeated until a refractory period of 2ms is reached after each spike. The number of simulated spikes thus varies from one run to another. Next, the noise-free spike train is generated by assigning randomly the spike templates to the locations. The polarity of each spike template also is chosen randomly. Finally, a randomly selected segment of the neural noise template is added. The noise is scaled to have a variance σ, according to the desired signal to noise ratio defined, as in [5]:

$$\text{SNR} = \frac{\|S_i\|_\infty}{\sigma} = \frac{1}{\sigma}. \tag{15}$$

Figure 2 displays 0.5s of such a simulated signal. The spike locations are indicated by the vertical bars on the bottom.

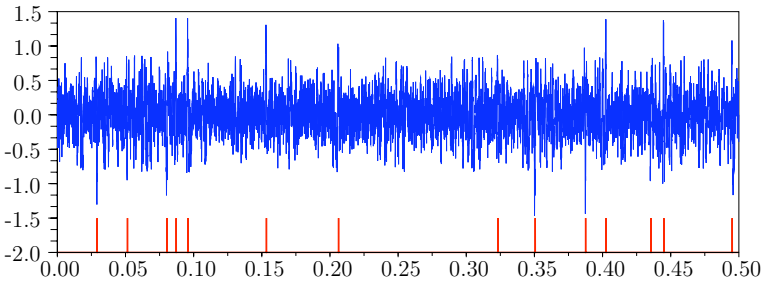


Figure 2: Noisy spike train simulation: SNR = 3.5, FR = 15 Hz.

3.2 Simulation settings and results

Throughout we set $\nu = 7$ for the order of iterated integrals and $T = 4ms$ for the estimation time. The parameters of WDM are chosen so as to have best performance for a number of scales set to 6.

Figure 3 below displays the receiver operating characteristic (ROC) curves: probability of correct detection (P_{CD}) against probability of false alarm (P_{FA}). These are computed by averaging results over 500 runs, for each

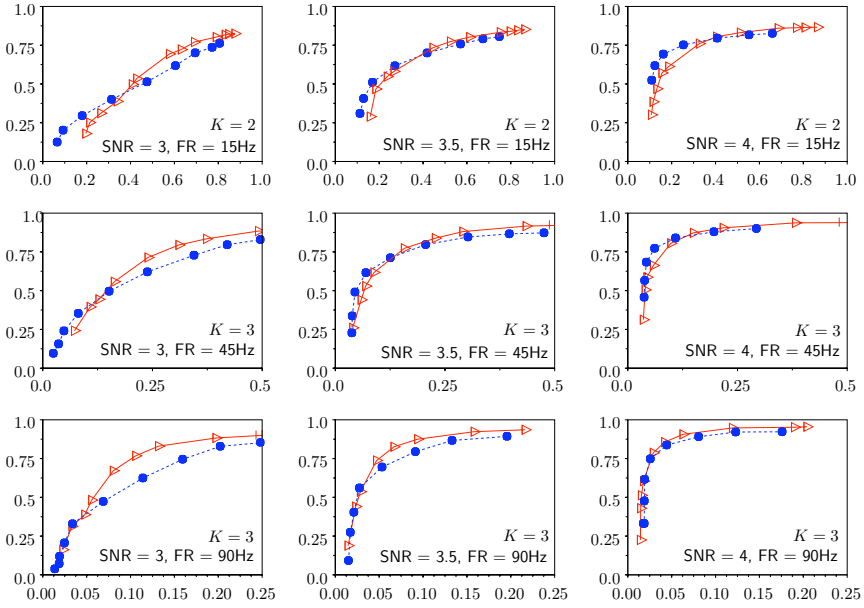


Figure 3: ROC curves: Presented method (▷ solid line) vs WDM (● dashed line).

pair (FR, SNR). Given a spike, correct detection is meant when the distance between the exact location and the estimated one is less than half the spike duration, *i.e.* $1.66ms$. As the threshold γ_J is expressed in percentage of $\{\mathcal{J}_n\}$, the performance of the method increase with the firing rate, like WDM. The ROC curves of the presented method ($\triangleright$ solid line) are almost always higher than that of the WDM ($\bullet$ dasehd line) for all pairs (FR, SNR). This is more apparent at high noise level (SNR = 3, first column in figure 3) for $\mathsf{P}_{\mathsf{CD}} > 0.5$. The presented method thus shows good robustness to noise perturbation. Moreover, it is easy to implement and it has low complexity: for a firing rate of FR $= 45Hz$, the average time per run was $0.48s$ against $1.52s$ for the WDM, *i.e.* more than three times.

