



Optimal bandwidth selection in stochastic regression of Bio-FET measurements

Luis A. Melara¹ · Ryan M. Evans² · Seulki Cho³ · Arvind Balijepalli⁴ · Anthony J. Kearsley² 

Received: 27 September 2021 / Revised: 27 September 2021 / Accepted: 5 March 2024 /

Published online: 10 June 2025

This is a U.S. Government work and not under copyright protection in the US; foreign copyright protection may apply 2025

Abstract

Biological field effect transistors (Bio-FETs) are modern bioelectronics instruments that offer rapid, low-cost and accurate point of care (POC) biomarker measurements. The time series data produced by these devices contain noise which interferes with quantitative analysis. Stochastic regression, which relies on modeling the measurement with a linear stochastic drift-diffusion equation with unknown coefficients, is employed to separate signal from noise. Coefficients are estimated through local weighted regression and maximum likelihood estimation, both of which depend on a kernel function and the size of a bandwidth parameter. In this work we determine the optimal bandwidth parameter associated with an experimental Bio-FET measurement, by considering three distinct but related kernel functions. Cross validation is performed with respect to different instrument aspect ratios. Results show optimal bandwidth parameters are surprisingly consistent across aspect ratios, and suggest a choice of kernel function.

Keywords Stochastic differential equation · Stochastic regression · Biological field effect transistors

Mathematics Subject Classification 60H10 · 62L20 · 62M10 · 92-10

✉ Anthony J. Kearsley
anthony.kearsley@nist.gov

¹ Department of Mathematics, Shippensburg University of Pennsylvania, 1871 Old Main Drive, Shippensburg 17257, PA, USA

² Applied and Computational Mathematics Division, National Institute of Standards and Technology, 100 Bureau Drive, Gaithersburg 20899, MD, USA

³ Chemical and Biomolecular Engineering, University of Maryland, College Park 20742, MD, USA

⁴ Microsystems and Nanotechnology Division, National Institute of Standards and Technology, 100 Bureau Drive, Gaithersburg 20899, MD, USA

1 Introduction

New point of care (POC) medical diagnostic instruments promise to fundamentally improve healthcare delivery by giving physicians an accessible way to measure biomarkers and thus make timely diagnostic decisions. Dire need for such technologies was illustrated by the recent COVID-19 pandemic wherein existing POC diagnostic tests proved less accurate than those based on quantitative polymerase chain reaction (Patrone et al. 2020). The latter requires significant sample preparation, specialized facilities and trained personnel. For this reason, test results are often reported days after being administered. The need for instruments that facilitate quality accessible care has led to an intense interest in their research and development.

Biological field effect transistors (Bio-FETs) are modern bio-electronics instruments that show particular promise, since they are low-cost and portable tools that yield highly precise measurements on a rapid time scale. The dimensions of these azimuthally symmetric instruments are on the order of millimeters. An idealized cross-sectional schematic is shown in Fig. 1. During a typical experiment, chemical reactants injected through a port at the top of the solution-well diffuse through the well to bind with receptors immobilized to the biochemical gate, which together with the semiconductor translate binding of chemical reactants into a signal, an example of which is shown in Fig. 2.

Figure 2 clearly reveals noise that can adversely affects quantitative analysis. Stochastic regression is a denoising approach that has successfully been applied to mass spectrometry (Kearsley et al. 2014) where a signal is modeled with a linear stochastic drift-diffusion equation that contains unknown coefficients. These coefficients are estimated through local weighted regression and maximum likelihood estimation, both of which depend on the local averaging function and the size of the local averaging window. Selection of the local averaging function, referred to as the *kernel function*, and the local averaging window size parameter or *bandwidth parameter* control the performance of the method. In this paper the optimal bandwidth parameter associated with an experimental Bio-FET measurement is calculated by considering three related kernel functions. We begin in Sect. 2 by describing our drift-diffusion model and outlining our regression procedure. In Sect. 3 three optimality measures are considered and employed to optimally reconstruct signal from Bio-FET measurement data. Cross validation of optimal bandwidths is described in Sect. 4 with different synthetic measurements corresponding to different aspect ratios ϵ , the ratio of the solution-well's height to the biochemical gate diameter. Concluding remarks will be given in Sect. 5.

2 Regression Model

Our drift-diffusion model for the signal X_t takes the form:

$$dX_t = M(t, X_t) dt + \Sigma(t, X_t) dW_t, \quad (2.1a)$$

$$X_t(0) = X_0, \quad (2.1b)$$

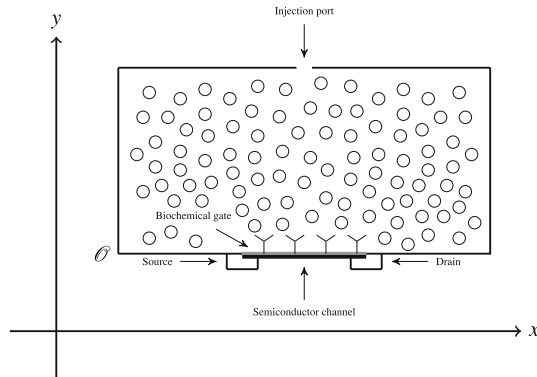
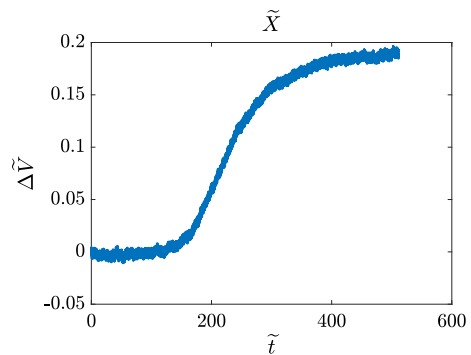


Fig. 1 Idealized cross-sectional schematic of a Bio-FET experiment. Chemical reactants injected at the top of the instrument diffuse to bind with receptors confined to the biochemical gate. In our schematic, chemical reactants in solution are represented as circles spread throughout the instrument, and receptor sites are depicted with Y's on the biochemical gate. This schematic is not to scale. The orientation of the x and y axes have been shown on the left, along with the location of the origin $\mathcal{O}$

