Kernel Estimates for One- and Two-Dimensional Ion Channel Dwell-Time Densities

Rafael A. Rosales,* William J. Fitzgerald,[†] and Stephen B. Hladky[‡]

*Departamento de Matemáticas, Instituto Venezolano de Invetigaciones Científicas, Caracas 1020-A, Venezuela, and Departments of [†]Engineering and [‡]Pharmacology, University of Cambridge, Cambridge, United Kingdom

ABSTRACT   In this paper, we compare nonparametric kernel estimates with smoothed histograms as methods for displaying logarithmically transformed dwell-time distributions. Kernel density plots provide a simpler means for producing estimates of the probability density function (pdf) and they have the advantage of being smoothed in a well-specified, carefully controlled manner. Smoothing is essential for multidimensional plots because, with realistic amounts of data, the number of counts per bin is small. Examples are presented for a 2-dimensional pdf and its associated dependency-difference plot that display the correlations between successive dwell times.

INTRODUCTION

Dwell-time histograms and the probability density functions (pdfs) estimated from them are an important means for presenting the results of patch-clamp investigations of ion channel currents. Blatz and Magleby (1986), McManus et al. (1987) and Sigworth and Sine (1987) demonstrated that constructing the distribution with a logarithmic rather than a linear time scale can greatly reduce the computations required for subsequent fitting and has a number of advantages for display of the data: it produces a pdf for a single exponential whose peak height and position on the log axis specify the amplitude and time constant of the exponential decay; it allows convenient display of dwell times spanning several orders of magnitude; and it provides a much better visual impression of the number of components in the exponential mixture. A major disadvantage has been that it is not straightforward to construct a histogram with log bins when the original data have been digitized. The mapping of dwell times for data recorded at a fixed sample interval to log bins is highly irregular: some bins for short durations can never receive counts, and significant irregularities between adjacent bins extend beyond the bin corresponding to 100 times the sample interval. Stark and Hladky (2000) have shown that simple approximations based on the nature of the binning process combined with the transformation rules for probability density functions make it possible to adjust the display of the counts to compensate for these irregularities. We report here that construction of kernel density plots (Bowman and Azzalini, 1997) rather than histograms provides a simpler method for displaying estimates of the pdf. These kernel-density plots also have the

advantage of producing an estimate that has been smoothed in a well-specified, carefully controlled manner. The advantages of smoothing are most marked in multidimensional plots such as the 2-dimensional histograms used to display the occurrence of pairs of dwell-times (Magleby and Song, 1992) where the number of data counts per bin in a traditional histogram would be small.

In a traditional histogram, each sojourn in a state or aggregate of states with the same conductance, i.e., a conductance class, is represented by a count in the bin corresponding to the dwell time. The contribution to the plot is a small rectangle with unit height and the same width as the bin. The placing of the rectangle is the same regardless of where the dwell time falls within the range of times within the bin. A nonparametric kernel density estimate is constructed by representing each sojourn by a kernel function positioned at the actual value of the sampled dwell time. A common choice of function is a normal with width still to be specified. The kernel density estimate is then the sum of the increments for all sojourns in the class. Because, for a log plot, the increment can be chosen to have unit area, the total area under the resulting curve is equal to the number of counts, and the pdf can be obtained by dividing all values on the curve by the total number of counts.

DWELL-TIME KERNEL ESTIMATES

Single class dwell times

Let $t = (t_i)$, $i = 1, 2, \ldots, n$ be a series of dwell times at a particular class. Also let the kernel function ϕ be a normal density on the log axis $z = \ln(t)$, with mean located at the value of the log of the dwell time, and scale dependent width (standard deviation), σ_i,

$$\phi_i(z) = \frac{1}{\sigma_i \sqrt{2\pi}} \exp\left[-\frac{(z - x_i)^2}{2\sigma_i^2} \right], \qquad (1)$$

then the kernel estimate for the pdf of $x = (x_i)$, $i = 1, \ldots,$ n, is

$$h(z) = \frac{1}{n} \sum_{i=1}^{n} \phi_i(z). \qquad (2)$$

For long dwell times, σ_i can be independent of the dwell time. Its value affects the appearance of the plot in a manner analogous to the way the bin width affects a histogram. If the width is much too narrow, the plot consists of many sharp spikes with a visual impression reflecting the accident of the particular samples taken rather than the underlying pdf. In contrast, if the width is much too wide, the plot becomes a single smear with loss of most information about the data. Fortunately, σ_i can be chosen simply by considering the properties of a log dwell-time plot.

