

On a set of data for the membrane potential in a neuron

Reinhard Höpfner (hoepfner@mathematik.uni-mainz.de)

Institute of Mathematics, University of Mainz, Staudingerweg 9, D-55099 Mainz, Germany

Abstract: We consider a set of data where the membrane potential in a pyramidal neuron is measured almost continuously in time, under varying experimental conditions. We use nonparametric estimates for the diffusion coefficient and the drift in view to contribute to the discussion which type of diffusion process is suitable to model the membrane potential in a neuron (more exactly: in a particular type of neuron under particular experimental conditions).

Key words: diffusion processes, neuron, membrane potential, nonparametric inference

2000 Mathematics Subject Classification: 06 J 60, 62 M 05

1. Introduction

A popular model for the membrane potential in a neuron between successive spikes is a mean-reverting Ornstein-Uhlenbeck (OU) process

$$(OU) \quad d\xi_t = (a - b\xi_t) dt + \gamma dW_t$$

with parameters $a \in \mathbb{R}$, $b, \gamma > 0$. In OU type models, the volatility (the constant γ) is independent of the present state of the process. In many papers, the main emphasis is on level crossing times of this process, in order to describe a membrane potential starting from some fixed reset level and evolving up to the random time where some fixed excitation threshold is attained (out of many references, we quote [T 89, Ch. 5.3] and [D-L 05] on the mathematical and [S-S-F 99] on the biological side); this will not be our approach in the present paper. A second class of models for the membrane potential are mean-reverting stochastic processes of Cox-Ingersoll-Ross (CIR) type (here and the following, 'type' means 'up to suitable linear transformation of the state space', in order to obtain the biologically relevant scaling). The CIR diffusion, well known in stochastic finance since [C-I-R 85], is

$$(CIR) \quad d\xi_t = (a - b\xi_t) dt + \gamma \sqrt{\xi_t \vee 0} dW_t$$

with parameters $a, b, \gamma > 0$. In difference to OU type models, volatility increases in the CIR diffusion with the state of the process, and the law of ξ_t is concentrated on the half axis $[0, \infty)$, for every choice of a starting point $\xi_0 \equiv x_0 > 0$. (CIR) type models for the membrane potential appear – under the name of Feller neuronal models – in [L-L 87, (4.12) and Thm. 4] and [D-L 06]. With a view to information transmission in large systems of neurons, [B-H 06] use (time-inhomogeneous) CIR type diffusions for neurons processing a deterministic signal. To compare essential features of OU and/or CIR type modelization, [G-L-N-R 88] look to differential equations for moments of level crossing times, and [L-S-T 95] compare first passage time densities. The conclusion of [L-S-T 95] is that suitable choice of parameters allows to obtain from both OU or CIR type models very similiar shapes for first passage time densities. Hence, in this respect, there is no clear distinction, and one needs a different approach to make appear the difference. We shall make use of a third model, similiar to [W-T 79, formula (4)] or [L-L 87, Thm. 2], which we write (again up to a suitable shift on the state space) as

$$(Q) \quad d\xi_t = -\beta\xi_t dt + \sqrt{\tau^2 + \gamma^2\xi_t^2} dW_t$$

with parameters $\beta, \tau, \gamma > 0$. The typical feature of (Q) is a 'bowl-shaped' diffusion coefficient $\sigma^2(x) = \tau^2 + \gamma^2x^2$ whose center coincides with the zero of the drift $b(x) = -\beta x$. The invariant density is symmetric around 0 with polynomially decreasing tails, and has finite moments only up to order $< 1 + \frac{2\beta}{\gamma^2}$. We term models of type (Q) 'quadratic'.

One could think of quite arbitrary classes of diffusion processes to model the membrane potential in a neuron between successive spikes. The question which type of model is adequate, more precisely, is adequate for a particular neuron under specified experimental conditions, can only be answered from data. The aim of the present paper is to contribute to this discussion by exploring a given set of membrane potential data. We view spikesless segments of the membrane potential – recorded at high frequency over a long time interval (in this sense 'almost continuously in time') – as trajectory of a diffusion process solving a stochastic differential equation (SDE)

$$(1) \quad d\xi_t = b(\xi_t) dt + \sigma(\xi_t) dW_t$$

with drift coefficient $b(\cdot)$ and diffusion coefficient $\sigma^2(\cdot)$; both $b(\cdot)$ and $\sigma^2(\cdot)$ are considered as unknown \mathcal{C}^1 functions. We estimate $b(\cdot)$ and $\sigma^2(\cdot)$ using nonparametric estimators to make appear typical shapes of drift and diffusion coefficient. This gives an indication which type of model

might be appropriate to describe the membrane potential between successive spikes in this neuron, under the given experimental conditions. In some of our membrane potentials, the neuron emits spikes, and a modelization (1) can be valid only on suitable intervals between successive spikes. Whenever spikes are present, we exclude neighbourhoods of these (in the sense of a suitable time interval) from our considerations, in order to avoid contamination of the estimates for $b(\cdot)$ and $\sigma^2(\cdot)$ induced by the typical (almost deterministic) shape of the spikes. We stress that the remaining segments of membrane potential on which we work in this case should not be understood in the spirit of level crossing times ([G-L-N-R 88], [L-S-T 95], [D-L 06]): in our data, spikes occur at a very low frequency, and we select long segments of membrane potential between successive spikes on which the maximal value in general sensibly exceeds the terminal value.

The surprise contained in our data was the following. The same neuron (one pyramidal neuron in a cortical cell slice preparation observed *in vitro*) was flexible enough to show essentially different features under different experimental conditions (exposition of the whole cell slice to different time-constant levels of potassium (K^+) which stimulates the networking properties of *all* neurons in the brain slice). Not only location characteristics – e.g., location of the overall median of the membrane potential data in one experiment – were changing with increasing level of potassium, but also the very shape of the diffusion coefficient.

When under high level of potassium (10, 12 or 15 [mM]) the observed neuron was able to emit spikes, we always got one type of answer: a linear shape (positive slope) of the diffusion coefficient combined with a linear shape (negative slope) of the drift. This characterizes a CIR type model.

In contrast to this, the outcome of the experiments under lower potassium concentration (from 3 to 9 [mM] in steps of 1 [mM]) was far from unique:

- i*) In one experiment (5 [mM] of K^+), the diffusion coefficient was constant, and the shape of the drift linear (negative slope). This characterizes an OU type model.
- ii*) Another experiment (8 [mM] of K^+) produced a picture close to a CIR type model.
- iii*) In some experiments (7,9 [mM] of K^+), a new feature appeared: on a large interval in the state space, containing points which were visited reasonably often, the diffusion coefficient seemed in first rough approximation constant on a large part of the state space, or linear ha-

ving very small (positive) slope; on a subinterval however, containing those points which were visited extremely often, the estimates made appear a bowl-shaped form of the diffusion coefficient together with a linear shape of the drift (negative slope) where the center of the 'bowl' was coinciding with the zero of the drift. This suggests a model of type (Q), at least on the corresponding (perhaps small) part of the state space.

Let us finally mention that our data – even in regimes where at a first glance we might believe to see a stationary behaviour – are far from being truly stationary processes, and different features may prevail in different parts of the same trajectory.

The paper is organized as follows. In section 2, we present the structure of our data and explain the nonparametric estimators which we will apply to these. More background concerning the estimators is given in section 5. The statistical analysis of the membrane potential data is presented in two sections: section 3 is concerned with high levels of potassium where the neuron was able to generate spikes (in our data, these are few and isolated), and section 4 with membrane potentials where (low levels of potassium) no spikes were emitted within the time interval of observation. A conclusion is given in section 6.

Acknowledgement: I would like to thank H. Luhmann and W. Kilb (Institute of Physiology and Pathophysiology, University of Mainz) for a lot of stimulating discussions on neurons, spikes and membrane potentials, and for the set of data investigated in this paper, recorded by W. Kilb. Financial support from Rheinland-Pfalz Stiftung für Innovation is gratefully acknowledged. Thanks are also due to two anonymous referees for careful reading and critical questions.

2. Our data and our estimators

We have data from 10 experiments, named 'level 1' to 'level 10', recording the membrane potential in *the same neuron* – one pyramidal neuron belonging to a cortical slice preparation from a mouse, observed in vitro – under different experimental conditions, see figure 1. In every experiment, the whole brain slice is exposed to some level of potassium (K^+). This level is kept constant during the experiment, and stimulates the networking properties of all neurons belon-

ging to the brain slice. Increasing the potassium level from 'level 1' to 'level 10' (from 3 to 10 [mM] in steps of 1 [mM], then 12 [mM] and 15 [mM]), we hope to observe effects exerted by the network on the neuron which is part of it.

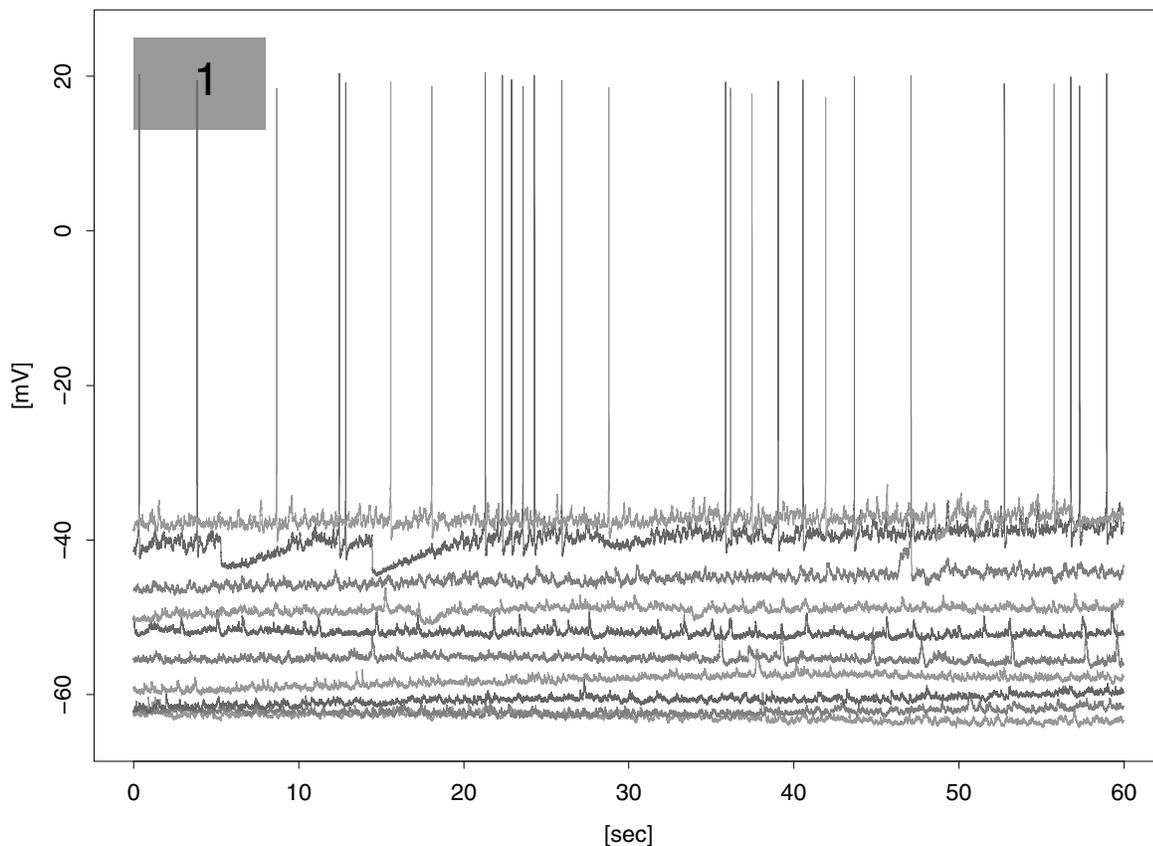


Figure 1: Membrane potential in one neuron (in a cortical slice preparation, observed in vitro, under different levels of potassium).