Fig. 2 Experimental Bio-FET signal. Streptavidin injected into the solution-well of a Bio-FET has diffused through the solution-well onto the biochemical gate to bind with immobilized biotin, creating the displayed signal. More information can be found in (Evans et al. 2020)



where $M(t, X_t)$ and $\Sigma(t, X_t)$ are respectively the drift and diffusion coefficients, and W_t is a standard Brownian motion term with independent and normally-distributed increments such that

$$\mathbb{E}[dW_t] = 0, \quad (2.2a)$$

$$\text{Var}[dW_t] = dt. \quad (2.2b)$$

In this work we assume that (2.1) holds on a finite interval $[0, T]$, and from (2.1) it follows that we are also assuming that noise is additive and normally-distributed; no assumptions are made on the autocorrelation. Each of the independent and dependent variables in (2.1) are dimensionless, and may readily be cast in dimensional form by

introducing unit characteristic scales for time $\tilde{t}_c$ and the signal $\tilde{X}_c$ and setting

$$\tilde{t} := \tilde{t}_c t, \quad (2.3a)$$

$$\tilde{X}_{\tilde{t}} := \tilde{X}_c X_t. \quad (2.3b)$$

We work with the dimensionless formulation (2.1) and convert to dimensional variables only when discussing solutions of (2.1), for ease of comparison with physical experiments.

Our model (2.1) is a drift-diffusion equation requiring the drift and diffusion coefficients to be specified. We take the drift term to be a linear function of X_t

$$M(t, X_t) = a_0(t) + a_1(t)X_t, \quad (2.4)$$

where $a_0(t)$ and $a_1(t)$ are functions to be determined. There are many possible forms for the diffusion coefficient, among them are

$$\sigma(t, X_t) = b(t), \quad (2.5a)$$

$$\sigma(t, X_t) = b(t)X_t, \quad (2.5b)$$

$$\sigma(t, X_t) = b_0(t)X_t^{b_1(t)}, \quad (2.5c)$$

where $b(t)$, $b_0(t)$, and $b_1(t)$ are functions to be determined. We shall comment on each of the choices in (2.5); first considering (2.5b), we observe that this choice renders (2.1) singular if the data are continuous and not strictly positive or negative. Similar problems may arise with (2.5c) if $\lim_{t \rightarrow t_i} X_t = 0$ and $b_1(t)$ does not approach zero at the appropriate rate. Furthermore (2.5c) will result in non-real coefficients if for some t_i during $[0, T]$ the datum X_{t_i} is negative and $b_1(t_i)$ assumes a non-integral value. These difficulties could ostensibly be overcome by mapping the data X_{t_i} to $X_{t_i} + c$ for some constant c , but we nevertheless prefer the choice (2.5a) for which there are no restrictions.

Now if we had the exact coefficients we could partition the interval $[0, T]$ into M subintervals with a set of nodes $t_0 < t_1 < \dots < t_M$ and apply an Euler-Maruyama discretization of the form

$$\Delta X_i = [a_0(t_i) + a_1(t_i)]\Delta t + b(i)\Delta W_i, \quad (2.6a)$$

$$X_0 = X_t|_{t=0}, \quad (2.6b)$$

for $i = 0, \dots, M-1$. In (2.6) we have let $\Delta t = t_{i+1} - t_i$, $\Delta X_i = X_{i+1} - X_i$, and $\Delta W_i = W_{i+1} - W_i$, where X_i and W_i are discrete approximations to their continuous counterparts X_{t_i} and W_{t_i} . From the properties in (2.2) it follows that $\Delta W_i \sim \mathcal{N}(0, \Delta t)$. Formula (2.6) suggests that we can recover the true signal through the difference equation

$$\Delta X_i^{(d)} = [a_0(t_i) + a_1(t_i)]\Delta t_i, \quad (2.7a)$$

$$X_0^{(d)} = X_t|_{t=0}, \quad (2.7b)$$

for $i = 0, \dots, M-1$. In (2.7) we have used $X^{(d)}$ to denote the deterministic component of the signal. We could recover the stochastic component through the formula

$$X_i^{(w)} = b(t_i) \Delta W_i, \quad (2.8)$$

for $i = 1, \dots, M$, since our initial condition is exact.

Although we do not have the exact form of $a_0(t)$, $a_1(t)$, or $b(t)$, following (Kearsley et al. 2014) we can estimate $a_0(t)$ and $a_1(t)$ through local weighted regression and $b(t)$ through maximum-likelihood estimation. To this end we assume that the drift and diffusion coefficients are well-approximated by constants over a set of narrow windows

$$(t_j - h, t_j + h), \quad (2.9)$$

where each one of the nodes t_j is chosen from the set $\{t_i\}_{i=1}^M$. The parameter h in (2.9) is the *bandwidth parameter*, and in this work we take it to be an integral multiple k of the time step

$$h = \Delta t k. \quad (2.10)$$

We note that we may not be able to evenly partition the interval $[0, T]$ into subintervals of the form (2.9); suppose this is the case, that we cannot fully partition the interval $[0, T]$ into subintervals of the form (2.9) and that there are l such subintervals in total. From (2.9) and (2.10) it follows that $(t_k - h, t_k + h)$ will be the first subinterval to lie in the interior of $[0, T]$, and that the remaining $l - 1$ will take the form

$$(t_{k+j(2k)} - h, t_{k+j(2k)}) \quad (2.11)$$

for $j = 1, \dots, l-1$. There will also be points in a neighborhood of the right boundary of $[0, T]$ that do not fall within any of the aforementioned l subintervals, and it is the case that the procedure described below for estimating the drift and diffusion coefficients applies only to nodes that lie within one of the l subintervals of the form (2.11) for $j = 0, \dots, l-1$. To estimate drift and diffusion coefficients for points near the right boundary that do not fall within any of these l subintervals, we linearly extrapolate the data to r fictitious points, where r is the least positive integer such that $2k$ divides $M + r$. Then we apply the procedure described below to estimate the drift and diffusion coefficients in that last extended subinterval, and discard the fictitious points when finished, since they are unnecessary for time stepping. This extrapolation procedure is justified for the signals herein since data points (t_i, X_{t_i}) for which t_i is close to the terminal time T are in a steady regime. In general, our extrapolation procedure is expected to work for signals that have the property that data points (t_i, X_{t_i}) for which t_i is close T exhibit linear behavior.