When plotted on a logarithmic scale, the width of a exponential component is independent of its amplitude or time constant. Hence, the principles for choosing the kernel for a mixture are the same as those for a single exponential. Estimates constructed with several choices of constant width, $\sigma_i = \sigma$, are compared in Fig. 1 A with the underlying theoretical pdf for an exponential with time constant 250 μs sampled at 5 μs intervals. Data points for the estimates (1000 per curve) were obtained deterministically in proportion to the value of the pdf (see Stark and Hladky, 2000, for details). For each data point, the kernel increment was evaluated at values of z on a grid with 20 steps per e-fold change in dwell time, which is sufficiently fine to avoid distorting the shape and size of the kernel. (The grid spacing must be less than σ.) Evaluations were restricted to the range of z for which ϕ is significant. Consider first the range $x_i = \ln(t_i) > -3$. For these values, any $\sigma < 0.2$ preserves the shape. Larger values of σ smear out the distribution, lowering the peak and adding weight for longer and shorter dwell times. The present example is not suitable to demonstrate the effect of making σ too small because there are too many data points and they are deterministically spaced.

The series of peaks at short dwell time in the trace for $\sigma = 0.1$ in Fig. 1 A represent the increments added for dwell times of 1, 2, . . . sample intervals. Because the data have been sampled at discrete times, all increments to the kernel estimate for short dwell times are centered at just these precise values. Thus these peaks are an artifact of the sampling, not a feature of the underlying pdf. As σ is increased, the increments are smeared out over the possible range, producing a better representation of the underlying signal.

It makes no sense to set the width of the kernel increment to a value smaller than the uncertainty in the dwell times. This uncertainty must be at least of the order of a sample interval. A constant uncertainty on a linear scale transforms to a variable uncertainty on a log scale. The value of σ_i
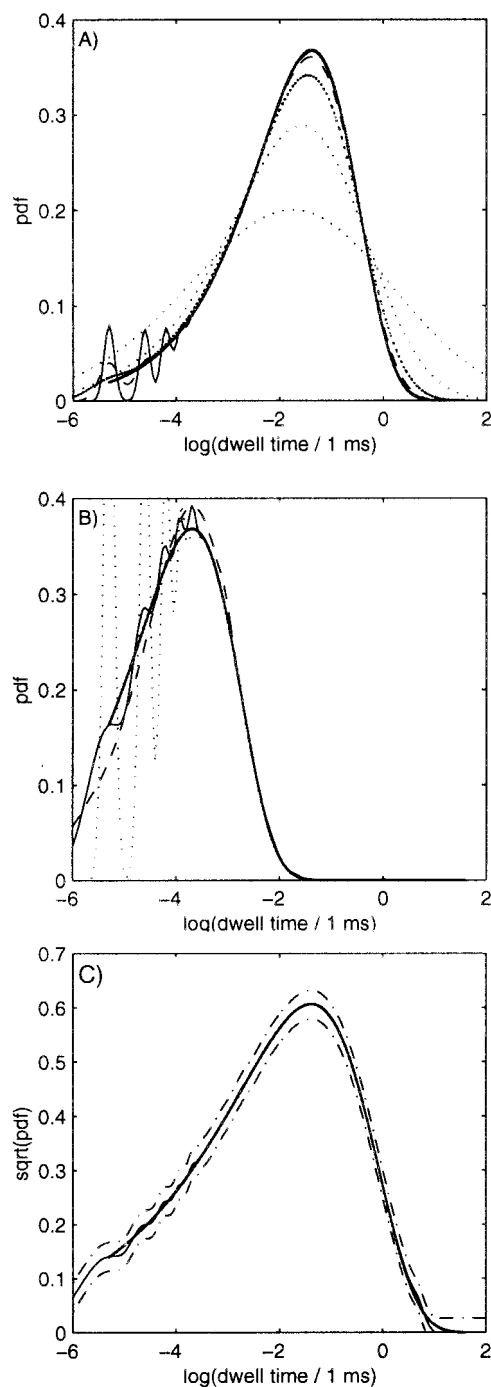

FIGURE 1 Comparisons of kernel estimates with a theoretical exponential probability density function of $\tau = 250$ μs. (A) Estimates for different normal kernels with constant width ζ. ζ values are: *solid curve*, 0.1; *dashed*, 0.2; *heavy dots*, 0.4; *light dots*, 0.8 and 1.6. The theoretical curve is shown as a heavier line. (B) Estimates with different corrections for the uncertainty for very short dwell-times. $\tau = 50$ μs. $\zeta = 0.1$. The *m* values are dotted line, 0; *solid*, 1; and *dashed*, 2. (C) Demonstration of a close match between estimate and theory (*thin* and *heavy solid lines*). τ is 250, $\zeta = 0.1$ $m = 1$. $\pm$SD confidence bands for the estimate, computed via Eq. A8, are also shown as dash-dotted lines. In this and the next two figures, "log" means logarithm to the base e.