In three experiments (high level 10, 12, 15 [mM] of potassium, 'level 8' to 'level 10'), the neuron generates isolated spikes at a very low frequency; the remaining seven experiments (low levels of potassium) were spikeless. The duration of every experiment is 60 seconds. The time grid for successive measurement of the membrane potential has mesh $6 \cdot 10^{-4}$ [sec]. Data are in [mV] and are exported formally with three decimals; we were adverticed however that the last decimal is

not biologically reliable. This is the structure of our data.

We explain the nonparametric estimators which we will apply to our data. We view spikeless segments of the membrane potential as trajectories of a diffusion process (1)

$$d\xi_t = b(\xi_t) dt + \sigma(\xi_t) dW_t \quad \text{on some time interval } [t_0, t_1]$$

driven by standard Brownian motion $W = (W_t)_{t \geq 0}$, for which we dispose of a set of discrete observations

$$(2) \quad X_{i\Delta}, \quad i_0 \leq i \leq i_1, \quad i_0 := \lceil \frac{t_0}{\Delta} \rceil, \quad i_1 := \lfloor \frac{t_1}{\Delta} \rfloor, \quad \Delta = 6 \cdot 10^{-4} \text{ [sec]}.$$

From data (2), we wish to estimate drift

$$(3) \quad x \longrightarrow b(x)$$

and diffusion coefficient

$$(4) \quad x \longrightarrow \sigma^2(x)$$

in equation (1). We estimate values of the functions (3)–(4) at points a using the kernel estimators

$$(5) \quad \widehat{\sigma}^2(a) := \widehat{\sigma}_{(\Delta, M, h)}^2(a) = \frac{\sum_{i=i_0}^{i_1-M} K\left(\frac{X_{i\Delta}-a}{h}\right) \left(\frac{X_{(i+M)\Delta}-X_{i\Delta}}{\sqrt{\Delta \cdot M}}\right)^2}{\sum_{i=i_0}^{i_1-M} K\left(\frac{X_{i\Delta}-a}{h}\right)}$$

$$(6) \quad \widehat{b}(a) := \widehat{b}_{(\Delta, M, h)}(a) = \frac{\sum_{i=i_0}^{i_1-M} K\left(\frac{X_{i\Delta}-a}{h}\right) \left(\frac{X_{(i+M)\Delta}-X_{i\Delta}}{\Delta \cdot M}\right)}{\sum_{i=i_0}^{i_1-M} K\left(\frac{X_{i\Delta}-a}{h}\right)}$$

based on suitable choice of a bandwidth $h > 0$, on two given kernels

$$K(y) = \frac{1}{2} 1_{(-1, +1)}(y) \text{ (rectangular)}, \quad K(y) = (1 - |y|) 1_{(-1, +1)}(y) \text{ (triangular)},$$

and on M -step Δ -increments in the trajectory, for suitable integer M . In all cases where spikeless segments (more than one) are extracted from the same membrane potential trajectory, we calculate numerators resp. denominators of (5) or (6) separately in every segment, then we add up corresponding terms from different segments. Our choice of h and M (comments on robustness with respect to choice of h and M are given in remark 4.7 and figure 10 below) is

$$(7) \quad h = 0.01 \quad , \quad M = 20.$$

This choice (7) makes two quantities in (5) and (6) simultaneously 'small':

i) the length $M\Delta$ of time intervals over which we calculate increments of the process ξ :

$$M\Delta = 0.012 ;$$

ii) the bandwidth h in comparison to (a certain power of) this length $M\Delta$:

$$\frac{h}{(M\Delta)^{1/3}} = 0.04 .$$

Our data set provides values for the $X_{i\Delta}$ in [mV] with formal accuracy ± 0.001 [mV], the third decimal being biologically unreliable. Hence the bandwidth h in (7) is close to a lower bound for possible choices of bandwidth in our data set (compare with (8) below). The step-size M for Δ -increments in (7) is then subject to some trade-off in view to get the quantities in both criteria i) and ii) simultaneously small. The power $\frac{1}{3}$ relating the bandwidth to the length of time intervals for ξ -increments as well as the choice of 'kernels of order one' above (see [T 04]) is adequate if we think of (3) and (4) as unknown \mathcal{C}^1 functions. A detailed discussion relating our estimates (5)–(7) to rigorous results on convergence of nonparametric estimates in diffusions (e.g. [FZ 93], [H 01], [K 03]) is postponed to section 5. Here we stress only one major point: the (observable) random function

$$(8) \quad a \quad \longrightarrow \quad \sum_{i=i_0}^{i_1} 1_{(a-\frac{h}{2}, a+\frac{h}{2})}(X_{i\Delta})$$

plays the role of an 'information' in the data (2) in the sense that it allows to appreciate the variance which we can expect for estimators $\widehat{\sigma^2}(a)$ or $\widehat{b}(a)$ at different points a , depending on the data and the choice of the bandwidth. We will call (8) 'local time' – by strong abuse of language, and appearing systematically in quotation marks – in this note.

In view of (8), we estimate $\sigma^2(a)$ and $b(a)$ only at points a where the variance of our estimators can be expected to be not too large; hence we fix some minimal level for (8) and use only points a with 'local time' > 25 . At points a where 'local time' in the sense (8) is large (e.g., > 100 or > 300), we expect a better concentration of the estimates (in terms of variance, note that some bias is remaining). In some cases with 'local time' > 25 , a few obvious outliers (never more than 3 points $(a, \widehat{\sigma^2}(a))$ or $(a, \widehat{b}(a))$, far away from the cloud formed by all other points) are not represented graphically. Points $(a, \widehat{\sigma^2}(a))$ or $(a, \widehat{b}(a))$ calculated with rectangle kernel are always represented as black diamonds \diamond . Points calculated with triangle kernel are shown solid and grey. We always calculate both versions. Whenever a regression (linear or quadratic)

is shown within a cloud of points $(a, \widehat{\sigma^2}(a))$ or $(a, \widehat{b}(a))$, it has been calculated with respect to both types of points taken together (the results using any of the two kernels separately being quite similar), and including possible outliers in the above sense of graphical representation.

3. Higher potassium levels: membrane potentials with spikes

In this section, we consider the higher levels of potassium where the membrane potentials contain spikes, i.e. 'level 8' (10 [mM] of K^+), 'level 9' (12 [mM]), and 'level 10' (15 [mM]). The spikes which occur over the total observation time of 60 seconds (1 spike in 'level 8', 18 spikes in 'level 9', 8 spikes in 'level 10') are few and isolated. We analyse the membrane potential in segments between successive spikes, and exclude for every spike – writing τ for the time of its maximal amplitude – data corresponding to a neighbourhood $(\tau - 0.12, \tau + 0.18)$ [sec] of the spike time from our considerations (see figure 4H for an illustration). This guarantees that the resulting segments of membrane potential between successive spikes are not affected by the particular (almost deterministic) shape of the spikes. The maximal amplitude of the spikes is $\approx +20$ [mV]. We start by formulating the résumé which we draw from our data for the potassium levels 10, 12, 15 [mM].

Resumé: For the neuron under consideration, at potassium level 10, 12, 15 [mV], the diffusion coefficient $\sigma^2(a)$ is an increasing function of the value a of the membrane potential. The shape of the drift is linear with negative slope. Here (CIR) seems to be an adequate modelization for the membrane potential between successive spikes.

In the remaining part of this section, we give all details concerning 'level 8' (10 [mM] of K^+ , figure 2), 'level 10' (15 [mM], figure 3), 'level 9' (12 [mM], figure 4).

3.1 Discussion of 'level 8': One single spike (at time $\tau = 47.133$ [sec]) together with his neighbourhood $(\tau - 0.12, \tau + 0.18)$ [sec] has been removed, and we work over the two remaining segments of the membrane potential.

a) **Diffusion coefficient in 'level 8':** Figure 2A shows estimated values for $\sigma^2(a)$ calculated by (5)+(7) on a grid of points a having 'local time' > 25 . The shape of 'local time' defined by (8) is shown in figure 2D. It indicates that the most reliable estimates for $\sigma^2(a)$ (in the sense of

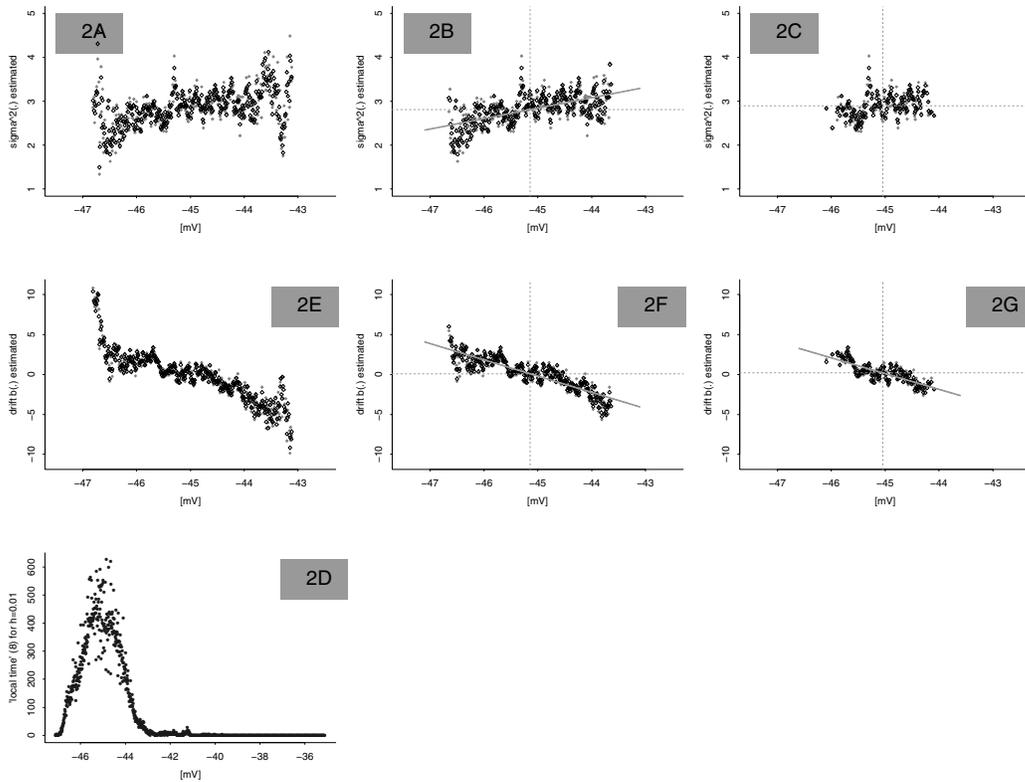


Figure 2: Diffusion coefficient and drift estimated in 'level 8' (10 [mM] of K^+) by means of (5)–(7). We estimate diffusion coefficient $\sigma^2(a)$ and drift $b(a)$ at points a having 'local time' > 25 in figures 2A and 2E, > 100 in 2B and 2F, > 300 in 2C and 2G. The shape of this 'local time' (8) – with $h = 0.01$ – is shown in 2D.

variance; some bias is remaining) are obtained for a -values in the region from -46 to -44 [mV]. Whenever 'local time' at a is low, estimated values $\widehat{\sigma^2}(a)$ present large fluctuations; this is seen in figures 2A for a -values close to -47 and -43 [mV].