Turning our attention to the drift coefficients, we take their approximation to be the solution to the local weighted regression problem

$$\{\hat{a}_0(i), \hat{a}_1(i)\} = \operatorname{argmin}_{a_0(t_i), a_1(t_i)} \sum_{j=0}^{M-1} \left(\frac{\Delta X_j}{\Delta t} - a_0(t_i) - a_1(t_i) X_j \right)^2 K_h \left(\frac{t_j - t_i}{h} \right). \quad (2.12)$$

In (2.12) we have let $\hat{a}_0(i)$ and $\hat{a}_1(i)$ denote approximations to the their continuous counterparts $a_0(t_i)$ and $a_1(t_i)$. Here $K_h(\cdot)$ is the *kernel function*, and determines the weighting. In this work we shall consider three, the first of which is the classic *Epanechnikov kernel* Epanechnikov (1969)

$$K_h(u) = \begin{cases} \frac{3}{4h}(1 - u^2), & -1 \leq u < 0, \\ 0, & 0 \leq u. \end{cases} \quad (2.13a)$$

Since this kernel is supported on the half-interval $(t_j - h, t_j)$ it is a *left-sided* kernel, and amounts to using only historical information. We shall also consider *right-sided*

$$K_h(u) = \begin{cases} \frac{3}{4h}(1 - u^2), & 0 \leq u < 1, \\ 0, & \text{otherwise.} \end{cases} \quad (2.13b)$$

and *centered*

$$K_h(u) = \begin{cases} \frac{3}{4h}(1 - u^2), & -1 \leq u \leq 1, \\ 0, & \text{otherwise.} \end{cases} \quad (2.13c)$$

versions of (2.13a). Then the solution to (2.12) is found by taking the derivative of the right-hand side of (2.12) with respect to $a_0(i)$ and $a_1(i)$, setting these expressions equal to zero, and solving the resulting system to obtain

$$\hat{a}_0(i) = \frac{\sum_{j=1}^N (\Delta X_j / \Delta t - \hat{a}_1(i) X_j) K_h(i, j)}{S(i)}, \quad (2.14a)$$

$$\hat{a}_1(i) = \frac{S \sum_{j=1}^N \frac{\Delta X_j}{\Delta t} X_j K_h(i, j) - \left[\sum_{j=1}^N \frac{X_j}{\Delta t} K_h(i, j) \right] \left(\sum_{j=1}^N K_h(i, j) X_j \right)}{S(i) \sum_{j=1}^N X_j^2 K_h(i, j) - \left(\sum_{j=1}^N X_j K_h(i, j) \right)^2}. \quad (2.14b)$$

In (2.14) we have introduced the notation

$$K_h(i, j) = K_h\left(\frac{t_i - t_j}{h}\right), \quad (2.15a)$$

$$S(i) = \sum_{j=1}^N K_h\left(\frac{t_i - t_j}{h}\right). \quad (2.15b)$$

Equation (2.20) gives us an estimate for the drift coefficients, so we now turn our attention to estimating the diffusion coefficient. We begin by observing that equation (2.1a) leads to the following model for the residual

$$\Delta X_i - [\hat{a}_0(i) + \hat{a}_1(i)]\Delta t \approx b(t_i)\Delta W_i. \quad (2.16)$$

Next we normalize the left-hand side of (2.16) and introduce the variable

$$\hat{E}_i := \frac{\Delta X_i - [\hat{a}_0(i) + \hat{a}_1(i)]\Delta t}{\sqrt{\Delta t}}. \quad (2.17)$$

From (2.16) and the fact that ΔW_i has zero mean and variance $\sqrt{\Delta t}$, it follows that $\hat{E}_i \sim \mathcal{N}(0, b^2(t_i))$ and that its conditional density given information up to time t_i is

$$[2\pi b^2(i)]^{-1/2} \exp\left(-\frac{\hat{E}_i^2}{2b(t_i)^2}\right). \quad (2.18)$$

Equation (2.18) suggests the local weighted log-likelihood function

$$L(b(t_i), t_i) = -\frac{1}{2} \sum_{j=1}^N K_h(i, j) \left(\log b^2(t_i) + \frac{\hat{E}_j^2}{b(i)^2} \right), \quad (2.19)$$

and taking $\hat{b}(i)$ to be the argument maximum of (2.19) we find

$$\hat{b}(i) = \sqrt{\frac{\sum_{j=1}^N K_h(i, j) \hat{E}_j^2}{\sum_{j=1}^N S(j)}}. \quad (2.20)$$

With the estimated drift (2.14) and diffusion (2.20) coefficients, the estimated reconstruction Y_i to the original signal X_{t_i} is given by

$$\Delta Y_i = [\hat{a}_0(i) + \hat{a}_1(i)]\Delta t_i + \hat{b}(i)\Delta W_i, \quad (2.21a)$$

$$Y_0 = X_t|_{t=0}. \quad (2.21b)$$

The deterministic component of the estimated signal is then given by

$$\Delta Y_i^{(d)} = [\widehat{a}_0(i) + \widehat{a}_1(i)]\Delta t, \quad (2.22a)$$

$$Y_0^{(d)} = X_t|_{t=0}, \quad (2.22b)$$

for $i = 0, \dots, M - 1$, and the stochastic component through the formula

$$\Delta Y_i^{(w)} = \widehat{b}(i)\Delta W_i, \quad (2.23)$$

for $i = 1, \dots, M$. In writing (2.22) and (2.23) we have let $Y^{(d)}$ and $Y^{(w)}$ denote the deterministic and stochastic components of the estimated signal respectively. So given a measurement X_t from a Bio-FET experiment, we can reconstruct it through the formula (2.21) and extract its deterministic (2.22) and stochastic (2.23) components. The estimated deterministic component (2.22) gives us an approximation to what in the original measurement corresponds to true signal, and (2.23) gives us an approximation to what in the original measurement corresponds to noise. These approximations depend on the estimated coefficients (2.14) and (2.20), which in turn depend upon the kernel function and the bandwidth parameter. It is clear that some choices of the bandwidth parameter will result in more accurate reconstructions of the signal than others. If the bandwidth parameter is taken to be too large, then our approximation to the coefficients (2.14), (2.20) will be poor and so will the reconstruction Y_i . If the bandwidth parameter is taken to be too small, then our approximation to the coefficients will (2.14), (2.20) also be poor and so will the reconstruction Y_i .

