Statistical Assessment of Change Point Detectors for Single Molecule Kinetic Analysis

Sean P. Parsons · Jan D. Huizinga

Abstract Change point detectors (CPDs) are used to segment recordings of single molecules for the purpose of kinetic analysis. The assessment of the accuracy of CPD algorithms has usually been based on testing them with simulated data. However, there have not been methods to assess the output of CPDs from real data independent of simulation. Here we present one method to do this based on the assumption that the elementary kinetic unit is a stationary period (SP) with a normal distribution of samples, separated from other SPs by change points (CPs). Statistical metrics of normality can then be used to assess SPs detected by a CPD algorithm (detected SPs, DSPs). Two statistics in particular were found to be useful, the z-transformed skew (S_Z) and z-transformed kurtosis (K_Z). $K_Z(S_Z)$ plots of DSP from noise, simulated data and single ion channel recordings showed that DSPs with false negative CP could be distinguished. Also they showed that filtering had a significant effect on the normality of data and so filtering should be taken into account when calculating statistics. This method should be useful for analyzing single molecule recordings where there is no simple model for the data.

Keywords Central moment · Change point detection · Kurtosis · Single molecule · Skewness

Introduction

Time series of single molecules are often characterized by near instantaneous changes intervening between periods of apparent stability. Examples include currents through ion channels, fluorescence intermittence in photosynthetic complexes and quantum dots, atomic force measurements of proteins, single molecule fluorescence resonance energy transfer and molecular dynamics simulations. This behavior can be interpreted as the switching of the molecule between stable conformations (Fig. 1). Thus, the sine qua non of the kinetic and thermodynamic analysis of the molecule is the accurate determination of change points (CPs) and as a corollary the stable (or stationary) periods (SPs). To this end there are a large number of change point detection (CPD) algorithms. These algorithms usually perform a simple mathematical operation on the series to produce a change point signal, which indicates the likely presence of a CP. This operation can be the derivative or some local (window) statistic. The CP signal can then be thresholded to segment the series into detected CPs and SPs, DCPs and DSPs respectively (Fig. 1).

The application of CPDs usually falls between two philosophical posts. At one end the series appears to conform to some simple model, for instance there are only two amplitude levels at which SPs occur, corresponding to (at least) two conformational states. The object of CPD is simply to automate the fitting of this data according to the model and only a short section of the fitted series has to be visually inspected to judge how good this fit is. At the other end there may be no simple model for the data or if there is, there may be factors limiting the effectiveness of the CPD such as short, bandwidth limited events. In these cases assessment has been made with simulated time series. CPs and SPs are generated, either purely randomly or based on some kinetic model, and to this is added Gaussian distributed noise. Assessment has often been limited to a visual comparison of the fit of the DSPs and DCPs with the unnoised simulation. However, a more quantitative approach has been taken by several authors, for instance

S. P. Parsons (✉) · J. D. Huizinga
McMaster University, Hamilton, ON, Canada
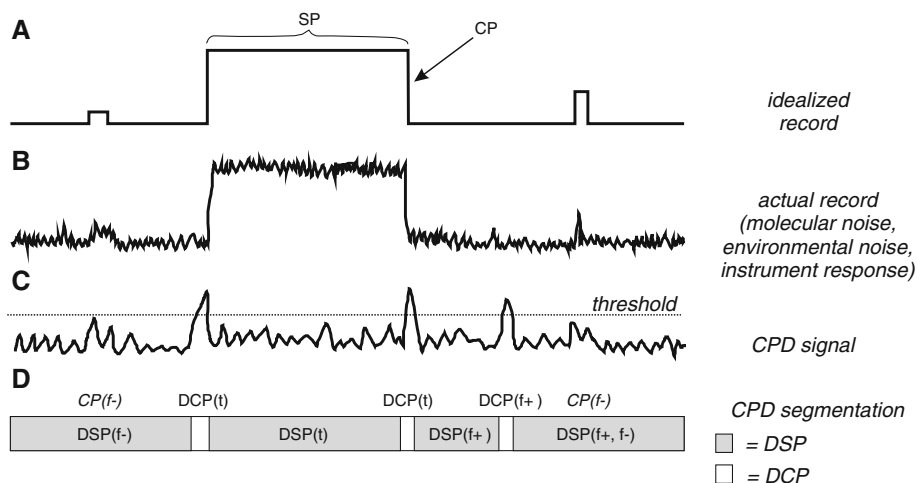e-mail: sparsonslab@gmail.com