corresponding to an uncertainty of m sample intervals for a dwell time of duration t_i is

$$\sigma_i = \ln\left(1 + \frac{m\delta}{2t_i}\right), \tag{3}$$

where δ is the sample interval. The value of σ_i calculated from this relation is $\ln(1.5)$ for $m = 1$ and a dwell time equal to one sample interval, but it decreases rapidly, falling below 0.1 for dwell times exceeding five sample intervals. Figure 1 B illustrates results obtained with several values of m. For $m = 1$, the kernel estimate is not completely smooth but the variations occur centered about the underlying pdf. For larger m, the curve is smoother but it lies below the pdf for small dwell times and above it for somewhat larger values. The optimal value of m is close to 1 regardless of the time constant. In the following, we have set $m = 1$.

Eq. 3 cannot be used on its own as a means of specifying the width, because σ_i becomes much too small for large dwell times. A simple, effective procedure is to use instead

$$\sigma_i = \max\left[\ln\left(1 + \frac{\delta}{2t_i}\right), \zeta\right]. \tag{4}$$

Figure 1 C displays the result of applying this relation with $\zeta = 0.1$, which produces essentially no distortion of the pdf. Although freedom from distortion favors this small ζ, elimination of the roughness resulting from random sampling favors large ζ. The more data available the smaller the value of ζ that should be used. However, because $\zeta = 0.2$ produces negligible distortion, although that produced by 0.8 would normally be unacceptable, in practice, it is the range 0.2–0.4 that will be useful. A procedure for the optimal estimation of ζ given the number of dwell times available, or, alternatively, an estimate of the minimum number of times required for a chosen value, is outlined in the Appendix. For 900 dwell times and an underlying pdf with two spaced exponential components, the optimal value is between 0.26 and 0.34.

Joint dwell times

Let $x_{i,1} = \ln(t_{i,1})$ and $x_{i,2} = \ln(t_{i,2})$ for $i = 1, 2, \ldots, n$, denote the log-transformed sojourns of consecutive dwell times in any two classes $\mathfrak{F}$ and $\mathfrak{C}$. Bivariate kernel estimates for the joint density $h(z_1, z_2)$ of adjacent dwell-time pairs, $(x_{i,1}, x_{i,2})$, may be defined by extending Eq. 1 as the product of univariate kernels along each direction,

$$\phi_i(z_1, z_2) = \phi_i(z_1)\phi_i(z_2). \tag{5}$$

This function is defined for all points (z_1, z_2) on the real plane, because, in principle, $(\ln(t_{i,1}), \ln(t_{i,2})) \in \mathbf{R}^2$. This construction leads to the estimate

$$h(z_1, z_2) = \frac{1}{2n\pi} \sum_{i=1}^{n} \frac{1}{\sigma_{i,1}\sigma_{i,2}} \exp[f(z_1, z_2)] \tag{6}$$

with

$$f(z_1, z_2) = -\left[\frac{(z_1 - x_{i,1})^2}{2\sigma_{i,1}^2} + \frac{(z_2 - x_{i,2})^2}{2\sigma_{i,2}^2}\right]. \tag{7}$$

Alternatively, it is possible to use a bivariate kernel with correlated coordinates to allow for smoothing along particular directions, see Bowman and Azzalini (1997) or Wand and Jones (1995).

Estimates based on Eq. 5 can also be used to generate plots of the dependency difference function introduced by Magleby and Song (1992),

$$D(z_1, z_2) = h(z_1, z_2)^{1/2} - h(z_1)^{1/2}h(z_2)^{1/2}, \tag{8}$$

with $h(z_1, z_2)$ given by Eq. 6 and $h(z_1)$, $h(z_2)$, the one-dimensional (1D) density estimates defined in Eq. 2, which can be calculated from $h(z_1, z_2)$ as the marginals
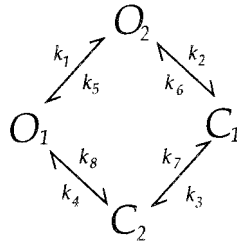
$$h(z_1) = \sum_{k=1}^{K_2} h(z_1, z_{2,k}), \quad h(z_2) = \sum_{k=1}^{K_1} h(z_{1,k}, z_2), \tag{9}$$

with K_1 and K_2 the sizes of the discrete grid used to evaluate the joint density along each coordinate. The dependency difference indicates the importance of the discrepancy between the observed distribution of dwell-time pairs and that which would arise if the dwell-times in each class were independent of each other.