Figure 2B shows the subcloud of points $(a, \widehat{\sigma^2}(a))$ where a has 'local time' > 100 : we find a correlation coefficient of 0.59 (this is not more than just acceptable) and a regression line

$$x \longrightarrow 0.24(x + 56.95) .$$

Figure 2C restricts further to points $(a, \widehat{\sigma^2}(a))$ where 'local time' at a is > 300 . Here the correlation coefficient (only 0.37) is too low to credit a linear regression. There is some lack of stationarity in the trajectory 'level 8' – this can be seen from occupation time measures calcu-

lated separately for the first and the last 20 [sec] of the experiment – which may be responsible for this last result.

b) **Drift in 'level 8'**: Figure 2E shows estimated values for $\hat{b}(a)$ calculated by (6)+(7) on the grid of points a with 'local time' > 25 . As above, we get large fluctuations of $\hat{b}(a)$ whenever 'local time' at a is low. A linear shape of the drift emerges on the segment where 'local time' is high enough. In figure 2F ('local time' > 100) we find a correlation coefficient of -0.88 and a regression line

$$x \longrightarrow -2.04(x + 45.10) ;$$

in figure 2G ('local time' > 300) the correlation still attains -0.80 , and the regression line remains almost unchanged. Hence the drift in 'level 8' is in good approximation linear. \square

3.2 Discussion of 'level 10': 8 spikes and their neighbourhoods have been removed, and we work on the remaining 9 segments. Figures 3D for the 'local time', 3A – 3C for the diffusion coefficient, 3E – 3G for the drift have been calculated and represented graphically in exact analogy to the corresponding parts of figure 2.

a) **Diffusion coefficient in 'level 10'**: Given the shape of 'local time' as a function of a (figure 3D), we expect to find the most reliable estimates for $\sigma^2(a)$ (in terms of variance) at a -values between ≈ -38.5 and ≈ -36.5 [mV]. Figure 3A shows estimated values of $\sigma^2(a)$ at points a with 'local time' > 25 . There is some intriguing convexity for a -values in the interval $(-36.5, -35)$ [mV], still present in figure 3B ('local time' > 100) where the best quadratic approximation is

$$(9) \quad x \longrightarrow 4.13 + 1.77(x + 38.10)^2 .$$

In figure 3C ('local time' > 300), a linear regression

$$(10) \quad x \longrightarrow 2.23(x + 39.75)$$

is justified by the high value 0.931 of the correlation coefficient; both correlation and (10) remain almost unchanged if we restrict further to points a having 'local time' > 500 .

A closer look to the trajectory 'level 10' - see figure 3H – shows that the interval $(-36.5, -35)$ [mV] where convexity in 3A and 3B was most visible is visited mainly at times which correspond to strong 'pulses' (significant peaks in the trajectory which are not part of a spike) in the membrane potential. Such 'pulses', absent in 'level 8', are quite frequent in spikeless segments of 'level 10'. We comment on this problem in remark 3.4 at the end of this section. Influence

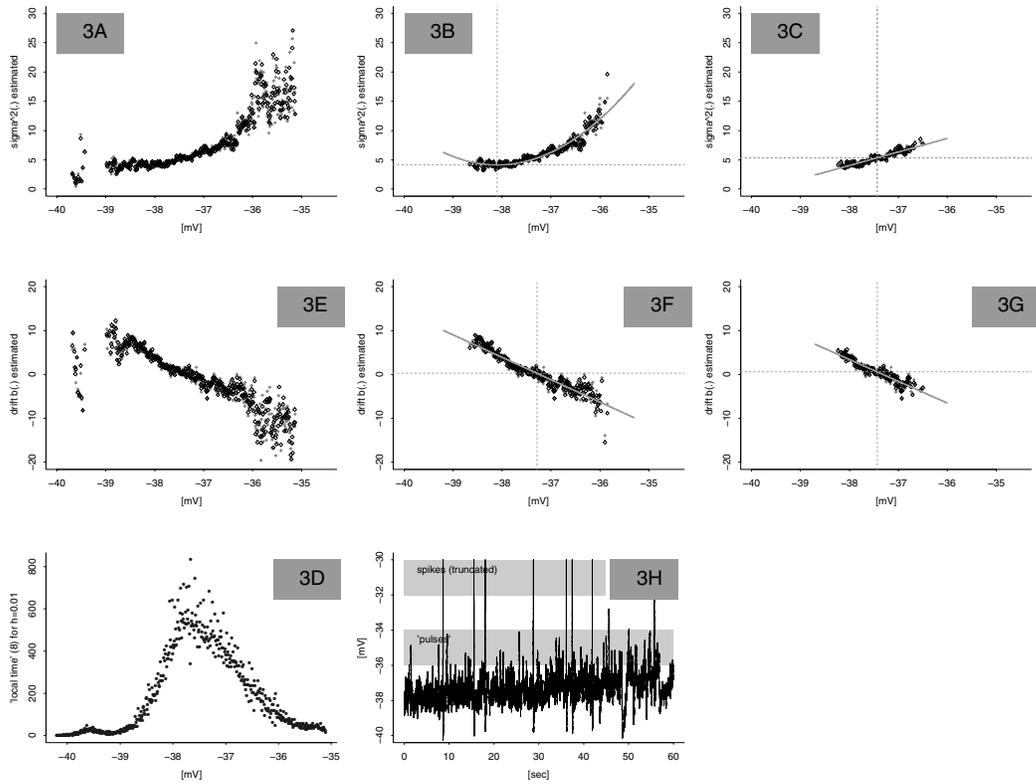


Figure 3: Diffusion coefficient and drift estimated in 'level 10' (15 [mM] of K^+). Graphics 3A – 3G have been calculated in exact analogy to 2A – 2G. The region above -36 [mV] is visited mainly through 'strong pulses' (significant peaks in the trajectory which are not part of a spike) as seen from figure 3H.

of 'pulses' is avoided when we calculate in (10) the regression for the subset of points where a has 'local time' > 300 as in figure 3C. Note that the derivative of (9) at the mean value -37.43 [mV] of the a -points appearing in figure 3C is close to the slope in (10). Hence we prefer a linear modelization for the central part of the cloud of points $(a, \widehat{\sigma^2}(a))$.

b) **Drift coefficient in 'level 10':** For 'local time' > 100 (figure 3F), we find a correlation coefficient of -0.94 and a regression line

$$x \longrightarrow -5.10(x + 37.23) .$$

In figure 3G ('local time' > 300), the correlation is still -0.92 , and the regression line

$$x \longrightarrow -4.97(x + 37.31)$$

almost unchanged: hence linearity of the drift is fairly evident. \square

3.3 Discussion of 'level 9': Neighbourhoods of 18 spikes have been removed. Figure 4H illustrates this for a part of the trajectory (spikes 13–15, truncated at -30 [mV]). Again values of the membrane potential close to -37 [mV] are attained mainly through 'strong pulses' (figure 4H). There is one striking particularity in 'level 9', visible already from figure 1: within the first 20 seconds of observation, there are two segments in the trajectory – one of these shown in detail in figure 4K – where the membrane potential suddenly breaks down to ≈ -44 [mv] and then recovers slowly until values close to -40 [mv] are reached; during both 'recovery periods', the trajectory itself suggests that volatility should be increasing with the state reached by the process.

a) **Diffusion coefficient in 'level 9':** The shape of local time in figure 4D indicates that the most reliable (in the sense of variance) estimates are expected when a is close to -40 or -39 [mV]. For 'local time' > 100 (figure 4B), a quadratic regression yields a curve which fits badly the left part of the point cloud (a -values below -41.5 [mV]), and is not at all compatible with the shape of the point cloud in figure 4A. A linear regression in 4B, with correlation of 0.8 and regression line

$$x \longrightarrow 1.32(x + 43.46) ,$$

overestimates the slope for the part of the point cloud with a -values < -38.5 [mV] (due to the convexity at ≈ -38 [mv], a region corresponding to 'pulses' in the trajectory, see figure 4H). For 'local time' > 300 (figure 4C), linear regression yields a line

$$(11) \quad x \longrightarrow 0.69(x + 46.41)$$

but remains questionable due to the low value of the correlation (only 0.58) in 4C.

The truth is probably best caught 'by eye' from figure 4A. First, the regression line (11) fits surprisingly well into the cloud of points $(a, \widehat{\sigma^2}(a))$ in figure 4A if one restricts attention to a -values below -38.5 [mV]. Second, an inspection of the whole trajectory (figure 1) shows that a -values below -42 [mv] are visited exclusively during the two 'recovery periods' mentioned above: (the second one is represented in figure 4K). These two 'recovery periods' being short, 'local time' for a -values below -42 [mV] is necessarily low. An estimation of $\sigma^2(\cdot)$ based only on data belonging to one of the 'recovery segments' – we have selected the time interval [15, 20] [sec] represented in 4K – supports (11) in a rigorous way: this small segment yields new points

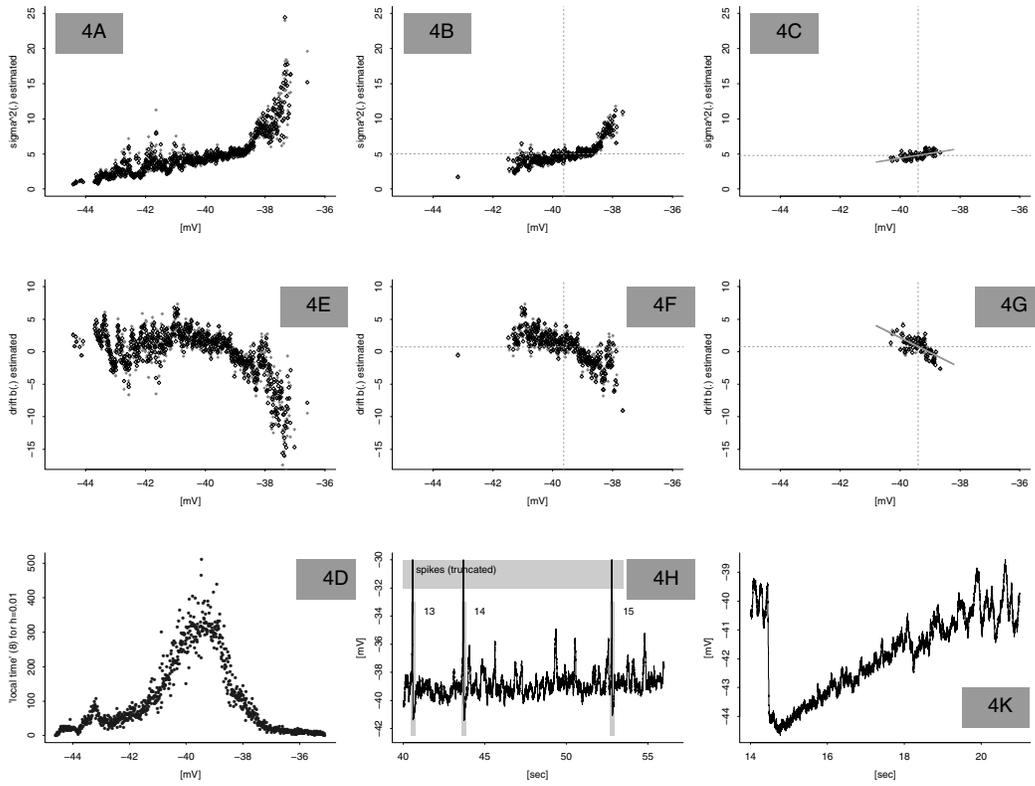


Figure 4: Diffusion coefficient and drift estimated in 'level 9' (12 [mM] of K^+). Graphics 4A – 4G have been calculated in exact analogy to 2A – 2G. Figure 4K shows a detail of the trajectory whereas 4H illustrates segments of membrane potentials between successive spikes (excluding neighbourhoods $(\tau - 0.12, \tau + 0.18)$ around spike times τ).