Fig. 1 The idealized single molecule and change point detection (CPD). **a** The signal from the idealized single molecule consists of periods where the amplitude does not change, stationary periods (SPs), separated by instantaneous changes in amplitude, change points (CPs). During each SP the molecule is in a particular conformational state and at a CP the molecule switches from one state to another. **b** In any conformational state the molecule will not be completely motionless, so there will be some "molecular noise" in the recorded signal. Of course there is (or can be) a thermodynamic continuum between "molecular noise" and "conformational state change," but in our idealization we make a dichotomy. This molecular noise, together with environmental and instrument noise and the properties of the recording apparatus (its impulse response), add to the effect of giving a less than ideal recorded signal. However, the noise during an SP, its amplitude distribution, should be unchanging during an SP and

therefore the SP is stationary in the strictly mathematical definition of stationary. **c** Most CPD algorithms perform a simple linear mathematical operation on the recording to produce a CPD signal, the amplitude of which should (we hope) increase at a CP. **d** It is then simply a matter of thresholding the CPD signal to segment the recording into detected CPs and detected SPs (DCPs and DSPs, respectively). Of course with all the noise in the recorded signal, no CPD is going to be perfect in its matching of DSP to SP, DCP to CP. A DSP may overlap an undetected CP; such a DSP is tagged as false negative, f−, because the false-negative detection of the CP. One of the two DCPs bordering a DSP may not coincide with a CP—as such one of those DCPs must be falsely positive, and so the DSP is tagged f+. If a DSP does not overlap any CP and neither of its bordering DCPs is falsely positive than the DSP is regarded as true and is tagged t

counting the number of false-positive and false-negative DCPs; the accuracy of DSP amplitudes; or with kinetic models, the accuracy of the kinetic data derived from the DSPs (Carter et al. 2008; Riessner et al. 2002). Either way, once a CPD has been tested with simulated data, its application to real data is then a matter of trust. If it works okay with simulated data, and is perhaps found to be the best CPD with such data, then it can only be applied blindly to the real data—its output must be taken on trust as a whole or not at all. But no CPD is perfect, they will all produce false-negatives and false-positives, and there has been no method of testing for these with real data. Here we present one such method.

Two assumptions made with single molecule data are, first, that the SPs are in fact stationary in the strict sense—that is, the distribution of amplitudes does not change with time during the SP; and second, that this distribution is Gaussian (normal), reflecting contamination of the actual SP amplitude by instrumental and environmental noise. It would seem, then, that a simple way of assessing the output of CPDs is to look at the statistics of DSPs, specifically statistics that measure how much the DSP approximates a normal distribution. This is the basic idea of our method.

Methods

Sources of Noise

Noise was recorded from an inside-out patch without channel activity. The patch clamp electrophysiology set up consisting of an CV201A headstage, Axopatch 200A amplifier and 1322A Digidata (all Molecular Devices, Sunnyvale, CA). The recording was made at 10 kHz sampling frequency with the amplifier's low pass Bessel filter cutoff set to 5 kHz. White noise with a standard normal distribution (mean of zero, standard deviation of one) was generated in silico by the Mersenne Twister random number generator available in the GNU Scientific Library (GSL) for C++.