RESULTS

It is routinely possible to obtain enough experimental data to construct 1D log dwell-time distributions with both a relatively large number of bins per decade and sufficient counts per bin to obtain a reasonable estimate of the pdf. The principal advantage of the use of kernel estimation in the construction of these 1D estimates is its simplicity. Binning, correction for the binning artefacts (see Stark and Hladky, 2000) and smoothing of the histogram are all replaced with a single, easily implemented summation of increments. The major disadvantage is that the smearing correction for the binning artifacts for dwell-times less than ~5 sample intervals is not as accurate (see for example Fig. 1 B) as the correction factors for log-binned histograms (see Stark and Hladky, 2000).

Two-dimensional (2D) pdfs or difference dependency plots require the square of the number of bins needed for the 1D plot, and, for these, the smoothing becomes critical. To illustrate the use of kernels for the estimation of open–

SCHEME 1

FIGURE 2 Joint dwell-time densities for dwell-time pairs for consecutive sojourn in $\mathfrak{F}$ (*right axis*) and $\mathfrak{C}$ (*left axis*). (*A*) Theoretical density predicted by the spectral decomposition of Q, (*B*) kernel density estimate, and (*C*) histogram smoothed as described in the text. Both (*B*) and (*C*) are generated from 900 dwell-time pairs fitted to synthetic data generated by the mechanism described by Q.

closed dwell-time densities, we have generated synthetic noisy data using the following four-state reversible cyclic mechanism (see Scheme 1).

The states were grouped into two classes, $\mathfrak{F}$ and $\mathfrak{C}$, with $\mathfrak{F} = \{O_1, O_2\}$ as the class of "open" states and $\mathfrak{C} = \{C_1, C_2\}$ the class of "closed" states. This decomposition induces the following partition for the Q matrix,

$$Q = \begin{bmatrix} Q_{\mathfrak{F}\mathfrak{F}} & Q_{\mathfrak{F}\mathfrak{C}} \\ Q_{\mathfrak{C}\mathfrak{F}} & Q_{\mathfrak{C}\mathfrak{C}} \end{bmatrix}, \qquad (10)$$

with $Q_{\mathfrak{F}\mathfrak{F}}$ as the submatrix of rate constants between the states of the $\mathfrak{F}$ class, $Q_{\mathfrak{F}\mathfrak{C}}$ the rates corresponding to the transitions from the $\mathfrak{F}$ class to the $\mathfrak{C}$ class, and so on. The rate constants k_i; $i = 1, \ldots, 8$, were chosen to ensure states with widely separated life times within each aggregate. These were set to $k_1 = 6.183$, $k_2 = 0.454$, $k_3 = 2.697$, $k_4 = 1.665$, $k_5 = 0.446$, $k_6 = 13.163$, $k_7 = 0.182$, and $k_8 = 11.812$ (all in arbitrary units). In this case, both O_1 and C_1 are short-lived states ($\tau_{O_1} = 0.0551$, $\tau_{C_1} = 0.0629$), whereas O_2 and C_2 are long ($\tau_{O_2} = 1.351$, $\tau_{C_2} = 0.5519$). For this setting, long sojourns at each class originate from events of the type $O_1 \rightleftharpoons O_2$ or $C_1 \rightleftharpoons C_2$, which are then interrupted by brief transitions to the other class by events of the form $C_2 \rightarrow O_1 \rightarrow C_2$ or $O_2 \rightarrow C_1 \rightarrow O_2$.