3 Optimal bandwidth for a Bio-FET measurement

To find the optimal bandwidth parameter, we need to introduce a notion of optimality. In the present work we consider three, the first of which is the root-mean-square error (RMSE)

$$e_{RMSE} = \sqrt{\frac{1}{M+1} \sum_{i=0}^M A_i^2}, \quad (3.1)$$

where

$$A_i = |X_{t_i} - Y_i(h)|. \quad (3.2)$$

We will also consider the mean average error (MAE)

$$\mu = \frac{1}{M+1} \sum_{i=0}^M A_i, \quad (3.3)$$

and the standard deviation (STD_X)

$$\sigma = \sqrt{\frac{1}{M} \sum_{i=0}^M (A_i - \mu)^2}. \quad (3.4)$$

For each of the signals considered herein the time step Δt is fixed, and k is a free parameter that may be chosen at our liberty when applying stochastic regression. Since for each signal Δt is fixed and k is free, the bandwidth parameter (2.10) is a function of k . So to optimize (3.1)–(3.4) we take of these optimality measures to be functions of k , and we also report their respective argument minimums in terms of k^* rather than $h^* = \Delta t k^*$.

To optimize (3.1)–(3.4), an exhaustive search was performed over all integral values of k in the range

$$1 < k < \lfloor M/2 \rfloor. \quad (3.5)$$

In many cases, physically reasonable behavior of our stochastic regression algorithm may require a bandwidth parameter that is not too small. For the bandwidth parameter can be thought of as the degree of smoothing that is applied to the signal, and many signals may require a minimum amount of smoothing before physically reasonable results are produced. Signals such as the one shown in Fig. 2 may have too much variation within the window (2.9) to obtain a reasonable approximation for the coefficients $a_0(t)$, $a_1(t)$, and $b(t)$ when k is too small. On the other hand k cannot be too large, for this would violate our assumption that the coefficients are well-approximated by constants over the window (2.9). So we formally restricted ourselves to values of k within (3.5) that were large enough to produce a physically reasonable approximation to the signal, and small enough to not violate our assumption that the coefficients are well-approximated by constants over the window (2.9). This choice was confirmed by the fact that, for all kernel and optimality measure pairings, the global minimum over (3.5) was consistently and uniformly within our restriction.

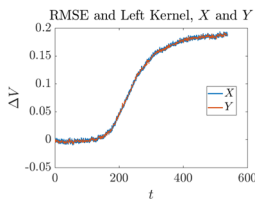
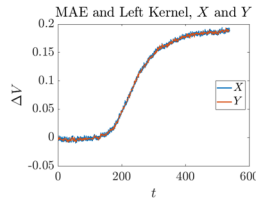
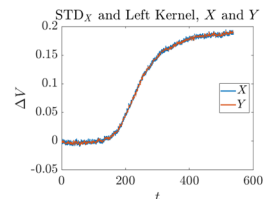
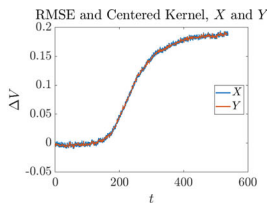
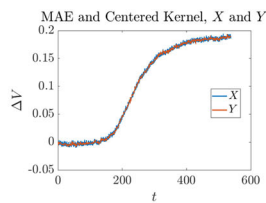
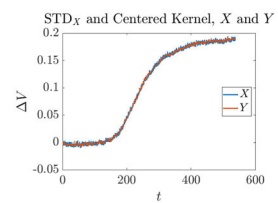
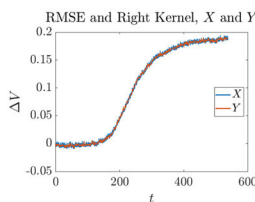
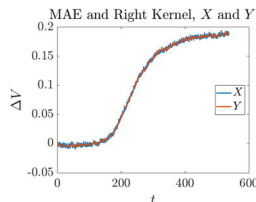
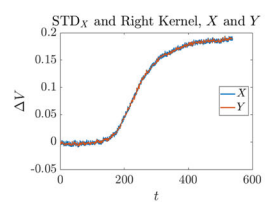
This procedure was applied to find optimal reconstructions of the signal shown in Fig. 2, and the associated optimal bandwidth values. The results are tabulated in Table 1, and shown in Figs. 3–5. From the latter, it is clear that all of the pairings between each kernel and each optimality measure produce signals that agree very well with the original signal X . The data in Table 1 shows that the curves match up quantitatively as well as qualitatively. We see that for each of the optimality measures, the centered kernel performs the best. This is to be expected, since centered kernels use information from both sides of the window (2.9), while its left and right-sided counterparts use information from only one side.

4 Cross validation

Although the signal studied in Sect. 3 came from a single instrument, in practice it is not uncommon to perform many experiments with instruments that have different

Table 1 Optimal bandwidth parameters and associated objective function values for the signal shown in Figs. 3 to 5

Kernel	RMSE		MAE		STD _X	
	k^*	(3.1)	k^*	(3.3)	k^*	(3.4)
Left Kernel	12	1.7611×10^{-3}	12	1.3467×10^{-3}	12	1.1348×10^{-3}
Centered Kernel	4	1.7190×10^{-3}	4	1.3038×10^{-3}	4	1.1204×10^{-3}
Right Kernel	7	1.7686	6	1.3259×10^{-3}	12	1.1705×10^{-3}

(a) An overlay of X and Y , for the left-sided kernel with the RMSE (3.1).(b) An overlay of X and Y , for the left-sided kernel with the MAE (3.3).(c) An overlay of X and Y , for the left-sided kernel with the STD_X (3.4).**Fig. 3** Overlay of the signal X and its optimal reconstructions Y for the left-sided kernel, for each of the objectives (3.1)–(3.4)(a) An overlay of X and Y , for the centered kernel with the RMSE (3.1).(b) An overlay of X and Y , for the centered kernel with the MAE (3.3).(c) An overlay of X and Y , for the centered kernel with the STD_X (3.4).**Fig. 4** Overlay of the signal X and its optimal reconstructions Y for the centered kernel, for each of the objectives (3.1)–(3.4)(a) An overlay of X and Y , for the right-sided kernel with the RMSE (3.1).(b) An overlay of X and Y , for the right-sided kernel with the MAE (3.3).(c) An overlay of X and Y , for the right-sided kernel with the STD_X (3.4).**Fig. 5** Overlay of the signal X and its optimal reconstructions Y for the right-sided kernel, for each of the objectives (3.1)–(3.4)

dimensions. One important way in which these instruments may differ from each other is with respect to their *aspect ratio* ϵ , the ratio of the solution-well's height to its diameter. To examine the performance of our algorithm when applied to instruments with different aspect ratios, cross validation was conducted by first selecting ten numbers in a uniformly random manner between the minimum $\epsilon_{\min} = 0.1250$ and maximum $\epsilon_{\max} = 5$ aspect ratios, shown in Table 2. Then using the physical model presented in Evans et al. (2020) ten curves were created, one for each of the aspect ratios. Normally distributed noise, with mean and variance equal to that of the estimated stochastic component of the signal analyzed in Sect. 3, was added to each of the curves. This resulted in ten synthetic measurements, each corresponding to an instrument with a different aspect ratio. By using the model presented in Evans et al. (2020) to create these synthetic measurements, the time and expense of creating ten different instruments and performing a different experiment with each of them was obviated.