Monte Carlo Estimation of Noise Statistics

Noise time-series were generated as described above. DSPs (Fig. 1) of random length were taken randomly from the series. As length (L) was typically plotted on a log axis, we wanted $\log L$ to be uniformly distributed between a minimum and maximum value (logs of L_{min} and L_{max}). The probability density function for this is

$$P(x) = \begin{cases} 1/(x_{\max} - x_{\min}) & x_{\min} < x < x_{\max} \\ 0 & \text{otherwise} \end{cases}$$

$$x = \log L, \quad x_{\min} = \log L_{\min}, \quad x_{\max} = \log L_{\max}$$

Thus, the cumulative density function is

$$C(x) = \int_{-\infty}^{x} P(x)\mathrm{d}x = u$$
$$= \begin{cases} 0 & x < x_{\min} \\ (x - x_{\min})/(x_{\max} - x_{\min}) & x_{\min} < x < x_{\max} \\ 1 & x > x_{\max} \end{cases}$$

and so the inverse cumulative density function is

$$C^{-1}(u) = x = \begin{cases} -\infty & u < 0 \\ u(x_{\max} - x_{\min}) + x_{\min} & 0 < u < 1 \\ +\infty & u > 1 \end{cases}$$

Thus, values of L were calculated as

$$L = 10^x = 10^{C^{-1}(u)} = L_{\min}(L_{\max}/L_{\min})^u$$

where u was a uniformly distributed pseudo random number between 0 and 1 (generated by the GSL Mersenne Twister generator). Similar derivations have been made for logarithmic binning of dwell times (Sansom et al. 1989; Sigworth and Sine 1987). DSPs were taken at random points in the series using uniformly distributed pseudo random numbers (again using the Mersenne Twister routine).

Random SP Simulation

Noise time-series were generated as described above. To these series, SPs were added with random lengths and amplitudes and in random sequence. Lengths were determined as above. Amplitudes were determined from uniformly distributed pseudo random numbers generated by the GSL Mersenne Twister routine. The series were then digitally filtered with a low pass, eighth-order Bessel at 3 kHz cutoff (the series being "sampled" at 10 kHz) to simulate some filtering by the amplifier (or other recording device) and then at 250 Hz cutoff before CPD.

Statistical Measures

Statistical measures of normality are often based on the third and fourth order standardized central moments—skew and kurtosis, respectively. For a DSP (Fig. 1) consisting of the data samples $\{x_1, x_2, x_3, \ldots x_N\}$,

$$S = \gamma_3 \quad K = \gamma_4$$

$$\gamma_p = \mu_p/\mu_2^{p/2}$$

$$\mu_p = \frac{1}{N}\sum_i^{N}(x_i - \bar{x})^p$$

where S is the skew (or skewness); K is the kurtosis; γ_p is the standardized pth order central moment; μ_p is the pth order central moment and; $\bar{x}$ is the DSP mean. Moments were calculated with the single-pass online method of Pebay (2008). K has the value of three for a normal distribution. The excess kurtosis (K_E) is defined as $K-3$.

If several DSPs are taken from normally distributed data, then the values of S and K of these DSPs will be distributed nonnormally (Pearson 1931). However, transformations can be made to S and K so that the transformed values have a standard normal distribution (mean of zero, variance of one). Such transformed values are often written as Z(X) where X is the untransformed variable, the Z (or z) distribution being another name for the standard normal distribution. However, for simplicity of describing functions involving Z(X), we use X_Z instead. For the Z transforms of S and K we used the equations given by D'Agostino et al. (1990). These are quite lengthy, so we will not repeat them here. (See Eqs. 8–13 of D'Agostino et al. for the transform of skew, and Eqs. 14–19 for the transform of kurtosis.) Note that in their article, S and K are given by their rather dated terms, $\sqrt{b_1}$ and b_2, respectively.

Skew or kurtosis can be used alone to measure the normality of a distribution. However, there are two measures that combine them: the Jarque–Bera statistic (Jarque and Bera 1987) and the D'Agostino–Pearson omnibus statistic, or K^2 test (D'Agostino et al. 1990).

$$J = \frac{N}{6}\left(S^2 + \frac{K_E^2}{4}\right)$$

$$A = S_Z^2 + K_Z^2$$

where J is the Jarque–Bera statistic and; A is D'Agostino–Pearson omnibus statistic.

Change Point Detectors

Infinitely Good

This CPD works with simulated data. The idealized (unnoised) record is provided to the CPD and it uses this to detect all CPs.