To assess the performance of joint histograms and the estimate in Eq. 6, a noisy trace of 900 dwell-time pairs that correspond to the consecutive sojourns in $\mathfrak{F}$ and $\mathfrak{C}$ was generated by sampling at $\delta = 0.005$. These data were idealized by using the methods described in Rosales et al. (2001). Kernel estimates, together with the theoretical joint density, are shown in Fig. 2. The highest peaks, located right foreground, originate from brief sojourns in $\mathfrak{F}$ followed by long durations in $\mathfrak{C}$. In principle, these pairs could arise from events of the type $O_1 \rightarrow C_2$ or $O_1 \rightarrow C_2 \rightleftharpoons C_1$. The second maxima originates from long sojourns in $\mathfrak{F}$ followed by short ones in $\mathfrak{C}$. These correspond to the transitions $O_2 \rightarrow C_1$ or $O_1 \rightleftharpoons O_2 \rightarrow C_1$. A third minor peak corresponds to long consecutive sojourns in both classes, which might appear to correspond to transitions $O_2 \rightarrow C_2$. Figure 2 *B* presents the kernel estimate obtained from the 900 simulated dwell-time pairs. This estimate was obtained using Eq. 6 with $\zeta = 0.35$ on a 60×60 grid, showing close resemblance to the theoretical density in Fig. 2 *A*. Figure
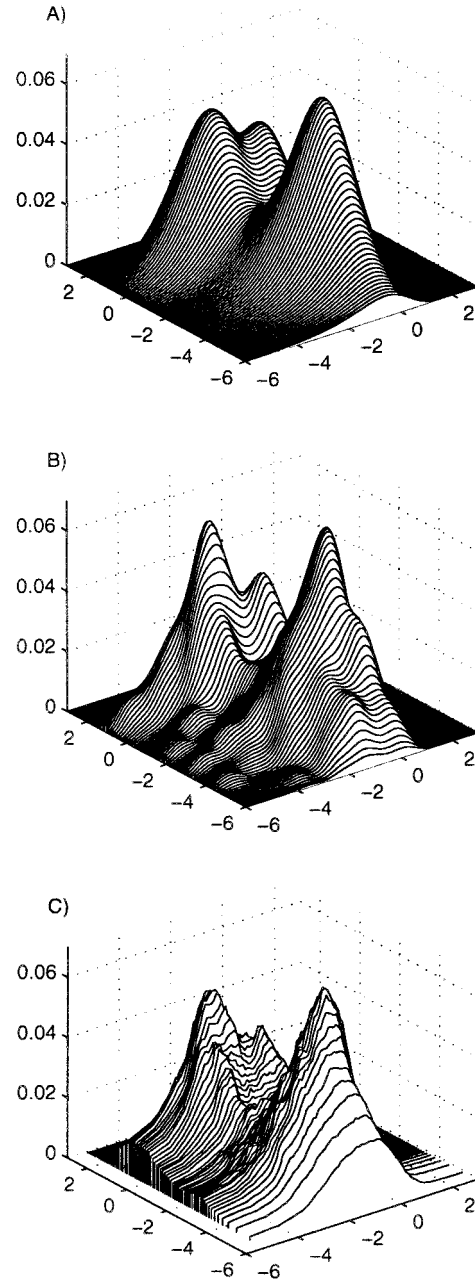
2 *C* shows a smoothed histogram constructed using the same sequence of dwell times by considering a 70×70 grid with 10 bins per e-fold change along each axis. The counts were corrected for the log binning artefacts as described by Stark and Hladky (2000). Without smoothing, both the raw histogram and the corrected plot had the appearance of a forest of trees. The histogram shown in Fig. 2 *C* was ob-

tained by replacing each point with an average of 361 points weighted according to the inverse of their distance to the center (the central point and the four nearest neighbors all have weight 1). Similar, but somewhat more ragged, results were obtained using a uniform average over 289 points instead of the inverse weighted average over 360. In both cases, the size of the averaging region were chosen so that the peak heights would be close to the theoretical pdf and the kernel estimates, but this agreement was achieved at the expense of considerable elevation in the distributions for short open or closed times. Smoother histogram plots could be produced using larger averaging regions, but, in these, there was severe loss of peak height and even larger tails.

Dependence-difference plots are shown in Fig. 3. Figure 3 *A* presents the theoretical results calculated from Eq. 8 using the theoretical joint and marginal densities obtained from the spectral decomposition of Q (Colquhoun and Hawkes, 1982). Figure 3 *B* shows the dependence-difference contour obtained using Eq. 8 and the estimate plotted in Fig. 2 *B*. Figure 3 *C* presents the corresponding dependence plot computed from the smoothed histogram displayed in Fig. 2 *C*. It can be seen from Fig. 3 *B*, that positive dependencies ($D(z_1, z_2) > 0$) are related to dwell-time pairs associated with events of the form $O_1 \rightarrow C_2$ and $O_2 \rightarrow C_1$, whereas negative dependencies are seen for the regions that correspond to the transitions $O_1 \rightarrow C_1$ and $O_2 \rightarrow C_2$. Therefore, the third minor maximum at Fig. 2 *B* corresponding to long sojourns in both $\mathfrak{F}$ and $\mathfrak{C}$ arises from composite transitions from O_2 to C_2, i.e., $O_2 \rightarrow C_1 \rightarrow C_2$ or $O_2 \rightarrow O_1 \rightarrow C_2$.