Cross validation on these measurements was then performed in the following manner. Let us fix our attention on a single measurement corresponding to one of the ϵ_i , given in Table 2. For each of the kernels given in (2.5), optimal values of (3.1), (3.3), and (3.4) were found as described in Sect. 3. For each optimality measure this resulted in an optimal value k_i^* , for every kernel. Then for each of the values of k_i^* , the average of all the *other* aspect ratios was calculated

$$\bar{k}_i = \frac{1}{9} \sum_{\substack{j=1, \\ j \neq i}}^{10} k_j^*. \quad (4.1)$$

and the absolute

$$|k_i^* - \bar{k}_i| \quad (4.2)$$

and relative

$$\frac{|k_i^* - \bar{k}_i|}{k_i^*} \quad (4.3)$$

differences were measured. The results for the centered kernel are given in Tables 3 to 5, the results for the left-sided kernel are given in Tables 6 to 8, and the results for the right-sided kernel are given in Tables 9 to 11. Illustrative overlays of synthetic measurements and the curves obtained by applying stochastic regression are shown in Figs. 6 to 8. The curves depicted are those for which the minimum error measure was achieved, for a fixed kernel. So for example, since ϵ_9 corresponds to the instrument with the smallest objective function value in Table 3, this is the curve that has been shown in Fig. 6a. When inspecting these figures one may notice that some curves reach steady state faster than others; for example, in Figs. 6a and 6b the terminal times are approximately $T = 7000$ and $T = 6000$ respectively. This is because instruments that have different aspect ratios differ in the time taken to achieve equilibrium. In general,

Table 2 Aspect ratios used for cross validation

i	1	2	3	4	5	6	7	8	9	10
ε_i	4.0968	4.5407	0.74406	4.5777	3.2078	0.60051	1.4827	2.7910	4.7928	4.8288

Table 3 Cross validation of stochastic regression with respect to aspect ratios, when using a centered kernel and the objective (3.1)

Aspect Ratio	Objective (3.1)		Cross Validation		
	k_i^*	$e\text{RMSE}$	$\bar{k}_i$	$ \bar{k}_i - k_i^* $	$\frac{ \bar{k}_i - k_i^* }{k_i^*}$
4.0968	4	1.6426×10^{-3}	$4.\bar{5}$	$0.\bar{5}$	$0.13\bar{8}$
4.5407	4	1.6741×10^{-3}	$4.\bar{5}$	$0.\bar{5}$	$0.13\bar{8}$
0.74406	4	1.6693×10^{-3}	$4.\bar{5}$	$0.\bar{5}$	$0.13\bar{8}$
4.5777	4	1.6578×10^{-3}	$4.\bar{5}$	$0.\bar{5}$	$0.13\bar{8}$
3.2078	4	1.6609×10^{-3}	$4.\bar{5}$	$0.\bar{5}$	$0.13\bar{8}$
0.60051	4	1.6612×10^{-3}	$4.\bar{5}$	$0.\bar{5}$	$0.13\bar{8}$
1.4827	6	1.6949×10^{-3}	$4.\bar{3}$	$1.\bar{6}$	$0.2\bar{7}$
2.7910	6	1.7784×10^{-3}	$4.\bar{3}$	$1.\bar{6}$	$0.2\bar{7}$
4.7928	4	1.6292×10^{-3}	$4.\bar{5}$	$0.\bar{5}$	$0.13\bar{8}$
4.8288	5	1.7217×10^{-3}	$4.\bar{4}$	$0.\bar{5}$	$0.\bar{1}$

Table 4 Cross validation of stochastic regression with respect to aspect ratios, when using a centered kernel and the objective (3.3)

Aspect Ratio	Objective (3.3)		Cross Validation		
	k_i^*	μ	$\bar{k}_i$	$ \bar{k}_i - k_i^* $	$\frac{ \bar{k}_i - k_i^* }{k_i^*}$
4.0968	3	1.2461×10^{-3}	$4.\bar{1}$	$1.\bar{1}$	$0.\bar{370}$
4.5407	4	1.2884×10^{-3}	4	0	0
0.74406	4	1.2872×10^{-3}	4	0	0
4.5777	4	1.2872×10^{-3}	4	0	0
3.2078	4	1.2734×10^{-3}	4	0	0
0.60051	4	1.2601×10^{-3}	4	0	0
1.4827	4	1.3099×10^{-3}	4	0	0
2.7910	3	1.3575×10^{-3}	$4.\bar{1}$	$1.\bar{1}$	$0.\bar{370}$
4.7928	4	1.2648×10^{-3}	4	0	0
4.8288	6	1.3314×10^{-3}	$3.\bar{7}$	$2.\bar{2}$	$0.\bar{370}$

those with smaller values of ϵ tend to reach steady state quicker because the initial injection of chemical reactant is physically closer to the biochemical gate, which is exactly what we see in Figs. 6a and 6b. The latter corresponds to an instrument with a smaller aspect ratio than the former, and so reaches equilibrium faster.