$(a, \widehat{\sigma^2}(a))$ admitting a correlation coefficient of 0.73 and a regression with parameters

$$(12) \quad x \longrightarrow 0.69(x + 45.36) .$$

Together, (11)+(12) support convincingly a linear modelization of the diffusion coefficient for a -values below -38.5 [mV].

b) **Drift in 'level 9':** For 'local time' > 100 (figure 4F), the correlation is -0.73 , and the regression parameters are

$$x \longrightarrow -1.77(x + 39.22) .$$

Here the slope seems not steep enough for the point cloud $(a, \widehat{b}(a))$ in 4F, perhaps due to the

influence of the isolated point at $a \approx -43$. For 'local time' > 300 (figure 5G), we obtain the correlation coefficient -0.66 and the regression line

$$(13) \quad x \longrightarrow -2.29(x + 39.05) .$$

If we put the regression line (13) obtained for a -points with 'local time' > 300 over the point cloud $(a, \widehat{b}(a))$ of picture 4F ('local time' > 100), the result seems perfectly acceptable to approximate the shape of the point cloud in 4F.

Comparing figures 4K and 4E, we see that the 'recovery periods' (corresponding to a -values below -41 [mv]) play a particular role: the drift there is significantly different from e.g. the drift in the neighbourhood of -40 [mV] appearing in figure 4F. \square

3.4 Remark: At the end of this section, we give some comments concerning the 'pulses' which we have encountered in 'level 10' and in 'level 9', and which will also appear in some of the experiments discussed in section 4.

i) Biologically, 'pulses' seem to be the effect of synaptic input near the recording site, i.e. via synaptic endpoints located directly at the soma of the neuron, or nearby at dendrites such that the usual decay of the postsynaptic potential on its way along the dendrite to the soma is negligible.

ii) Mathematically, a diffusion model (1) seems adequate only on subintervals of the range of the membrane potential which are not affected by 'pulses'. To model 'pulses', an extension of SDE (1) to a diffusion with jumps seems useful, e.g. in form

$$(14) \quad d\tilde{\xi}_t = b(\tilde{\xi}_{t-}) dt + \sigma(\tilde{\xi}_{t-}) dW_t + \int_{(0,\infty)} x \mu(dt, dx)$$

where $\mu(ds, dx)$ is a Poisson point process – independent of Brownian motion W – with intensity $\tilde{a} dt \nu(dx)$ for some constant $\tilde{a} > 0$ and some measure ν on $(0, \infty)$ such that $\int_{(0,\infty)} (x \wedge 1) \nu(dx)$ is finite. Recent results on bipower variation in semimartingales (see [BN-G-J-P-S 06], [BN-G-J-S 05], [P 06]) provide tools to detect presence of jumps in a discretely observed trajectory. A modelization including jumps is not the topic of the present paper.

iii) We give an additional comment related to i). Spikes can be suppressed in the brain slice by administration of tetrodotoxin (TTX). TTX is blocking the ion channels involved in spike generation; under TTX we observe an artificially silenced network. In cells not belonging to the data set considered here, the membrane potential under TTX was – at comparable level of

potassium – free of pulses. □

4. Lower potassium levels: membrane potentials without spikes

In this section, we consider the data 'level 1' to 'level 7' where the membrane potential is not high enough to allow for generation of spikes. Again we start by formulating a resumé.

Resumé: The data 'level 1' to 'level 7' (potassium concentration from 3 to 9 [mM]) present three essentially different types of pictures.

- i) In some cases, the diffusion coefficient is constant, and the drift linear with negative slope. This corresponds to an OU type model.
- ii) In some cases, diffusion coefficient and drift allow – at least on subintervals of the range of the membrane potential – for a linear approximation in the sense of CIR type models.
- iii) A third group of cases has a diffusion coefficient which seems in first rough approximation constant or linear with very small positive slope, combined with linear drift having strongly negative slope; however, in restriction to the set of points $(a, \widehat{\sigma^2}(a))$ where a has highest 'local time', a clearly expressed bowl-shaped form of the diffusion coefficient appears whose minimum is located at the zero of the drift. This is a model of type (Q).

Hence for lower potassium concentration, the neuron under observation seems flexible enough to switch between the three types of models (CIR), (OU) and (Q) presented in the introduction. In the remaining part of this section, we give details for some selected levels.

4.1 Discussion of 'level 3': The data 'level 3' (5 [mM] of K^+) are shown in figure 5. The membrane potential is not a stationary process. For values of the membrane potential close to -61.5 [mV] – this occurs in the first 10 seconds of the experiment only, see figure 5H – some irregularity is visible in 5A+5B and 5E+5F, and corresponds to a secondary peak in the shape of the 'local time' 5D at -61.5 [mV]. Up to this, 'level 3' shows the type (OU).

a) **Diffusion coefficient in 'level 3':** Figure 5A shows estimated values for $\sigma^2(a)$ calculated by (5)+(7) on a grid of points a with 'local time' > 25 . The shape of 'local time' as function of a (figure 5D) indicates that the most reliable estimates for $\sigma^2(a)$ are obtained for a between

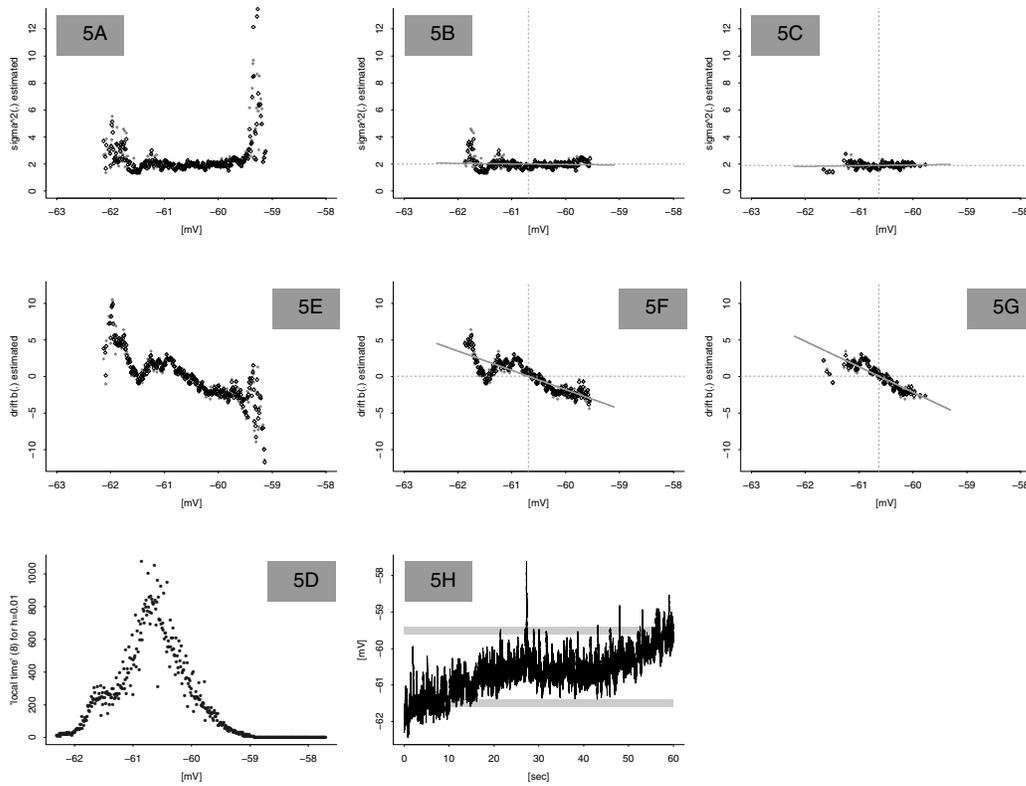


Figure 5: Diffusion coefficient and drift estimated in 'level 3' (5 [mM] of K^+). 5A – 5G are calculated in exact analogy to 2A – 2G. Figure 5H allows to associate the deviations from linearity seen in 5A and 5E (close to -59.5 and -61.5 [mV]) to the corresponding parts of the trajectory which is not stationary.

-61 and -60 [mV]. Between -61 and -60 [mV], the cloud of points $(a, \widehat{\sigma^2}(a))$ in figure 5A is flat. We have a correlation coefficient of -0.075 in figure 5B ('local time' > 100) and of $+0.1$ in figure 5C ('local time' > 300). The slope of the regression line is -0.05 in the first case, and $+0.06$ in the latter; hence the diffusion coefficient does not depend on the space variable, and takes a constant value close to 1.9. Restricting further to a -values having 'local time' > 500 the cloud of points $(a, \widehat{\sigma^2}(a))$ is flat.

b) **Drift in 'level 3':** The correlation coefficient in the cloud of points $(a, \widehat{b}(a))$ in figure 5F ('local time' > 100) is -0.85 , and the regression line with slope -2.36 is strongly influenced by the deviation from linearity at $a \approx -61.5$ [mV]. For 'local time' > 300 (figure 5G), the

correlation remains unchanged, the regression line

$$x \longrightarrow -3.48(x + 60.62)$$

is steeper; for a -values having 'local time' > 500 , we get a clear linear shape with regression line

$$x \longrightarrow -5.306(x + 60.57)$$

and correlation -0.92 . There is a strong time inhomeogeneity of the trajectory 'level 3' (in the drift, not in the diffusion coefficient), as seen from the first, second and third 20 seconds of the experiment considered separately. For the whole trajectory 'level 3', linearity of a drift function (3) with time-independent constant coefficients is only a rough approximation. \square

4.2 Discussion of 'level 6': Figure 6 shows 'level 6' (8 [mM] of K^+). We detect already from figure 1 a large number of 'strong pulses', in contrast to e.g. 'level 3' or 'level 8', see remark 3.4. For values of the membrane potential above -51.25 [mV], these 'pulses' add a seemingly chaotic part to the cloud of points $(a, \widehat{\sigma^2}(a))$ or $(a, \widehat{b}(a))$, see 6H, 6A, 6B. For values of the membrane potential below -51.5 [mV] we find – see 6C and 6G – a CIR type example.

a) **Diffusion coefficient in 'level 6':** The cloud of points $(a, \widehat{\sigma^2}(a))$ in figure 6A, with 'local time' > 25 , seems 'by eye' linear for a -values below ≈ -51.5 [mV]. According to figure 8D, the most reliable estimates are expected when a is close to -52 [mv]. Figures 6B ('local time' > 100) and 6C ('local time' > 300) support a linear approximation of the diffusion coefficient. The cloud of points in 6B is still influenced by 'pulses'; in 6C, we get a correlation of ≈ 0.85 , and a regression line

$$x \longrightarrow 2.0(x + 53.37) .$$

Restricting further to a -values having 'local time' > 500 – they range from -52.6 to -51.7 – we still find a correlation coefficient of 0.79 , and the slope of the regression line is 1.5 .