Finitely Bad

Again this CPD works with simulated data. The idealized (unnoised) record is provided but, unlike the infinitely good CPD, only every nth CP is detected, where n is a random number between one and four picked for each DSP.

Derivative

Signal derivative has been used as a CPD for a number of single molecule studies (Carter et al. 2008; Tyerman et al. 1992; VanDongen 1996). By plotting the derivative against amplitude it can be used as a graphical method for summarizing levels (Tyerman et al. 1992), similar to mean–variance plots (see below). Signal derivative was calculated by the method of Savitzky and Golay (1964) (Steinier et al. 1972). The signal is convolved with a centered window, the coefficients of which are calculated as

$$\mathbf{A} = [a_{i,j}]_{i=0\ldots2h;j=0\ldots d}$$

$$a_{i,j} = (i - h)^j$$

$$\mathbf{B} = [b_{i,j}]_{i=0\ldots d;j=0\ldots 2h} = \left(\mathbf{A}^{\mathrm{T}}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}$$

$$\mathbf{w} = b_{1,*}$$

where h is the half width of the convolution window; d is the degree of the polynomial (the convolution is really the fit of a polynomial) and; $\mathbf{w}$ is the vector (window) of coefficients. The absolute value of the signal derivative constitutes the CPD signal (Fig. 1) and is thresholded to detect CPs.

Welch's t-Statistic

Welch's t-test is a form of Student's t-test applicable where the two samples to be compared may not have the same variance. It has been employed by a number of authors for CPD (Carter et al. 2008; Moghaddamjoo 1988; Pastushenko and Schindler 1997). In Moghaddamjoo (1988), it is given as the F test. Where the size of the two samples are equal the F statistic and Welch's t-statistic are equivalent with $F = t^2$. A window is passed across the series such that,

$$t = \frac{|m_1 - m_2|}{\sqrt{((V_1 + V_2)/w)}}$$

$$v = (w - 1)\frac{(V_1 + V_2)^2}{V_1^2 + V_2^2}$$

$$p = 1 - B_{v/(t^2+v)}\left(\frac{v}{2}, \frac{1}{2}\right)$$

where w is the width of the window in samples; m_1 and m_2 are the means of the first and last half of the window; V_1 and V_2 are the variances of the first and last half of the window; t is Welch's t-statistic; v is the number of degrees of freedom; $B_x(a, b)$ is the normalized incomplete beta function; and $1-p$ is the probability that a CP has not occurred (that the distribution of the first and last half of the window are the same). t or p can constitute the CPD signal.

Cochrane's Statistic

Variance alone can be used as a simple CPD signal. Also if window variance is plotted against window mean (mean–variance plot), a step increase in amplitude will describe an inverted parabola. The form of this parabola can be derived from first principles and so can be useful for quantitative analysis as well as qualitative assessment of levels (Patlak 1988, 1993; Thompson et al. 2002; Traynelis and Jaramillo 1998). Window variance can also be transformed to a probability that it is sampled from a population with a defined variance (in this context, the noise variance). This is based on the theory of Cochrane (Cochrane 1934, 1954). It is sometimes known as the χ^2 test for variance, but we use Cochrane's test for better distinction from other χ^2 tests (similar to using Welch's t-test rather than just t-test). As far as we aware it has only been used in one CPD study (Riessner et al. 2002). A window is passed across the series such that,

$$p = 1 - Q\left(\frac{w}{2}, \frac{V(w - 1)}{\sigma^2}\right)$$

where w is the width of the window in samples; V is the window variance; σ_2 is the population (noise) variance; $Q(a, b)$ is the normalized incomplete gamma function and; $1-p$ is the probability that a CP has not occurred (the probability that the window is sampled from a population with variance of σ_2).

Ion Channel Data

Maxi channels were recorded from in situ preparations of mouse small intestine as described previously (Parsons et al. 2012). The recordings were made at 10 kHz sampling frequency with the amplifier's low pass Bessel filter cutoff set to 5 kHz. All procedures were carried out in accordance with regulations from the Animal Ethics Board of McMaster University.