DISCUSSION

The comparison of the estimate for the open–closed density shown in Fig. 2 *B* with the theoretical shape in Fig. 2 *A*, and, even more markedly, the comparison of the dependency differences computed from them in Fig. 3, show clearly the utility of kernel density estimation for log-transformed dwell times. To produce the plot at Fig. 2 *C*, the histogram was first constructed and then corrected to account for the aberrations induced by log binning (Stark and Hladky, 2000). The corrected counts were finally smoothed. This sequence is accomplished in a single step when evaluating the kernel directly from Eq. 6, making these estimates simpler to compute.

Inspection of Fig. 2 and Fig. 3 may suggest that the corrected and smoothed histogram is closer than the density estimate to the underlying pdf. However, the amount of smoothing for the histogram was chosen to make the highest peak agree with the theoretical distribution to display a result as favorable as possible for this method. In practice, this choice is not available because it depends on the knowledge of the underlying density. By constrast, the kernel method provides an automatic choice for the amount of smoothing found by minimizing the bias introduced by the
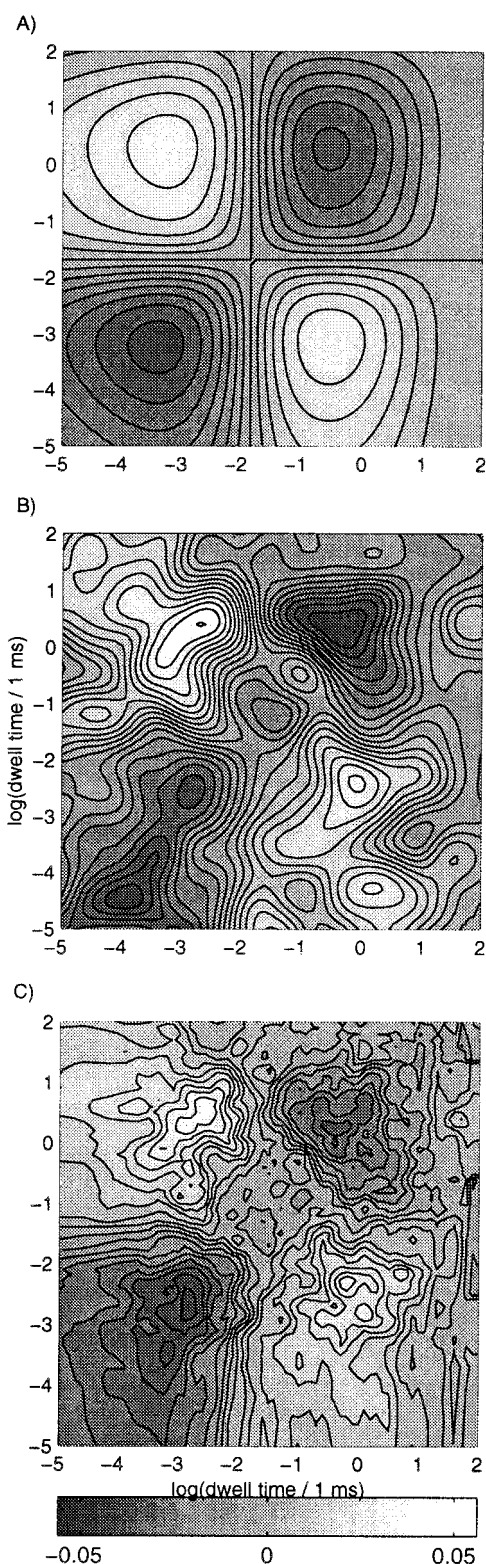

FIGURE 3 Dependence-difference contour plots. (*A*) is calculated from the theoretical joint and marginals given by Q and the dependency-difference definition in Eq. 8. (*B*) and (*C*) correspond to the dependencies calculated via Eq. 8 and Eq. 9 from the kernel estimate in Fig. 2 *B* and the histogram in Fig. 2 *C*, respectively. The vertical axis is used for the times in $\mathfrak{C}$, whereas the horizontal axis for those in $\mathfrak{F}$.

estimation procedure. It should be emphasized that the results presented show that estimates obtained with ζ in the range of 0.2–0.4 do not create or eliminate peaks nor do they shift peaks appreciably on the z axis or the (z_1, z_2) plane.