We mention that in both figures 6B and 6C, a best fitting quadratic does not catch well the shape of the point cloud $(a, \widehat{\sigma^2}(a))$ for a -values below -52 [mV].

b) **Drift in 'level 6':** In figure 6F ('local time' > 100), the correlation equals 0.95 , and the regression line is

$$x \longrightarrow -6.31(x + 52.03) ;$$

in figure 6G ('local time' > 300), the correlation is 0.94 , and the slope of the regression line changes to -5.65 . Hence, with respect to correlation coefficient or slope of regression line at

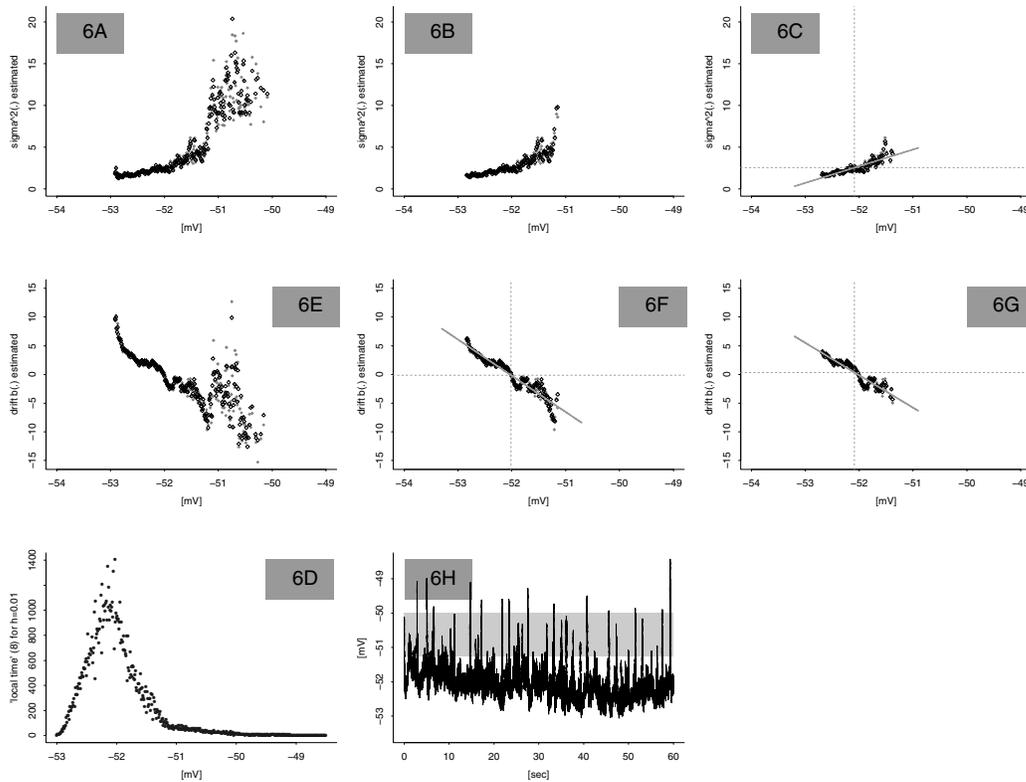


Figure 6: Diffusion coefficient and drift estimated in 'level 6' (8 [mM] of K^+). 6A –6G are calculated in exact analogy to 2A – 2G. Figure 6H allows to appreciate in comparison with 6A and 6E the effect of 'strong pulses' in the trajectory.

'local time' > 300 , 'level 6' is not much different from 'level 10' discussed in 3.2. \square

4.3 Discussion of 'level 5': At first glance, 'level 5' shown in figure 7 (potassium level of 7 [mM] in the brain slice) does not seem much different from 'level 6' discussed in 4.2; however, a difference appears when we look to a -values having 'local time' > 300 . There are some (few, but strongly expressed) pulses in the trajectory, see 7H and figure 1.

a) **Diffusion coefficient in 'level 5':** An inspection of the trajectory, cf. figure 7H, shows that a -values above -54.5 [mV] are attained at times of strong 'pulses', but not exclusively. In accordance with remark 3.4, we restrict the investigation to a -values below -54.5 [mV]. However, the part of figure 7A ('local time' > 25) with a -values < -54.5 [mV] seems difficult to judge: is

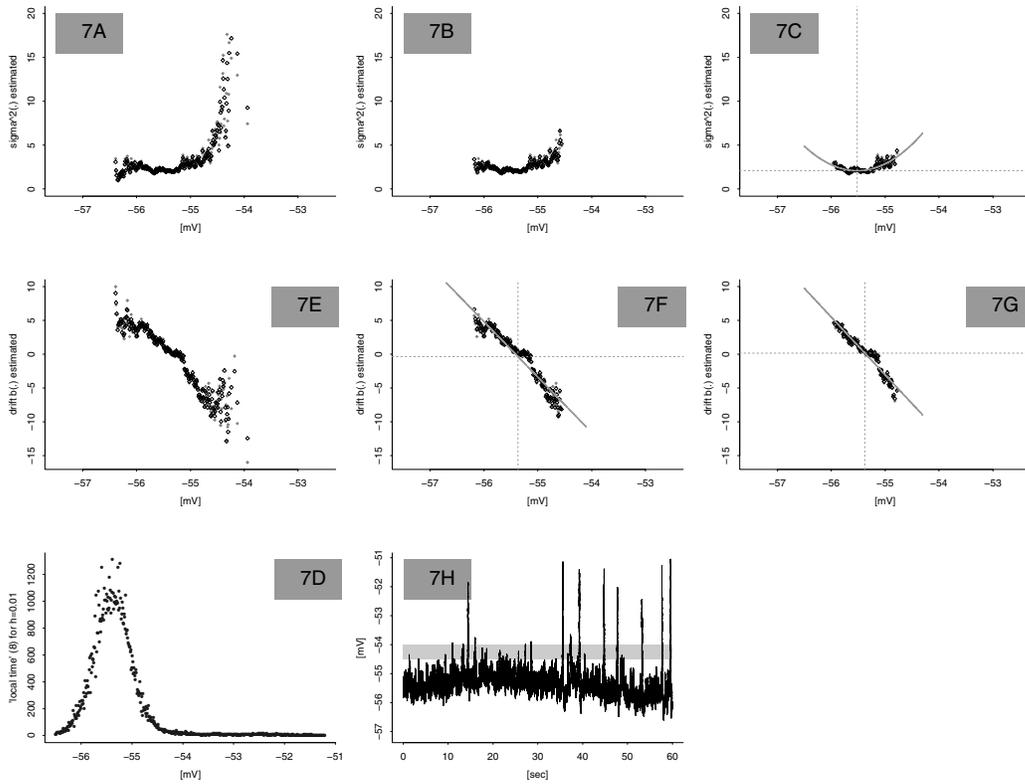


Figure 7: Diffusion coefficient and drift estimated in 'level 5' (7 [mM] of K^+). 7A –7G are calculated in exact analogy to 2A – 2G.

the cloud of points $\{(a, \widehat{\sigma^2}(a)) : a < -54.5\}$ 'close to flat', is there a very small positive slope, or should we view it as quadratic? According to the shape of 'local time' in figure 7D, we expect the most reliable (in the sense of variance) estimates for a -values between -55.75 and -55 [mV]. This coincides approximately with 'local time' > 300 where in 7C a clear 'bowl-shape' appears. Formally, we get for figure 7B ('local time' > 100) a correlation coefficient ≈ 0.59 (this is not more than just acceptable), and a linear regression with parameters

$$x \longrightarrow 0.98(x + 58.12)$$

for a -values with 'local time' > 300 (figure 7C), the correlation is reduced to 0.56, and the slope to 0.83. In both cases, the regression line does not catch the typical feature of the cloud of points $(a, \widehat{\sigma^2}(a))$ which is much better described by a quadratic function; note that all this happens

at points a having high 'local time' where our estimates are reliable. In figure 7C ('local time' > 300), the best fitting quadratic

$$(15) \quad x \longrightarrow 2.1 + 2.89(x + 55.52)^2$$

represents convincingly the cloud of points. In 7B ('local time' > 100) the quadratic fit is

$$(16) \quad x \longrightarrow 2.09 + 2.39 * (x + 55.57)^2 ;$$

for 'local time' > 500 , the conclusion concerning 'linear versus quadratic' remains unchanged.

Drift in 'level 5': For the points $(a, \widehat{b}(a))$, we see a well expressed linear shape in figures 7F and 7G. In 7G ('local time' > 300), the correlation is 0.97, and the regression is

$$(17) \quad x \longrightarrow -8.53(x + 55.35) ;$$

the values in 9F ('local time' > 100), with correlation 0.94 and regression line

$$(18) \quad x \longrightarrow -8.2(x + 55.41)$$

are similar. Note that the center of the quadratic approximation (15) or (16) to the diffusion coefficient is very close to the zero of the linear approximation (17) or (18) of the drift. Hence 'level 5' is an example for a model of type (Q). \square

Discussion of 'level 7': The results for 'level 7' (9 [mM] of K^+) are shown in figure 8. The situation is very similar to 'level 5', with quadratic shape of the diffusion coefficient at sufficiently high 'local time' even more clearly expressed than there.

a) **Diffusion coefficient in 'level 7':** In contrast to 'level 5', there is only one strong 'pulse' in 'level 7' at time 15.25 [sec], see 8H and figure 1. There are some breakdowns in the trajectory which influence the estimates for a -values between -51 and -50 [mV], cf. 8H, 8A and 8E. . Figure 8A presents a picture of estimates for the diffusion coefficient ('local time' > 25) which for a -values below -48 [mv] may seem roughly linear with small positive slope. However, already for 'local time' > 100 , a linear regression in 8B is not suitable, the correlation being as low as 0.45. At 'local time' > 300 (figure 8C), a bowl-shaped form of the cloud of points $(a, \widehat{\sigma^2}(a))$ appears very clearly and is well represented by the best fitting quadratic

$$(19) \quad x \longrightarrow 1.99 + 3.20(x + 49.07)^2 .$$

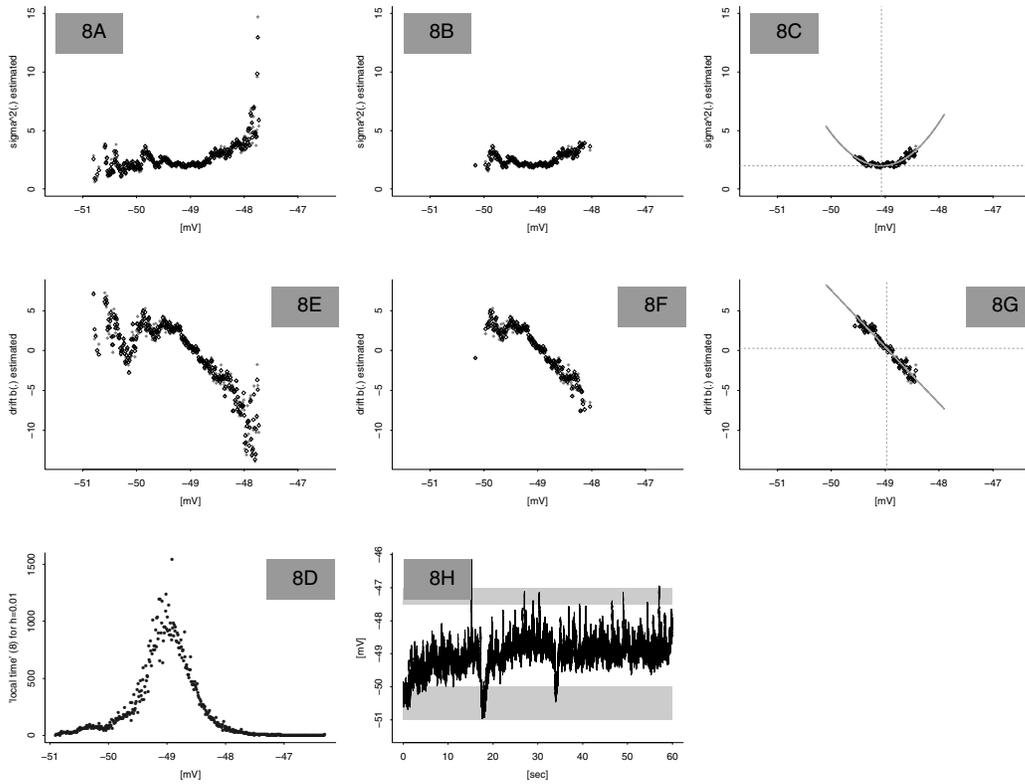


Figure 8: Diffusion coefficient and drift estimated in 'level 7' (9 [mM] of K^+). 8A –8G are calculated in exact analogy to 2A – 2G. It is seen in 8C and 8G that a Q type model describes well diffusion coefficient and drift in some neighbourhood of -49 [mV].