Results

Statistics of Noise

The assumption of models of single molecule data, and the basis of our idea for CPD assessment, is that each SP has a normal sample distribution. Put another way, a time series that is free of CPs should be normally distributed. Our first aim was to test this. Time series of pure noise were generated from two sources—an inside–out patch without any channel activity and a computer pseudo-random number generator set to generate white noise with a standard normal distribution. In some cases the series were low pass filtered at a

250 Hz cutoff as such filtering is an inevitable part of the recording and analysis of real single molecule data (of course the patch noise was already filtered to some extent by the recording system). Segments of random length, representing DSPs, were then sampled randomly and their statistical properties calculated. These were plotted as a function of DSP length (Fig. 2), as we wanted to get some statistical measure of normality that was invariant with length.

Two statistics are commonly used to assess normality: skew (S) and kurtosis (K). If we have a bell-shaped distribution, not necessarily normal by mathematical definition but of a similar shape to the undiscerning eye, then we might describe that bell in terms of its central bulge and two tails either side of this bulge. In these terms S is a measure of the symmetry of the bell, the relative size of the tails or any lean in the bulge, with an S of zero typically indicating perfect symmetry. A normal distribution has an

S of zero. K is a measure of the size of the tails relative to the bulge, K getting larger the tails or tail get. A normal distribution has an K of three. Therefore, $K-3$, the "excess kurtosis" (K_E), is often used for convenience.

The distribution of S as a function of DSP length, $S(L)$, was similar for both filtered and unfiltered, computer generated or patch generated, noise. The scatter of S decreased with L and was slightly greater with filtered than unfiltered data (Fig. 2). For unfiltered noise, both computer or patch generated, $K_E(L)$ was similarly distributed to $S(L)$. However, filtering of both noise series caused a marked asymmetric shift to negative K_E. The change in scatter (variance) of S and K with L would be a major impediment to using either quantity as a universal measure of normality—i.e., a single threshold of S or K could not be applied to DSPs of all lengths. However, there are formulas that, taking account of the number of samples (i.e., L), will
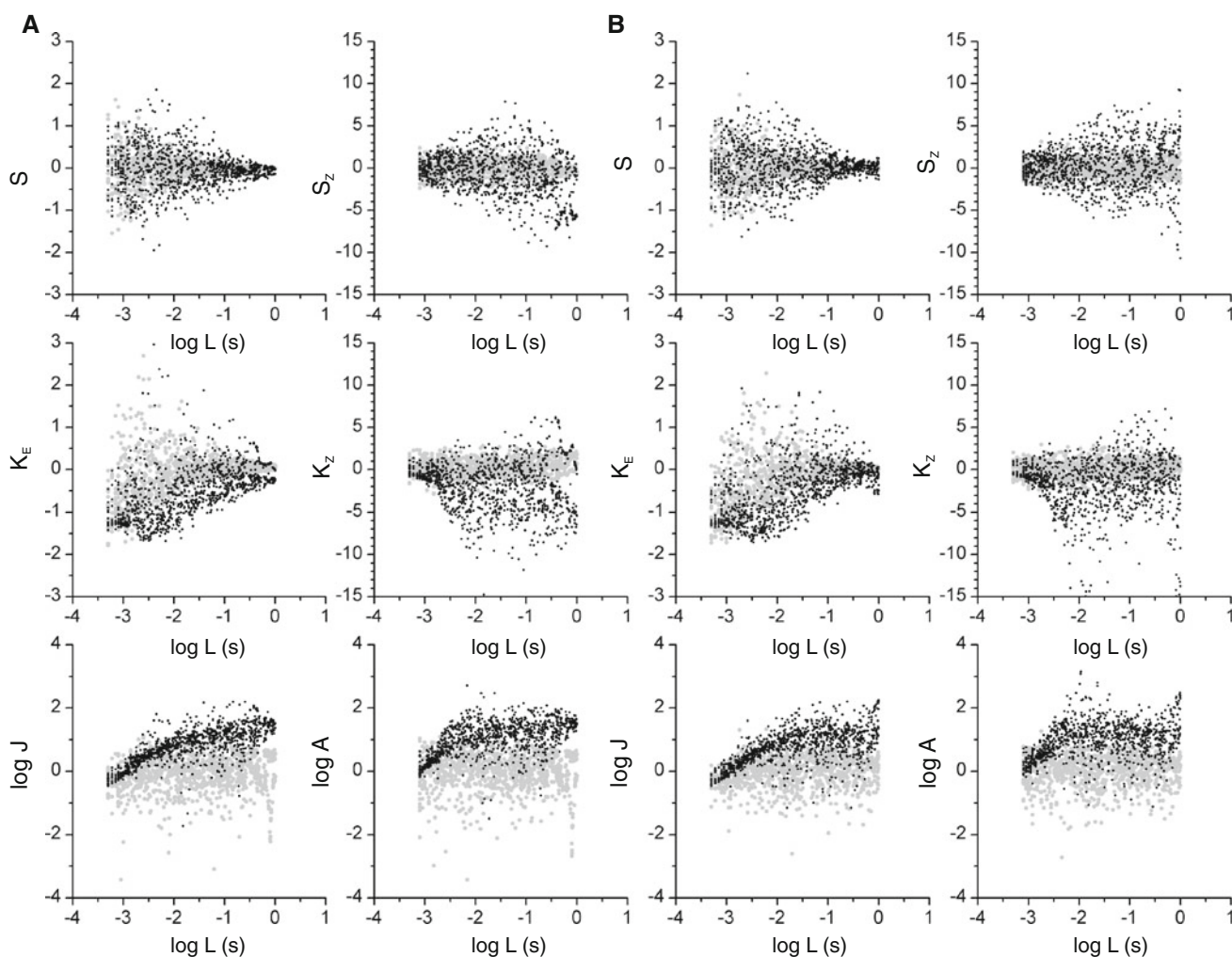