Construction of kernel estimates and histograms should produce similar results as the number of dwell times increases asymptotically toward infinity. However, the kernel estimator can be shown to be more efficient because its bias (in terms of the mean integrated squared error, see the Appendix) decreases as $\mathbb{O}(n^{-4/5})$ whereas the bias of a histogram follows $\mathbb{O}(n^{-2/3})$ (see Wand and Jones, 1995, p. 23). This property is particularly useful when estimating joint densities for the kinetic analysis of ion channels, since the observations available from a single molecule are usually brief or few when compared to the number of bins that would be required.

The errors incurred when fitting the kernels generated by Eqs. 4 and A5 with exponential mixtures gives similar results as those shown in Fig. 2 of Sigworth and Sine (1987). However, although the smoothing introduced by the kernel estimation can be chosen to minimize the mean integrated squared error, quantitative assessment of the effects of this smoothing on subsequent fitting remains a subtle, open question.

APPENDIX

Optimal Bandwidths

An optimal choice for a fixed scale bandwidth ζ follows by minimizing a measure for the overall effectiveness of h in estimating the underlying density g. A commonly used distance is provided by the mean integrated squared error. In the univariate case, this is

$$\epsilon(h) = \mathbf{E}\left\{\int (h(x) - g(x))^2 \, dx\right\}$$

$$= \int [\mathbf{E}\{h(x)\} - g(x)]^2 \, dx + \int \mathbf{V}\{h(x)\} \, dx, \quad \text{(A1)}$$

with $\mathbf{E}$, $\mathbf{V}$ denoting expectation and variance (Bowman and Azzalini, 1997, p. 26). Here $x \in R$ is used to denote an element of the domain of all the functions to be considered. Asymptotically, as n becomes large, Taylor series approximations applied to both the bias and the variance terms lead to

$$\epsilon(h) \approx \frac{1}{4}\zeta^4 \int [g''(x)]^2 \, dx + \frac{1}{n\zeta}\int \phi(x)^2 \, dx. \quad \text{(A2)}$$

An optimal width follows by minimizing $\epsilon(h)$ with respect to ζ. The second term in Eq. A2 is obtained by taking $\phi(x)$ as defined by Eq. 1 with $\sigma = \zeta$, and integrating over $\mathbf{R}$,

$$\int \phi(x)^2 \, dx = \frac{1}{2\sqrt{\pi}}. \quad \text{(A3)}$$

An optimal bandwidth is then given by

$$\zeta = \left(n2\sqrt{\pi}\int [g''(x)]^2 \, dx\right)^{-1/5}. \quad \text{(A4)}$$

By considering an exponential pdf with rate λ in the log metric,

$$g(x) = \lambda \exp[x - \exp(x)\lambda],$$

the required integral of $[g''(x)]^2$ in Eq. A2 yields ¼, and, hence, from Eq. A4, the optimal value is $\zeta = 1.02/n^{1/5}$.

More generally, by taking an exponential mixture of two components with equal amplitudes and rates $\lambda_1, \lambda_2 = \beta\lambda_1$, and varying β from 0 to ∞, the width is bounded as

$$\frac{1.02}{n^{1/5}} \leq \zeta \leq \frac{1.34}{n^{1/5}}.$$

The smallest value is obtained for isolated exponentials, i.e., for $\beta = 1$. The greatest is achieved at $\beta = 0.22$ (or $\beta = 4.5$). Because the data usually follows a multiple-component mixture, it seems reasonable to take the upper limit,

$$\zeta = 1.34n^{-1/5}. \quad \text{(A5)}$$

The bivariate case follows from a similar argument. Let $\bar{x} = (x_1, x_2)$, and $g(\bar{x})$ the joint pdf for successive sojourns x_1 and x_2. By considering Eq. 7 and assuming $\sigma_{i,1} = \sigma_{i,2}$, i.e., $\zeta_1 = \zeta_2 \equiv \xi$, a Taylor series approximation for Eq. A1 gives

$$\epsilon(h) \approx \frac{1}{4}\xi^4 \iint [\nabla^2 g(\bar{x})]^2 \, d\bar{x} + \frac{1}{n\xi^2}\iint \phi(\bar{x})^2 \, d\bar{x},$$

where the integrations are taken over $\mathbf{R}^2$, and

$$\nabla^2 g(\bar{x}) = \frac{\partial^2 g(\bar{x})}{\partial x_1} + \frac{\partial^2 g(\bar{x})}{\partial x_2}.$$