A consistent theme across all kernels and all optimality measures is that the optimal value k_i^* is within a factor two to three of the predicted value $\bar{k}_i$. This shows that there is not large variation of k_i^* between instruments with different aspect ratios, so that k_i^* for one instrument will be very close to that of another. There are exceptions; for

Table 5 Cross validation of stochastic regression with respect to aspect ratios, when using a centered kernel and the objective (3.4)

Aspect Ratio	Objective (3.4)		Cross Validation		
	k_i^*	σ	$\bar{k}_i$	$ \bar{k}_i - k_i^* $	$\frac{ \bar{k}_i - k_i^* }{k_i^*}$
4.0968	4	1.0485×10^{-3}	$5.\bar{2}$	$1.\bar{2}$	$0.30\bar{5}$
4.5407	5	1.0578×10^{-3}	$5.\bar{1}$	$0.\bar{1}$	$0.0\bar{2}$
0.74406	4	1.0631×10^{-3}	$5.\bar{2}$	$1.\bar{2}$	$0.30\bar{5}$
4.5777	7	1.0399×10^{-3}	$4.\bar{8}$	$2.\bar{1}$	0.30159
3.2078	5	1.0661×10^{-3}	$5.\bar{1}$	$0.\bar{1}$	$0.0\bar{2}$
0.60051	5	1.0413×10^{-3}	$5.\bar{1}$	$0.\bar{1}$	$0.0\bar{2}$
1.4827	6	1.0731×10^{-3}	5	1	$0.1\bar{6}$
2.7910	6	1.0976×10^{-3}	5	1	$0.1\bar{6}$
4.7928	4	1.0271×10^{-3}	$5.\bar{2}$	$1.\bar{2}$	$0.30\bar{5}$
4.8288	5	1.0755×10^{-3}	$5.\bar{1}$	$0.\bar{1}$	$0.0\bar{2}$

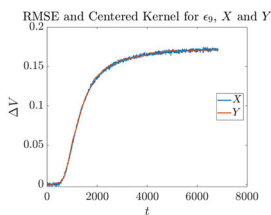
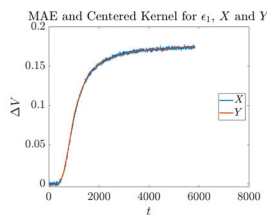
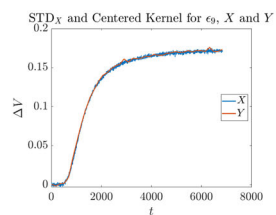
example, the results for the left-sided kernel show that $|\bar{k}_i - k_i^*|$ is typically around six for the aspect ratio $\epsilon_6 = 0.60051$. An examination of the results shown in Tables 3 to 11 reveals that this is indeed an exception, and that our stochastic regression algorithm enjoys remarkable consistency across Bio-FETs with different aspect ratios.

The results in Tables 3 to 11 may be used to guide the choice of kernel in our algorithm. For all aspect ratios and each optimality measure, the centered kernel always has a value of k_i^* less than or equal to that of either its left or right-sided counterparts. Smaller values of k_i^* correspond to smaller averaging windows, which increases the accuracy with which the coefficients $\hat{a}_0(i)$, $\hat{a}_1(i)$, and $\hat{b}(i)$ are estimated. Moreover, for every optimality measure, the centered kernel always has the smallest objective function value. This suggests that using a centered kernel may recover the true signal with the greatest accuracy.

It is also the case that the centered kernel is no less consistent than the others, when consistency is measured in terms of the absolute difference $|k_i^* - \bar{k}_i|$. This is the appropriate measure of consistency, rather than the relative difference $|k_i^* - \bar{k}_i|/k_i^*$. To see why this is the case, let us consider the absolute and relative differences corresponding to $\epsilon_3 = 0.74406$ for the STD_X , in the case of both the centered kernel and the left-sided kernel. The absolute difference for the former is $|5.\bar{2} - 4| = 1.\bar{2}$ while the absolute difference for the latter is $|9.\bar{4} - 8| = 1.\bar{4}$, suggesting that the centered kernel is superior in this case. However the relative difference for the centered kernel is $|5.\bar{2} - 4|/4 = 0.30\bar{5}$ while the relative difference for the left-sided kernel is $|9.\bar{4} - 8|/8 = 0.180\bar{5}$, suggesting that left-sided kernel is superior. The slightly larger magnitude of k_3^* for the left-sided kernel tips the scales in its favor when considering the relative difference. This is misleading, since in this context k_3^* for the left-sided kernel is not significantly larger than k_3^* for the centered kernel. For this reason the absolute difference is the appropriate measure of performance. For the RMSE, since

Table 6 Cross validation of stochastic regression with respect to aspect ratios, when using a left-sided kernel and the objective (3.1)

Aspect Ratio	Objective (3.1)		Cross Validation		
	k_i^*	$e\text{RMSE}$	$\bar{k}_i$	$ \bar{k}_i - k_i^* $	$\frac{ \bar{k}_i - k_i^* }{k_i^*}$
ε_i					
4.0968	10	2.0503×10^{-3}	$9.\bar{2}$	$0.\bar{7}$	$0.0\bar{7}$
4.5407	8	1.9183×10^{-3}	$9.\bar{4}$	$1.\bar{4}$	$0.180\bar{5}$
0.74406	8	1.9474×10^{-3}	$9.\bar{4}$	$1.\bar{4}$	$0.180\bar{5}$
4.5777	9	1.9504×10^{-3}	$9.\bar{3}$	$0.\bar{3}$	0.370
3.2078	10	2.0022×10^{-3}	$9.\bar{2}$	$0.\bar{7}$	$0.0\bar{7}$
0.60051	15	2.2262×10^{-3}	$8.\bar{6}$	$6.\bar{3}$	$0.4\bar{2}$
1.4827	10	2.2117×10^{-3}	$9.\bar{2}$	$0.\bar{7}$	$0.0\bar{7}$
2.7910	7	1.9821×10^{-3}	$9.\bar{5}$	$2.\bar{5}$	0.36508
4.7928	9	1.8670×10^{-3}	$9.\bar{3}$	$0.\bar{3}$	0.0370
4.8288	7	1.8151×10^{-3}	$9.\bar{5}$	$2.\bar{5}$	0.365079

(a) The aspect ratio $\varepsilon_9 = 4.7928$ is shown, which corresponds to $k_9^* = 4$.(b) The aspect ratio $\varepsilon_1 = 4.0968$ is shown, which corresponds to $k_1^* = 3$.(c) The aspect ratio $\varepsilon_9 = 4.7928$ is shown, which corresponds to $k_9^* = 4$.**Fig. 6** Three graphs corresponding to the optimal values of (3.1), (3.3), and (3.4), in the case of the centered kernel. In each graph there is an overlay of a synthetic measurement X and its approximation Y

there are seven aspect ratios for which the absolute difference of the centered kernel is less than that of the right-sided kernel, we can say that the centered kernel performs better than the left-sided kernel seven times and is superior when measured in the RMSE. The centered kernel performs better than the right-sided kernel five times, so the two are equal in the RMSE. Similar analyses can be given for the MAE and standard deviation. In the case of the former, the centered outperforms the left-sided eight times and the right-sided seven times. It is worth noting that for the MAE, *the relative difference of the centered kernel is zero for seven of the ten aspect ratios*. With respect to standard deviation, the centered kernel outperforms the left-sided kernel six times and the right-sided kernel five times. So there is no case in which the centered kernel performs worse, and it often performs better. In this sense the centered kernel is no less consistent than the others and it is often more so. It is for these reasons that we prefer the centered kernel over the others.