Fig. 2 Monte Carlo estimation of noise statistics. One thousand segments of random length were picked at random from a series of either **a** in silico simulated standard normal white noise or **b** noise recorded from an inside–out patch without ion channel activity (10 kHz sampling frequency, 5 kHz amplifier low pass). Scatter plots of segment statistic as a function of segment length (L). *Gray points* unfiltered series; *Black points* series digitally filtered with an eighth-order Bessel low pass with 250 Hz cutoff

transform a statistic such that it has a standard normal distribution (mean of zero and variance of one), irrespective of that number. As z (or Z) is another name for the standard normal distribution, these transforms are often called z-transforms and the transformed statistics, z-statistics. Equations have been given for z-transformation of both S and K. In our hands they seemed to work well. The scatter of both z-transformed S and K, S_Z and K_Z, did not vary with L. As with the untransformed statistics, S_Z was more widely distributed with the filtered data and the distribution of K_Z was shifted negative by filtering.

K and S can be combined to give a single measure of normality. Two common approaches of this kind are the Jarque–Bera statistic (J) and D'Agostino's K-squared statistic (we use A instead of the usual K^2 to avoid confusion). In both cases a value of zero indicates normality. With both J and A scatter did not change with L. With unfiltered data this scatter was centered around zero as may be expected. However, filtering caused both J and A to increase with L.

Statistics of DSPs from Simulated Data

With the statistics of noise established we could then compare this with the statistics of DSPs generated by CPD. We first did this with simulated data, not because this was required as a prerequisite to analyzing real data, but rather because simulation could test our analysis over a greater range of inputs—SPs of different length, amplitude and sequence—than might occur in a real data set. The same noise series as used before (patch or computer generated) were added to an idealized sequence of SPs of random length and amplitude in an entirely random sequence. (An example of simulated data is shown in Fig. 6a) The series were then low pass filtered, various CPD algorithms were applied and the statistics of the DSPs calculated. Because the statistics of pure noise (see above) suggested that excessive filtering introduced nonnormality we performed CPD on the filtered series, but used the unfiltered series to calculate each DSP's statistics. The same statistics were plotted as for the analysis of noise in Fig. 2.