The threshold $\gamma_{\mathcal{J}}$ is now determined as for WDM, by following [5]. It is fixed and we set $K = 4$. The top graph of figure 4 displays P_{CD} against P_{FA} for SNR $= 3.5, 3.6, \cdots, 4$ and for FR $= 15, 30$ and 45 Hz (solid, dotted and dashed line respectively). Although the two methods show quite different behaviors, it is possible to compare their performance by considering the ratio $\mathsf{P}_{\mathsf{FA}}/\mathsf{P}_{\mathsf{CD}}$ as a function of the SNR.

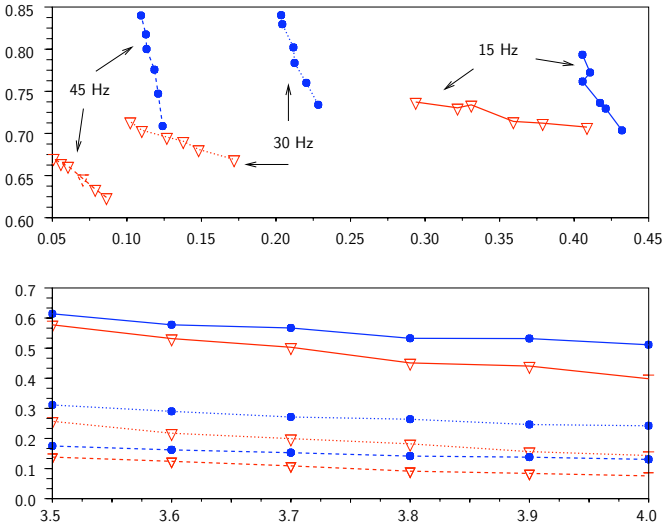


Figure 4: Comparison of the presented method ($\triangledown$ mark) *vs* WDM ($\bullet$ mark).

This is displayed in the bottom graph, for FR $= 15, 30$ and 45 Hz. Not only this ratio is always smaller for our method but also, it decreases faster when the SNR increases. Finally, let us mention that typical values of K are in the range 2..4, with very little differences for $K = 3$ and $K = 4$.

4 Conclusion

We have presented a joint spike detection and estimation method based on the operational calculus, a very classical tool in electrical engineering. The method is simple and easy to implement as a discrete Volterra filter. It compares favorably with one of the most successful spike detection method, based on continuous wavelet decomposition.

References

[1] F. Rieke, D. Warland, R. de Ruyter van Steveninck, and W. Bialek, *Spikes: Exploring the neural code*, MIT Press, 1999.

[2] Christophe Pouzat, *Methods and Models in Neurophysics. Les Houches 2003 Summer School.*, pp. 729–786, Elsevier, 2005, Available from: `http://fr.arxiv.org/abs/q-bio.QM/0405012`.

[3] W. Gerstner and W. Kistler, *Spiking Neuron Models. Single Neurons, Populations, Plasticity*, Cambridge University Press, 2002.

[4] Matthieu Delescluse and Christophe Pouzat, "Efficient spike-sorting of multi-state neurons using inter-spike intervals information," *J Neurosci Methods*, vol. 150, no. 1, pp. 16–29, 2006.

[5] Z. Nenadic and J. W. Burdick, "Spike detection using the continuous wavelet transform," *IEEE Trans. Biomed. Eng.*, vol. 52, pp. 74–87, 2005.

[6] B. S. Gutkin, B. Ermentrout, and M. Rudolph, "Spike generating dynamics and the conditions for spike-time precision in cortical neurons," *J. of Comput. Neuroscience*, vol. 15, no. 1, pp. 91–103, 2003.

[7] S. Roa, M. Bennewitz, and S. Behnke, "Fundamental frequency estimation based on pitch-scaled harmonic filtering," *ICASSP 2007*, vol. 4, pp. IV–397–IV–400, April 2007.

[8] Michael S. Lewicki, "A review of methods for spike sorting: the detection and classification of neural action potentials," *Network: Computation in Neural Systems*, vol. 9, no. 4, pp. 53–78, 1998.

[9] M. Basseville and I. V. Nikiforov, *Detection of Abrupt Changes: Theory and Application*, Prentice-Hall, 1993.

[10] N. Vaswani, "Additive change detection in nonlinear systems with unknown change parameters," *IEEE Trans. Signal Processing*, vol. 55, no. 3, pp. 859–872, 2007.

[11] L. Schwartz, *Théorie des distributions*, Hermann, 3rd edition, 1998.

[12] J. Mikusiǹski and T. K. Boehme, *Operational Calculus*, vol. 2, PWN Varsovie & Oxford University Press, Oxford, 1987.

[13] M. Fliess and H. Sira-Ramírez, "An algebraic framework for linear identification," in *ESAIM: COCV*, vol. 9, pp. 151–168. SMAI, 2003.

[14] M. Mboup, "Parameter estimation for signals described by differential equations," To appear in *Applicable Analysis*, 2008.

[15] S. P. Lloyd, "Least squares quantization in PCM," *IEEE Trans. Inform. Theory*, vol. 28, pp. 129–137, 1982.