This bowl-shape persists almost unchanged for local time > 500 (hence it must represent some intrinsic property of the diffusion coefficient), and vanishes only for 'local time' > 900 where a very small cluster of points $(a, \widehat{\sigma^2}(a))$ – compare with 8D – remains, concentrated almost exactly at the center of the quadratic approximation in figure 8C.

Drift in 'level 7': At 'local time' > 100 (figure 8F), even with correlation -0.92 , the linear approximation to the estimates for the drift is not completely satisfactory, still influenced by the breakdowns of the trajectory in 8H. At 'local time' > 300 (figure 8G), the linear regression is almost ideal, the correlation being -0.97 : the parameters of the regression are

$$(20) \quad x \longrightarrow -7.08(x + 48.93) .$$

The center of the quadratic fit (19) in 8C is very close to the zero of the linear approximat-

on (20) in 8G. Hence 'level 7' at 'local time' > 300 is a convincing example of a type Q model. \square

4.5 Remark: We mention that an algorithm written by the authors of [C-R 02] from a completely different point of view – nonparametric regression and penalized model selection in time series — applied recently to our data 'level 3', 'level 5' and 'level 7' came to the same conclusion as our analysis in 4.1, 4.3 and 4.4: it detected a constant diffusion coefficient for 'level 3', and a quadratic diffusion coefficient for 'level 5' and 'level 7'.

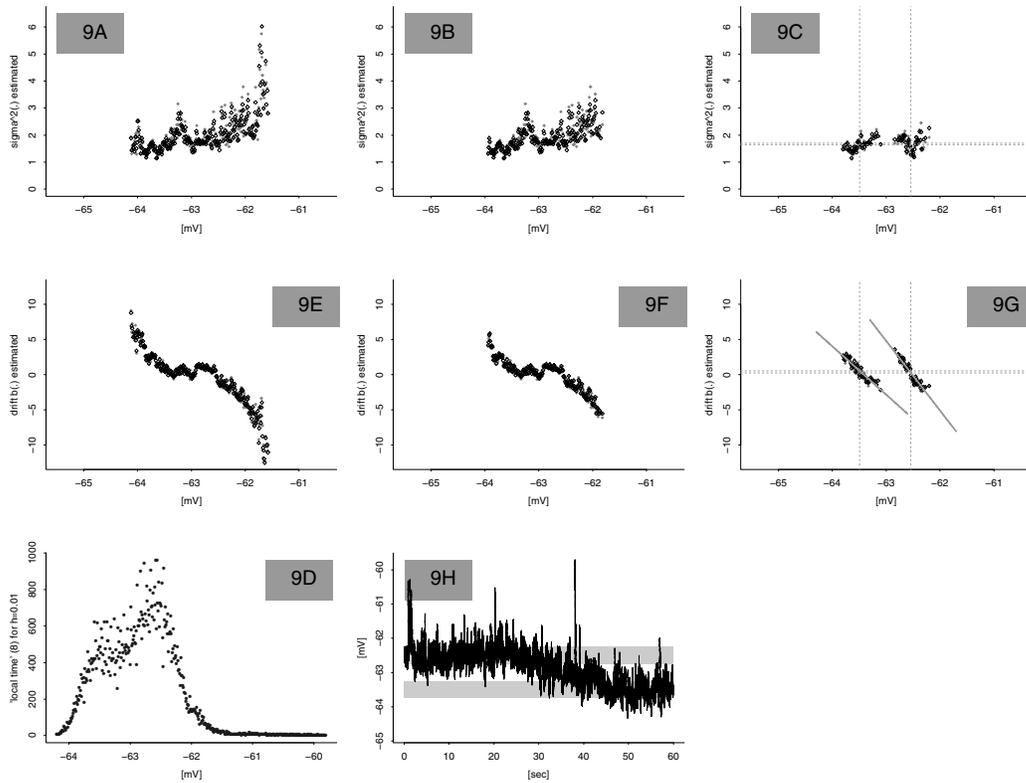


Figure 9: Time homogeneity is seriously violated in the data set 'level 1' (3 [mM] of K^+), cf. 9H. Figures 9A+B, 9E+F, 9D are in exact analogy to 2A+B, 2E+F, 2D. Figures 9C and 9G evaluate the estimates (5)–(7) separately on the time intervals $[0, 20]$ [sec] and $[40, 60]$ [sec], and show different regimes on different time intervals.

4.6 Discussion of 'level 1': The data set 'level 1' (potassium level of 3 [mM], figure 9) shows clearly expressed time inhomogeneities. Since – to some minor extend – inhomogeneities were also present in the previously discussed levels, we discuss this problem in detail here. Figure 9H indicates that the trajectory 'level 1' changes its regime during the total observation time of 60 seconds.

a) **Diffusion coefficient in 'level 1':** A first glance on figure 9A might suggest a linear diffusion coefficient with positive slope for a -values below -62 [mv]. For 'local time' > 100 (figure 9B), we get a correlation of 0.61 (not more than just acceptable) and a regression line which does not catch the doubly bowl-shaped structure of the cloud of points $(a, \widehat{\sigma^2}(a))$. Using only a -values with 'local time' > 300 (not shown), the double bowl persists, whereas the value of the correlation coefficient (as low as 0.48) is discrediting any linear approximation.

In fact, the shape of 'local time' (figure 9D) reveals the presence of two peaks (the larger one at ≈ -62.5 , the smaller one at ≈ -63.5 [mV]), and a closer look to the trajectory in 9H reveals the lack of stationarity: during the first 20 seconds of the experiment, the membrane potential fluctuates around -62.5 [mV]; then it diminishes slowly; during the last 20 seconds, it stabilizes again around -63.5 [mV]. In figure 9C, we have used only data from either the first 20 seconds or the last 20 seconds of the experiment, and – separately in the two new data sets – only a -values with 'local time' > 300 . Two well distinguished clusters appear, where points $(a, \widehat{\sigma^2}(a))$ with a -values close to -62.5 [mv] correspond to data from the first 20 seconds of the trajectory, the others to the last 20 seconds. Both clouds of points $(a, \widehat{\sigma^2}(a))$ are too small to support a rigorous further modelization; the mean value of all estimates for $\sigma^2(\cdot)$ is 1.71 for the first 20 seconds, and 1.62 for the last 20 seconds.

b) **Drift in 'level 1':** The picture for the drift is easy to interpret. Again we use for figure 9G only data either from the first 20 seconds or from the last 20 seconds of the experiment, and get a convincing linearity separately in each of both clouds of points $(a, \widehat{b}(a))$. During the first 20 seconds, the correlation is -0.94 and the regression parameters are

$$x \longrightarrow -9.92(x + 62.52) ,$$

during the last 20 seconds, we have a correlation of -0.92 and a regression line

$$x \longrightarrow -6.87(x + 63.41) .$$

This supports the interpretation of 9H and 9D in the sense of a diffusion process switching within the total observation time of 60 seconds from one regime to another. \square

It remains to discuss the crucial question whether – for the neuron under consideration and under the given experimental conditions – diffusion processes (1) are the right tool to model and to understand the membrane potential between successive spikes. Usually there is ‘microstructure noise’ in real data (2) recorded with time resolution Δ . It can be predominant when for small Δ , the Δ -increments are taken directly from a data set (2) and are used to estimate drift or diffusion coefficient in the underlying process (1). To overcome this problem, the usual device – well known in financial data – is to work for small Δ with M -step Δ -increments where M – not too small – can average out ‘microstructure noise’. Now, we can check whether or not our data are compatible with a diffusion process modelling. By definition, in a diffusion process limits (22) and (23) below must exist: at the level of the data, we mimic limits (22) and (23) by taking $M\Delta = s$ where we change M . Hence, for a true underlying diffusion process, estimators (5)–(6) which are build on (22) and (23) have to produce results which remain stable under moderate changes of M . The following remark 4.7 – together with figure 10 – shows that this is the case for our data provided M is not too small.

4.7 Remark: We discuss the question of robustness of our results with respect to choice of bandwidth h and M -step Δ -increments in the definition of the kernel estimator (5) for the diffusion coefficient. We illustrate this in the data sets ‘level 3’ (OU type), ‘level 7’ (Q type), ‘level 6’ (CIR type), discussed in 4.1, 4.4 and 4.3 above for $M = 20$ and $h = 0.01$, according to the choice made in (7). We change the bandwidth to $h = 0.007, 0.01, 0.013$, and the step size M to $M = 10, 20, 40$.

The result is seen in figure 10: we show regression lines or best fitting quadratics to clouds of points $(a, \widehat{\sigma^2}_{(\Delta, M, h)}(a))$ for varying M and h . In every subgraphs 10A – 10K, h is fixed, and we calculate the estimators at points a which have ‘local time’ (8) – defined with respect to h – at least 300. We represent the result obtained for $M = 20$ by a solid line. For $M = 10$ we use a dotted line, for $M = 40$ a dashed line. The outcome shows that our results do remain stable under suitable changes of M and h .

Even for $M = 80$ (not shown) we obtain in most cases only a small deviation from what we show in figure 10. Essential differences arise for $M = 5$ (not shown): the shape of the point cloud (linear or quadratic) remains unchanged, but the estimated values using $M = 5$ are about 1.5 times larger. Whereas the results for $M = 10, 20, 40$ are concordant, $M = 5$ is too small. \square

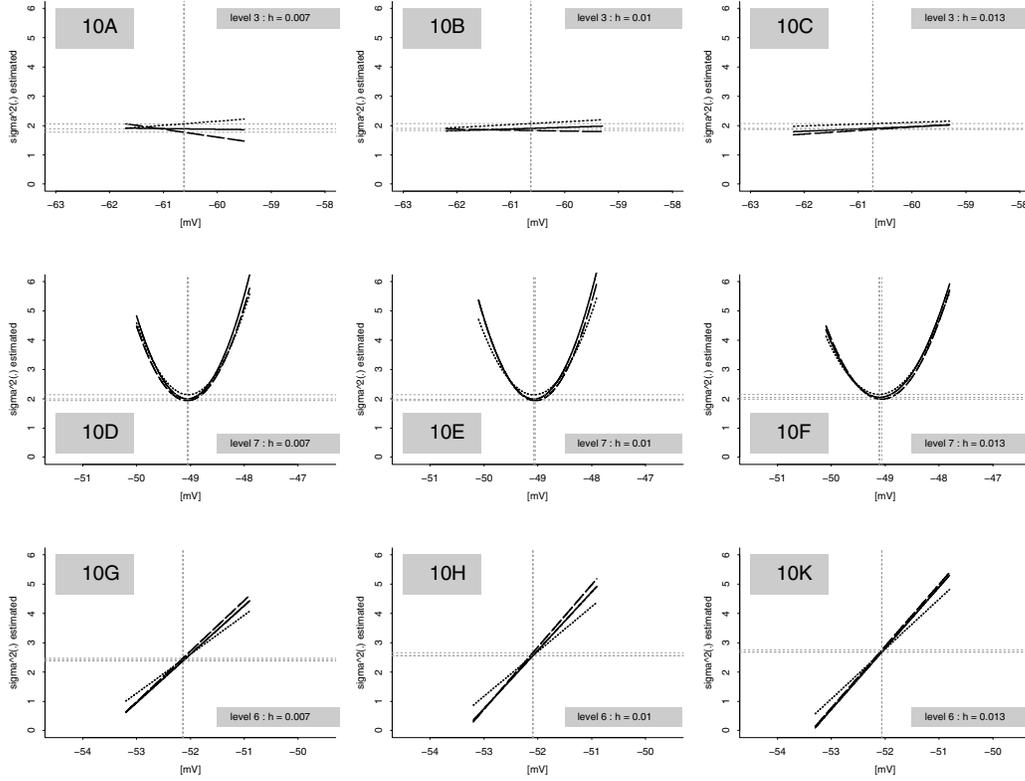


Figure 10: In 'level 3' (OU type), 'level 7' (Q type), 'level 6' (CIR type), we modify h and M in the kernel estimator (5) for the diffusion coefficient. For fixed h , we show best (linear or quadratic) approximations to the cloud of points $(a, \widehat{\sigma^2}_{(\Delta, M, h)}(a))$ with 'local time' > 300 in the cases $M = 10$ (dotted), $M = 20$ (solid), $M = 40$ (dashed).