We began by assessing two rather artificial CPDs, artificial in the sense that they were provided with the idealized sequence of SPs. The first, the infinitely good CPD, used the idealized sequence to detect all CPs perfectly—i.e., it produced only DCP(t) and DSP(t). The second, the finitely bad CPD, instead would only detect only every nth CP, n being a random number between one and four. In this way it produced a mix of DCP(t), DSP(f−) and DSP(t), but not DCP(f+) or DSP(f+). The infinitely good CPD produced distributions the same as for noise, as might be expected (Fig. 3, gray points). With the finitely bad CPD, the scatter of S, K_E, S_Z and K_Z increased with L (Fig. 3, black points). Also J and A increased with L. This can be explained by the fact

that as the number of CPs missed increases both the length of the DSP will increase and its nonnormality.

Given the dispersion of S, K_E, S_Z and K_Z with the finitely bad CPD, it was thought that it might be useful to plot K_Z against S_Z. Such plots (Fig. 4) showed a very characteristic distribution for the finitely bad CPD, rather like a swooping bird with its head at the origin, two wings extending into positive K_Z and a downwardly extended tail into negative K_Z. When the DSP(f−) and DSP(t) were distinguished (Fig. 4, bottom two panels) it could be seen that the DSP(t) were clustered around the origin, as might be expected, but that there were a number of DSP(f−) also around the origin that would not be distinguished from the DSP(t) by a simple threshold. These would occur where the CP(f−) was small enough that the DSP(f−) distribution appeared normal. For the infinitely good CPD all DSP were centered around the origin as would be expected (Fig. 4, gray points in top four panels).

The shape of the $K_Z(S_Z)$ distribution can be explained quite simply (Fig. 5a). The tips of the wings consist of DSP(f−) where one of the SPs is very short in length relative to the DSP—i.e., a short deflection. This gives a large tail to the DSP distribution, and therefore a large kurtosis. The fact that one tail is increased, also causes skew to be either negative (downward deflection, left wing) or positive (upward deflection, right wing). The tail of the bird consists of DSP(f−) with two SPs of near the same length so that skew is not too extreme. With the two SPs close in amplitude (i.e., a small CP(f−)) the DSP(f−)'s distribution will appear like a fat (large central bulge) bell, hence the negative kurtosis. At larger amplitude differences the effect on kurtosis is not so obvious as the DSP distribution becomes bimodal. Therefore, we simulated a DSP(f−) of fixed length with two SPs, the amplitude and length difference of which varied systematically. The $K_Z(S_Z)$ plot of this data (Fig. 5b) confirmed our interpretation of the wings and tail and showed that for the tail K_Z decreases almost infinitely with an increase in amplitude difference.