Table 7 Cross validation of stochastic regression with respect to aspect ratios, when using a left-sided kernel and the objective (3.3)

Aspect Ratio	Objective (3.3)		Cross Validation		
	k_i^*	μ	$\bar{k}_i$	$ \bar{k}_i - k_i^* $	$\frac{ \bar{k}_i - k_i^* }{k_i^*}$
ε_i					
4.0968	10	1.5403×10^{-3}	9	1	0.1
4.5407	8	1.4915×10^{-3}	$9.\bar{2}$	$1.\bar{2}$	$0.152\bar{7}$
0.74406	8	1.4491×10^{-3}	$9.\bar{2}$	$1.\bar{2}$	$0.152\bar{7}$
4.5777	9	1.5026×10^{-3}	$9.\bar{1}$	$0.\bar{1}$	0.012346
3.2078	10	1.4955×10^{-3}	9	1	0.1
0.60051	15	1.5832×10^{-3}	$8.\bar{4}$	$6.\bar{5}$	$0.437\bar{0}$
1.4827	8	1.5315×10^{-3}	$9.\bar{2}$	$1.\bar{2}$	$0.152\bar{7}$
2.7910	7	1.4773×10^{-3}	$9.\bar{3}$	$2.\bar{3}$	$0.\bar{3}$
4.7928	9	1.4440×10^{-3}	$9.\bar{1}$	$0.\bar{1}$	0.012346
4.8288	7	1.3931×10^{-3}	$9.\bar{3}$	$2.\bar{3}$	$0.\bar{3}$

Table 8 Cross validation of stochastic regression with respect to aspect ratios, when using a left-sided kernel and the objective (3.4)

Aspect Ratio	Objective (3.4)		Cross Validation		
	k_i^*	σ	$\bar{k}_i$	$ \bar{k}_i - k_i^* $	$\frac{ \bar{k}_i - k_i^* }{k_i^*}$
ε_i					
4.0968	10	1.3536×10^{-3}	$9.\bar{2}$	$0.\bar{7}$	$0.0\bar{7}$
4.5407	8	1.2067×10^{-3}	$9.\bar{4}$	$1.\bar{4}$	$0.180\bar{5}$
0.74406	8	1.3013×10^{-3}	$9.\bar{4}$	$1.\bar{4}$	$0.180\bar{5}$
4.5777	9	1.2437×10^{-3}	$9.\bar{3}$	$0.\bar{3}$	$0.0\bar{3}\bar{7}$
3.2078	10	1.3316×10^{-3}	$9.\bar{2}$	$0.\bar{7}$	$0.0\bar{7}$
0.60051	15	1.5664×10^{-3}	$8.\bar{6}$	$6.\bar{3}$	$0.4\bar{2}$
1.4827	10	1.5169×10^{-3}	$9.\bar{2}$	$0.\bar{7}$	$0.0\bar{7}$
2.7910	7	1.3218×10^{-3}	$9.\bar{5}$	$2.\bar{5}$	0.36508
4.7928	9	1.1837×10^{-3}	$9.\bar{3}$	$0.\bar{3}$	$0.0\bar{3}\bar{7}$
4.8288	7	1.1638×10^{-3}	$9.\bar{5}$	$2.\bar{5}$	0.36508

5 Concluding remarks

Herein we have found the optimal bandwidth parameter for stochastic regression denoising of an experimental Bio-FET measurement. This was done by performing an exhaustive search for the minimum values of (3.1), (3.3), and (3.4) over an appropriate range of values for k . By pairing these different optimality measures with different kernels, we found optimal decompositions of the experimental measurement into signal and noise. The centered kernel performed the best across all optimality measures.

Table 9 Cross validation of stochastic regression with respect to aspect ratios, when using a right-sided kernel and the objective (3.1)

Aspect Ratio	Objective (3.1)		Cross Validation		
	k_i^*	e_{RMSE}	$\bar{k}_i$	$ \bar{k}_i - k_i^* $	$\frac{ \bar{k}_i - k_i^* }{k_i^*}$
ε_i					
4.0968	7	1.8295×10^{-3}	$6.\bar{3}$	$0.\bar{6}$	9.5238×10^{-2}
4.5407	7	1.8016×10^{-3}	$6.\bar{3}$	$0.\bar{6}$	9.5238×10^{-2}
0.74406	7	1.9946×10^{-3}	$6.\bar{3}$	$0.\bar{6}$	9.5238×10^{-2}
4.5777	7	1.7855×10^{-3}	$6.\bar{3}$	$0.\bar{6}$	9.5238×10^{-2}
3.2078	6	1.7589×10^{-3}	$6.\bar{4}$	$0.\bar{4}$	0.074
0.60051	7	1.9269×10^{-3}	$6.\bar{3}$	$0.\bar{6}$	9.5238×10^{-2}
1.4827	6	1.8009×10^{-3}	$6.\bar{4}$	$0.\bar{4}$	0.074
2.7910	6	1.8604×10^{-3}	$6.\bar{4}$	$0.\bar{4}$	0.074
4.7928	5	1.7152×10^{-3}	$6.\bar{5}$	$1.\bar{5}$	0.31
4.8288	6	1.7909×10^{-3}	$6.\bar{4}$	$0.\bar{4}$	0.074

Table 10 Cross validation of stochastic regression with respect to aspect ratios, when using a right-sided kernel and the objective (3.3)