Two data sets remain, 'level 2' and 'level 4', which we will not discuss in detail; see figure 11 for the diffusion coefficient at 'local time' > 300 . In 'level 2', the bowl-shape appears which characterizes models of type (Q), whereas strong time inhomogeneities are present in 'level 4'.

5. On nonparametric estimates for diffusion coefficient and drift

Our choice (5)–(7) of nonparametric estimators for drift and diffusion coefficient differs from the usual form (e.g. [FZ 93] for the diffusion coefficient) in one important aspect: our estimators

work with overlapping time intervals over which increments of the process ξ are evaluated. In this section, we discuss this choice and relate it to known results on convergence of nonparametric estimates in diffusion processes. Consider a diffusion process (1)

$$d\xi_t = b(\xi_t) dt + \sigma(\xi_t) dW_t, \quad t \geq 0$$

observed at discrete time points

$$(21) \quad t_i := i\delta, \quad i \in \mathbb{N}_0$$

for some $\delta > 0$ (at first, δ of (21) is not related to Δ of (2)). With $(\mathcal{F}_t)_{t \geq 0}$ the filtration generated by the process ξ , we get from

$$(22) \quad b(\xi_t) = \lim_{s \downarrow 0} \frac{1}{s} E(\xi_{t+s} - \xi_t \mid \mathcal{F}_t)$$

$$(23) \quad \sigma^2(\xi_t) = \lim_{s \downarrow 0} \frac{1}{s} E((\xi_{t+s} - \xi_t)^2 \mid \mathcal{F}_t)$$

the usual form of kernel estimates based on the first m observations according to (21)

$$(24) \quad \widehat{\sigma}_{[\delta, m, h]}^2(a) := \frac{\sum_{i=0}^{m-1} K\left(\frac{X_{t_i} - a}{h}\right) \left(\frac{X_{t_{i+1}} - X_{t_i}}{\sqrt{\delta}}\right)^2}{\sum_{i=0}^{m-1} K\left(\frac{X_{t_i} - a}{h}\right)}$$

$$(25) \quad \widehat{b}_{[\delta, m, h]}(a) = \frac{\sum_{i=0}^{m-1} K\left(\frac{X_{t_i} - a}{h}\right) \left(\frac{X_{t_{i+1}} - X_{t_i}}{\delta}\right)}{\sum_{i=0}^{m-1} K\left(\frac{X_{t_i} - a}{h}\right)}$$

with some bandwidth $h > 0$ and some kernel $K(\cdot)$. In contrast to (5)+(6), the estimators (24)+(25) evaluate increments of the process ξ over disjoint time intervals. Two types of asymptotics for observation schemes (21) are of importance:

$$(26) \quad \text{for } \delta > 0 \text{ fixed, let } m \text{ tend to } \infty,$$

$$(27) \quad \delta \downarrow 0 \text{ and } m \uparrow \infty \text{ such that } m\delta \equiv 1 \text{ (say) remains constant,}$$

where (26) requires an additional ergodicity assumption. We think of $b(\cdot)$ and $\sigma^2(\cdot)$ as \mathcal{C}^1 functions, with $\sigma^2(\cdot)$ strictly positive (it is sufficient to have these properties on some subinterval of the state space on which we wish to estimate drift and diffusion coefficient), and use a kernel $K(\cdot)$ of order one (in the sense of [T 04]) with compact support.

For estimation of the diffusion coefficient we need asymptotics (27). The following result, under some smoothness conditions on $b(\cdot)$ and $\sigma(\cdot)$, is [FZ 93, thm. 1]: for shrinking bandwidth

$$(28) \quad h = h(m) = o\left(m^{-\frac{1}{3}}\right)$$

one has asymptotic normality

$$(29) \quad \sqrt{\sum_{i=0}^{m-1} K\left(\frac{X_{t_i} - a}{h}\right)} \left(\widehat{\sigma^2}_{[\delta, m, h]}(a) - \sigma^2(a)\right) \longrightarrow \mathcal{N}(0, \sigma^4(a)) ,$$

for the rectangle kernel. For rate-optimality of this estimator on certain nonparametric function classes see [H 01], [H 99]; some extensions are given in [J 00].

For drift estimation under asymptotics (26), assuming ergodicity of the process $(\xi_t)_{t \geq 0}$ and stationarity with invariant density $f(\cdot)$

$$f(x) = C \frac{1}{\sigma^2(x)} \exp \left\{ \int_0^x \frac{2b}{\sigma^2}(v) dv \right\}$$

($0 < C < \infty$ the norming constant), we take a continuous-time approach (see [K 03] for a large scope of methods and asymptotic optimality results in continuous time), and view the drift estimate $\widehat{b}_{[\delta, m, h]}(a)$ as a discrete approximation along (21) to the continuous-time estimator

$$(30) \quad \bar{b}_{T, h}(a) := \frac{\int_0^T K\left(\frac{\xi_s - a}{h}\right) d\xi_s}{\int_0^T K\left(\frac{\xi_s - a}{h}\right) ds}, \quad T := m\delta .$$

Under the condition

$$(31) \quad T \longrightarrow \infty, \quad h = h(T) = o\left(T^{-\frac{1}{3}}\right),$$

a standard scheme of proof (see [T 04] for the guideline in a classical iid setting; our proof proceeds as in [H-H-L 02, (56) and (60)–(66)] where the context is more general) writes

$$A^{T, h} := \frac{1}{T} \int_0^T \frac{1}{h} K\left(\frac{\xi_s - a}{h}\right) ds, \quad M^{T, h} := \frac{1}{\sqrt{T}} \int_0^T \frac{1}{\sqrt{h}} K\left(\frac{\xi_s - a}{h}\right) \sigma(\xi_s) dW_s$$

and yields

$$A^{T, h} \longrightarrow f(a) \quad (\text{in } L^2), \quad M^{T, h} \longrightarrow \mathcal{N}\left(0, \sigma^2(a) f(a) \int K^2(v) dv\right) \quad (\text{in law})$$

together with a decomposition of suitably rescaled estimation errors

$$\sqrt{T h} A^{T, h} (\bar{b}_{T, h}(a) - b(a)) = M^{T, h} + B^{T, h}, \quad B^{T, h} = o_P(1)$$

for a kernel of order one. The last assertions combine to weak convergence of randomly normed estimation errors

$$(32) \quad \sqrt{\int_0^T K\left(\frac{\xi_s - a}{h}\right) ds} (\bar{b}_{T, h}(a) - b(a)) \longrightarrow \mathcal{N}\left(0, \sigma^2(a) \int K^2(v) dv\right)$$

under condition (31). For rate-optimality of the continuous-time estimator on certain classes of \mathcal{C}^1 functions see [K 03, thm. 4.38]. Assertion (32) for the drift is a continuous-time analogue to (29); for discrete observations along (21), we replace the estimator $\bar{b}_{T,h}(a)$ of (32) by $\widehat{b}_{[\delta,m,h]}(a)$, and the random norming of (32) by

$$\sqrt{\delta \cdot \sum_{i=0}^{m-1} K\left(\frac{X_{t_i} - a}{h}\right)}$$

where δ is constant under asymptotics (26).

From (32) and (29) we see that the accuracy at a of estimators $\widehat{\sigma}_{[\delta,m,h]}^2(a)$ in (24) and $\widehat{b}_{[\delta,m,h]}(a)$ in (25) is measured in the sense of stochastic order by (square root of) the 'time spent near a '

$$(33) \quad \sum_{i=0}^m K\left(\frac{X_{t_i} - a}{h}\right)$$

which plays the role of an 'observed information' available from the data. Under both types of asymptotics, the specific choice of a kernel $K(\cdot)$ of order one plays no role (corresponding to asymptotics (26), by convergence of $A^{T,h}$ in L^2 above, under asymptotics (27), by [FZ 93, prop. 1]), up to a factor in the variance of the limit law.

We exploit this as follows. Fix a positive integer k , subdivide $]t_i, t_{i+1}]$ into k subintervals of length $\frac{1}{k}\delta$, and assume that – strengthening (21) – we have observations of ξ also on the refined time grid with step size $\frac{1}{k}\delta$. Then we might use

$$(34) \quad \widehat{\widehat{\sigma}}_{[k,\delta,m,h]}^2(a) := \frac{\sum_{j=0}^{k(m-1)} K\left(\frac{X_{j\delta/k} - a}{h}\right) \left(\frac{X_{(j\delta/k)+\delta} - X_{j\delta/k}}{\sqrt{\delta}}\right)^2}{\sum_{i=0}^{k(m-1)} K\left(\frac{X_{j\delta/k} - a}{h}\right)}$$

as a variant of (24) under asymptotics (27), with k fixed, and

$$(35) \quad \widehat{\widehat{b}}_{[k,\delta,m,h]}(a) := \frac{\sum_{j=0}^{k(m-1)} K\left(\frac{X_{j\delta/k} - a}{h}\right) \left(\frac{X_{(j\delta/k)+\delta} - X_{j\delta/k}}{\delta}\right)}{\sum_{i=0}^{k(m-1)} K\left(\frac{X_{j\delta/k} - a}{h}\right)}$$

as a variant of (25) under asymptotics (26), with k fixed. Now intervals on which we calculate increments of the process ξ overlap. However, the new estimators are of a type which can be related to the previous ones as follows. Define for $l = 0, 1, \dots, k-1$ time-delayed versions of the estimator (24) for the diffusion coefficient

$$(36) \quad \widehat{\widehat{\sigma}}_{[k,\delta,m,h;l]}^2(a) := \frac{\sum_{i=0}^{m-1} K\left(\frac{X_{(t_i + \frac{l}{k}\delta) - a}}{h}\right) \left(\frac{X_{(t_{i+1} + \frac{l}{k}\delta)} - X_{(t_i + \frac{l}{k}\delta)}}{\sqrt{\delta}}\right)^2}{\sum_{i=0}^{m-1} K\left(\frac{X_{(t_i + \frac{l}{k}\delta) - a}}{h}\right)}$$

then we have a decomposition

$$(37) \quad \widehat{\widehat{\sigma}}^2_{[k,\delta,m,h]}(a) = \sum_{l=0}^{k-1} \gamma_{[k,\delta,m,h;l]}(a) \widehat{\sigma}^2_{[k,\delta,m,h;l]}(a)$$

with (random) coefficients converging in probability (at least)

$$(38) \quad \gamma_{[k,\delta,m,h;l]}(a) = \frac{\sum_{i=0}^{m-1} K \left(\frac{X_{(t_i + \frac{1}{k}\delta)} - a}{h} \right)}{\sum_{j=0}^{k(m-1)} K \left(\frac{X_{j\delta/k} - a}{h} \right)} \longrightarrow c_{[k,\delta,m,h;l]}(a)$$

with a set of limiting constants

$$c_{[k,\delta,m,h;l]}(a) > 0, \quad 0 \leq l < k, \quad \sum_{l=0}^{k-1} c_{[k,\delta,m,h;l]}(a) = 1$$