The integral for $\phi(\bar{x})^2$ leads to $1/(16\pi^2)$, and, hence, the optimal width is

$$\xi = \left(n2^3\pi^2 \iint \nabla^2 g(\bar{x}) \, d\bar{x}\right)^{-1/6}.$$

By considering a bivariate mixture, $g(\bar{x})$, with two components along each coordinate with rates $(\lambda_1, \beta_1\lambda_1)$ for x_1 and $(\lambda_2, \beta_2\lambda_2)$ for x_2,

$$g(\bar{x}) = \frac{1}{4}\prod_{i=1,2}\{\lambda_i\exp[x_i - \exp(x_i)\lambda_i]$$

$$+ \beta_i\lambda_i\exp[x_i - \exp(x_i)\beta_i\lambda_i]\},$$

and letting β_1 and β_2 vary from 0, up to $\lim(\beta_1, \beta_2) \to +\infty$, gives

$$\frac{0.66}{n^{1/6}} \leq \xi \leq \frac{0.93}{n^{1/6}}.$$

Again, we propose as an optimal bandwidth the upper limit of this interval, that is,

$$\xi = 0.93n^{-1/6}. \quad \text{(A6)}$$

As one should expect, both widths ζ and ξ decrease monotonically with n. For the data shown in the Results section where $n = 900$, Eq. A5 gives $\zeta = 0.34$ and Eq. A6 $\xi = 0.299$.

Confidence Intervals

Following Bowman and Azzalini (1997), the scatter associated with the kernel-estimation procedure can be approximated by restricting attention to the kernel variance $\mathbf{V}\{h(x)\}$. This quantity can be made independent of the underlying density, g, by considering a square root transformation of the kernel estimate. More generally, denote by T a continuous differentiable transformation of $h(x)$. By using a Taylor expansion,

$$\mathbf{V}\{T(h(x))\} \approx \mathbf{V}\{h(x)\}[T'(\mathbf{E}\{h(x)\})]^2.$$

If $T(h(x)) = \sqrt{h(x)}$, then the principal term of the above expression becomes

$$\mathbf{V}\{\sqrt{h(x)}\} \approx \frac{1}{4}\frac{1}{n\zeta}\int\phi(u)^2\,du, \qquad (A7)$$

which is independent of the unknown density g. Thus, on the square root scale, the uncertainty associated with the estimate can be expressed by placing a band of width equal to two standard errors (sd) around $\sqrt{h(z)}$. From Eq. A7, this gives

$$2\text{sd} = 2\sqrt{\frac{1}{4n\zeta}\int\phi(u)^2\,du},$$

and, following Eq. A3,

$$\text{sd} = \frac{1}{4}\sqrt{\frac{2}{n\sqrt{\pi}\zeta}}. \qquad (A8)$$

These variance-regularizing procedures have also been considered by Sigworth and Sine (1987) in the histogram context.

REFERENCES

Blatz, A. L., and K. L. Magleby. 1986. Quantitative description of three modes of activity of fast chloride channels from rat skeletal muscle. *J. Physiol. (Lond.).* 378:141–174.

Bowman, A. W., and A. Azzalini. 1997. Applied Smoothing Techniques for Data Analysis. Oxford Statistical Science Series, 18. Clarendon Press, Oxford, UK.

Colquhoun, D., and A. G. Hawkes. 1982. On the stochastic properties of bursts of single ion channel openings and of clusters of bursts. *Phil. Trans. Roy. Soc. Lond. B.* 300:1–59.

Magleby, K. L., and L. Song. 1992. Dependency plots suggest the kinetic structure of ion channels. *Proc. R. Soc. Lond. B.* 249:133–142.

McManus, O. B., A. L. Blatz, and K. L. Magleby. 1987. Sampling, log binning, fitting and plotting durations of open and shut intervals from single channels and the effects of noise. *Pflügers Arch. Eur. J. Physiol.* 410:530–553.

Rosales, R., J. A. Stark, W. J. Fitzgerald, and S. B. Hladky. 2001. Bayesian restoration of ion channel records. *Biophys. J.* 80:1088–1103.

Sigworth, F. J., and S. M. Sine. 1987. Data tranformations for improved display and fitting of single-channel dwell time histograms. *Biophys. J.* 52:1047–1054.

Stark, J. A., and S. B. Hladky. 2000. Adjustments for the display of quantized ion channel dwell times in histograms with logarithmic bins. *Biophys. J.* 78:662–667.

Wand, M. P., and M. C. Jones. 1995. Kernel Smoothing. Monographs on Statistics and Applied Probability, 60. Chapman & Hall, London.