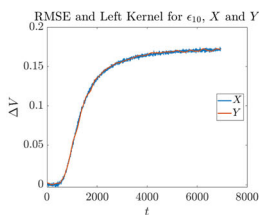
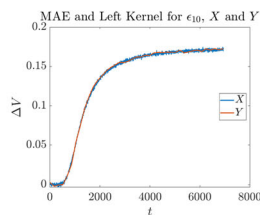
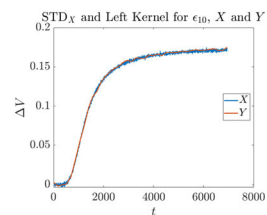
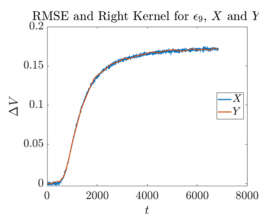
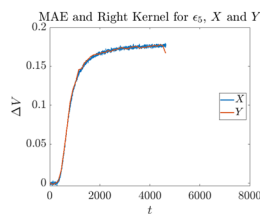
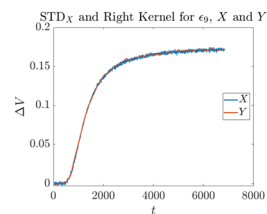
Aspect Ratio	Objective (3.3)		Cross Validation		
	k_i^*	μ	$\bar{k}_i$	$ \bar{k}_i - k_i^* $	$\frac{ \bar{k}_i - k_i^* }{k_i^*}$
ε_i					
4.0968	6	1.3821×10^{-3}	$6.\bar{3}$	$0.\bar{3}$	0.05
4.5407	7	1.3751×10^{-3}	$6.\bar{2}$	$0.\bar{7}$	0.1
0.74406	8	1.5024×10^{-3}	$6.\bar{1}$	$1.\bar{8}$	0.2361
4.5777	7	1.3938×10^{-3}	$6.\bar{2}$	$0.\bar{7}$	0.1
3.2708	6	1.3266×10^{-3}	$6.\bar{3}$	$0.\bar{3}$	0.05
0.60051	7	1.4302×10^{-3}	$6.\bar{2}$	$0.\bar{7}$	0.1
1.4827	6	1.3562×10^{-3}	$6.\bar{3}$	$0.\bar{3}$	0.05
2.7910	5	1.4431×10^{-3}	$6.\bar{4}$	$1.\bar{4}$	0.28
4.7928	5	1.3391×10^{-3}	$6.\bar{4}$	$1.\bar{4}$	0.28
4.8288	6	1.3549×10^{-3}	$6.\bar{3}$	$0.\bar{3}$	0.05

In addition, we have used cross-validation to demonstrate that optimal bandwidths are strikingly consistent across instruments with a range of different aspect ratios. Our cross-validation also suggests that the centered kernel is the most appropriate in this application, consistent with our findings from the experimental measurement.

Bio-FETs are a promising way to measure the biomarker levels, and we expect forthcoming Bio-FET measurements to lead to exciting applications of this work. In addition to Bio-FETs, we anticipate that other applications of stochastic regression are on the horizon. Biomedical technology will continue to play an increasingly impor-

Table 11 Cross validation of stochastic regression with respect to aspect ratios, when using a right-sided kernel and the objective (3.4)

Aspect Ratio	Objective (3.4)		Cross Validation		
	k_i^*	σ	$\bar{k}_i$	$ \bar{k}_i - k_i^* $	$\frac{ \bar{k}_i - k_i^* }{k_i^*}$
ε_i					
4.0968	7	1.7763×10^{-3}	$6.\bar{3}$	$0.\bar{6}$	0.095238
4.5407	7	1.1643×10^{-3}	$6.\bar{3}$	$0.\bar{6}$	0.095238
7.4406	7	1.3042×10^{-3}	$6.\bar{3}$	$0.\bar{6}$	0.095238
4.5777	7	1.1162×10^{-3}	$6.\bar{3}$	$0.\bar{6}$	0.095238
3.2078	6	1.1552×10^{-3}	$6.\bar{4}$	$0.\bar{4}$	$0.\overline{074}$
0.60051	7	1.2916×10^{-3}	$6.\bar{3}$	$0.\bar{6}$	0.095238
1.4827	6	1.1851×10^{-3}	$6.\bar{4}$	$0.\bar{4}$	$0.\overline{074}$
2.7910	6	1.1603×10^{-3}	$6.\bar{4}$	$0.\bar{4}$	$0.\overline{074}$
4.7928	5	1.0721×10^{-3}	$6.\bar{5}$	$1.\bar{5}$	$0.3\bar{1}$
4.8288	6	1.1715×10^{-3}	$6.\bar{4}$	$0.\bar{4}$	$0.\overline{074}$

(a) The aspect ratio $\varepsilon_{10} = 4.8288$ is shown, which corresponds to $k_{10}^* = 7$.(b) The aspect ratio $\varepsilon_{10} = 4.8288$ is shown, which corresponds to $k_{10}^* = 7$.(c) The aspect ratio $\varepsilon_{10} = 4.8288$ is shown, which corresponds to $k_{10}^* = 7$.**Fig. 7** Three graphs corresponding to the optimal values of (3.1), (3.3), and (3.4), in the case of the left-sided kernel. In each graph there is an overlay of a synthetic measurement X and its approximation Y (a) The aspect ratio $\varepsilon_9 = 4.7928$ is shown, which corresponds to $k_9^* = 5$.(b) The aspect ratio $\varepsilon_9 = 4.7928$ is shown, which corresponds to $k_9^* = 5$.(c) The aspect ratio $\varepsilon_9 = 4.7928$ is shown, which corresponds to $k_9^* = 5$.**Fig. 8** Three graphs corresponding to the optimal values of (3.1), (3.3), and (3.4), in the case of the right-sided kernel. In each graph there is an overlay of a synthetic measurement X and its approximation Y

tant role in healthcare and the economy. Such technology often produces times series measurements like those analyzed in the present work, and we expect that optimizing stochastic regression in these specific situations will recover the most accurate reconstruction of the true signal and in so doing give us an accurate estimate of the noise. Moreover, for many instrument applications, having an estimate of the noise is useful for calibration.

Disclaimer

This paper is an official contribution of the National Institute of Standards and Technology, and is not subject to copyright in the United States. Commercial products are identified in order to adequately specify certain procedures. In no case does such identification imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the identified products are necessarily the best available for the purpose.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Epanechnikov VA (1969) Nonparametric estimation of a multidimensional probability density. *Theory of Probability and its Applications* 14(1):153–158
- Evans RM, Balijepalli A, Kearsley AJ (2020) Transport phenomena in biological field effect transistors. *SIAM J Appl Math* 80(6):2586–2607
- Kearsley AJ, Gadhyan Y, Wallace WE (2014) Stochastic regression modeling of chemical spectra. *Chemom Intell Lab Syst* 139:26–32
- Patrone PN, Romsos EL, Cleveland MH, Vallone PM, Kearsley AJ (2020) Affine analysis for quantitative pcr measurements. *Anal Bioanal Chem* 412(28):7977–7988

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.