(under asymptotics (27), use again [FZ 93, prop. 1]; note that k is fixed). From an assertion analogous to (29) for all estimators $\widehat{\sigma}^2_{[k,\delta,m,h;l]}(a)$ and from (37)+(38), we get for the new estimator $\widehat{\widehat{\sigma}}^2_{[k,\delta,m,h]}(a)$ a tightness result

$$(39) \quad \sqrt{\sum_{j=0}^{k(m-1)} K \left(\frac{X_{j\delta/k} - a}{h} \right)} \left(\widehat{\widehat{\sigma}}^2_{[k,\delta,m,h]}(a) - \sigma^2(a) \right) \text{ is tight under (27)+(28).}$$

The quantity under the square root in (39) plays the role of a measure of accuracy – in stochastic order – for the new estimator (34); this measure of accuracy can be read from the data. Similiar arguments apply to the drift estimator (35).

Now we can justify our choice of estimators in section 2. In a real data set as our (2)

$$X_{i\Delta}, \quad i_0 \leq i \leq i_1, \quad i_0 := \lceil \frac{t_0}{\Delta} \rceil, \quad i_1 := \lfloor \frac{t_1}{\Delta} \rfloor, \quad \Delta = 6 \cdot 10^{-4} \text{ [sec]}$$

asymptotics (26) or (27) are never feasible. For our $X_{i\Delta}$, the given accuracy of measurement (formally, 0.001 [mV] – where the last decimal is biologically not reliable) forbids to shrink the bandwidth down to an order of 10^{-3} . So we have to mimick both assumptions (28) and (31) by choosing a bandwidth $h > 0$ and M -step Δ -increments such that both

$$M\Delta =: \delta, \quad \frac{h}{(M\Delta)^{1/3}}$$

are 'small', whence our choice $M := 20$ and $h := 0.01$ in (7) of section 2.

Next, using estimates (24) and (25) with $\delta = M\Delta$ where $M = 20$ in the data set (2) would imply that we use only 5 per cent of our data; instead of an amount of 'visits near a ' of order

100, 300, 1000 (cf. part D of all figures in sections 3 and 4) we could only work with amounts of 5, 15 or 50 which is completely insufficient.

This is why in section 2 we have switched to estimates of form (34) and (35), paying a price in terms of explicit limit laws, in order to have a tractable – in terms of tightness – (random) speed of convergence which can be read from the data.

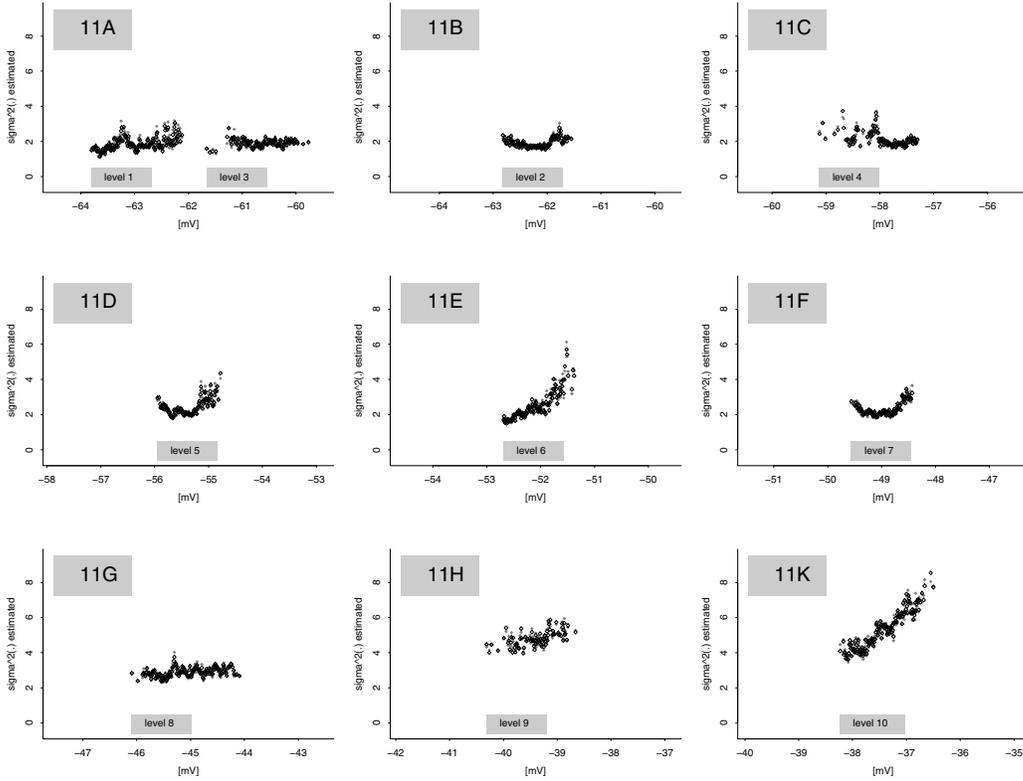


Figure 11: For a -values having 'local time' > 300 , we compare clouds of points $(a, \widehat{\sigma^2}_{(\Delta, M, h)}(a))$ obtained from (5)+(7). All subgraphs 11A – 11K have identical scaling. The different shape of the cloud of points in cases 'level 3' (OU type), 'level 5' and 'level 7' (Q type), 'level 9' and 'level 10' (CIR type) appears clearly. Strong time inhomogeneities are present e.g. in 'level 1'.

6. Conclusion

We have addressed the question of suitable stochastic models for the membrane potential through nonparametric estimates for diffusion coefficient and drift in discretely observed diffusion processes, applied to segments of the membrane potential between successive spikes. The same neuron (one pyramidal neuron in a cortical slice preparation observed in vitro) under different experimental conditions (exposition of the whole slice to different constant levels of potassium, stimulating the networking properties of all neurons in this slice, figure 1) was flexible enough to show different types of behaviour (for the diffusion coefficient, see figure 11). We have encountered examples for models of Ornstein-Uhlenbeck type, Cox-Ingersoll-Ross type, and of a third 'quadratic' type: 'level 8' (10 [mM] of K^+), 'level 9' (12 [mM]) and 'level 10' (15 [mM]) are examples of type CIR; 'level 3' (5 [mM]) is of type OU; 'level 5' (7 [mM]) and 'level 7' (9 [mM]) are of type Q. Strong time inhomogeneities are present in some levels, e.g. 'level 1' (3 [mM]).

References

- [BN-G-J-P-S 06] Barndorff-Nielsen, O., Graversen, S., Jacod, J., Shephard, N.: Limit theorems for bipower variation in financial econometrics. Preprint PMA 973, Laboratoire de Probabilités et Modèles Aléatoires, Université Paris VI, 2005.
- [BN-G-J-P-S 06] Barndorff-Nielsen, O., Graversen, S., Jacod, J., Podolskij, M., Shephard, N.: A central limit theorem for realised power and bipower variations of continuous semimartingales. *From stochastic calculus to mathematical finance* (Shiryaev Festschrift), pp. 33–68, Springer 2006.
- [B-H 06] Brodda, K., Höpfner, R.: A stochastic model and a functional central limit theorem for information processing in large systems of neurons. *J. Math. Biol.* **52**, 439 - 457 (2006)
- [C-I-R 85] Cox, J., Ingersoll, J., Ross, S.: A theory of the term structure of interest rates. *Econometrica* **53**, 385–407 (1985).
- [C-R 02] Comte, F., Rozenholc, Y.: Adaptive estimation of mean and volatility functions in (auto-)regressive models. *Stochast. Proc. Appl.* **97**, 111–145 (2002).
- [D-L 05] Ditlevsen, S., Lánský, P.: Estimation of the input parameters in the Ornstein-Uhlenbeck neuronal model. *Physical review E* **71**, 011907 (2005).

- [D-L 06] Ditlevsen, S., Lánský, P.: Estimation of the input parameters in the Feller neuronal model. *Physical review E* **73**, 061910 (2006).
- [FZ 93] Florens-Zmirou, D.: On estimating the diffusion coefficient from discrete observations. *J. Appl. Probab.* **30**, 790–804 (1993).
- [G-L-N-R 88] Giorno, V., Lánský, P., Nobile, A., Ricciardi, L.: Diffusion approximation and first-passage-time problem for a model neuron. *Biol. Cybern.* **58**, 387–404 (1988).
- [H-H-L 02] Höpfner, R., Hoffmann, M., Löcherbach, E.: Non-parametric estimation of the death rate in branching diffusions. *J. Scand. Statist.* **29**, 665–693 (2002).
- [H 99] Hoffmann, M.: L_p estimation of the diffusion coefficient. *Bernoulli* **5**, 447–481 (1999).
- [H 01] Hoffmann, M.: On estimating the diffusion coefficient: parametric versus nonparametric. *Ann. Inst. H. Poincaré Probab. Statist.* **37**, 39–372 (2001).
- [J 00] Jacod, J.: Non-parametric kernel estimation of the coefficient of a diffusion. *Scand. J. Statist.* **27**, 83–96 (2000).
- [K 03] Kutoyants, Yu.: *Statistical inference for ergodic diffusion processes*. Springer 2003.
- [L-L 87] Lánský, P., Lánská, V.: Diffusion approximation of the neuronal model with synaptic reversal potentials. *Biol. Cybern.* **56**, 19–26 (1987).
- [L-S 01] Lánský, P., Sacerdote, L.: The Ornstein-Uhlenbeck neuronal model with signal-dependent noise. *Physics Letters A* **285**, 132–140 (2001).
- [L-S 99] Lánský, P., Sato, S.: The stochastic diffusion models of nerve membrane depolarization and interspike interval generation. *J. Peripheral Nervous System* **4**, 27–42 (1999).
- [L-S-T 95] Lánský, P., Sacerdote, L., Tomassetti, F.: On the comparison of Feller and Ornstein-Uhlenbeck models for neural activity. *Biol. Cybern.* **73**, 457–465 (1995).
- [P 06] Podolskij, M.: New theory on estimation of integrated volatility. PhD thesis, University of Bochum 2006.
- [T 04] Tsybakov, A.: *Introduction à l'estimation non-paramétrique*. Springer-SMAI 2004.
- [S-S-F 99] Shinomoto, S., Sakai, Y., Funahashi, S.: The Ornstein-Uhlenbeck process does not reproduce spiking statistics of neurons in prefrontal cortex. *Neural Comput.* **11**, 935–951 (1999).
- [T 89] Tuckwell, H.: *Stochastic processes in the neurosciences*. CBMS-NSF conference series in applied mathematics, SIAM 1989.
- [W-T 79] Wan, F., Tuckwell, H.: The response of a spatially distributed neuron for white noise current injection. *Biol. Cybernetics* **33**, 39–55 (1979).