In addition to the artificial CPDs (infinitely good and finitely bad) we assessed three CPDs commonly used for real single molecule data—derivative, Welch's t-test and Cochrane's test. Welch's t-test appeared overly sensitive as could be seen from the low signal to noise ratio of its CPD signal (Fig. 6c). As a result it broke the SPs up into many short DSP(f+) and produced very few DSP(f−) (Fig. 6c, Table 1). Cochrane's test had a CPD signal with very good signal to noise ratio (Fig. 6d) so there were very few DSP(f+) (Table 1). However, there were a substantial number of DSP(f−) (Table 1) and these were not distinguishable from DSP(t) in $K_Z(S_Z)$ plots, suggesting that the CP(f−) were small (Fig. 6d). Similar to Cochrane's test, the derivative had a fairly high signal-to-noise ratio (though not as good; Fig. 6b) and so again had few
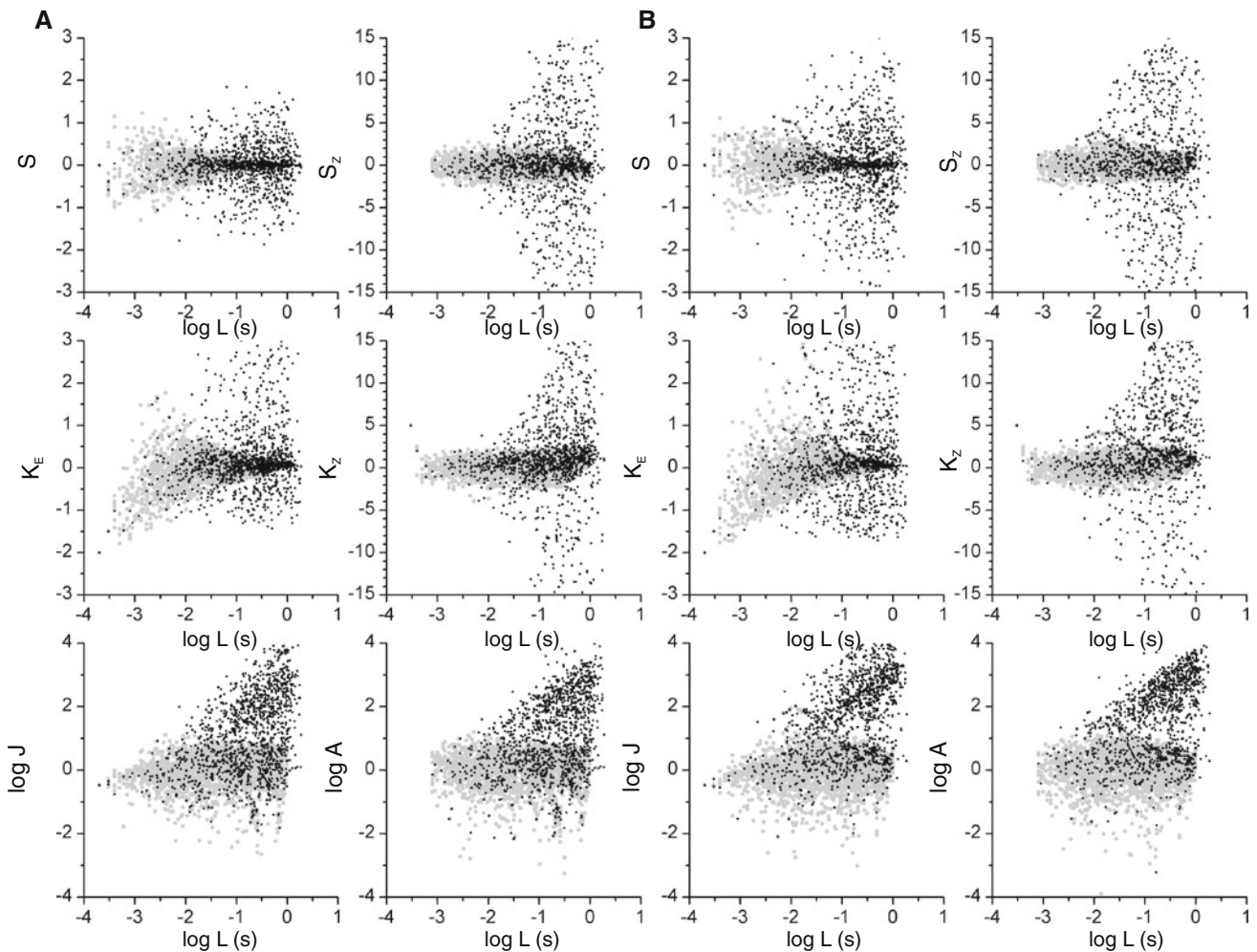
Fig. 3 Performance of the infinitely good CPD and finitely bad CPD with simulated SPs. 4010 SPs of random length and amplitude were simulated and then added to either **a** in silico simulated standard normal white noise (4.5 kHz digital low pass) or **b** noise recorded from an inside–out patch without ion channel activity (10 kHz sampling frequency, 5 kHz amplifier low pass). Scatter plots of DSP statistic as a function of segment length (L). *Gray points* infinitely good CPD; *Black points* finitely bad CPD

DSP(f+) (Table 1). It had a similar percentage of DSP(f−) to Cochrane's (Table 1) and to a certain extent these were more discernible in $K_Z(S_Z)$ plots (Fig. 6b).
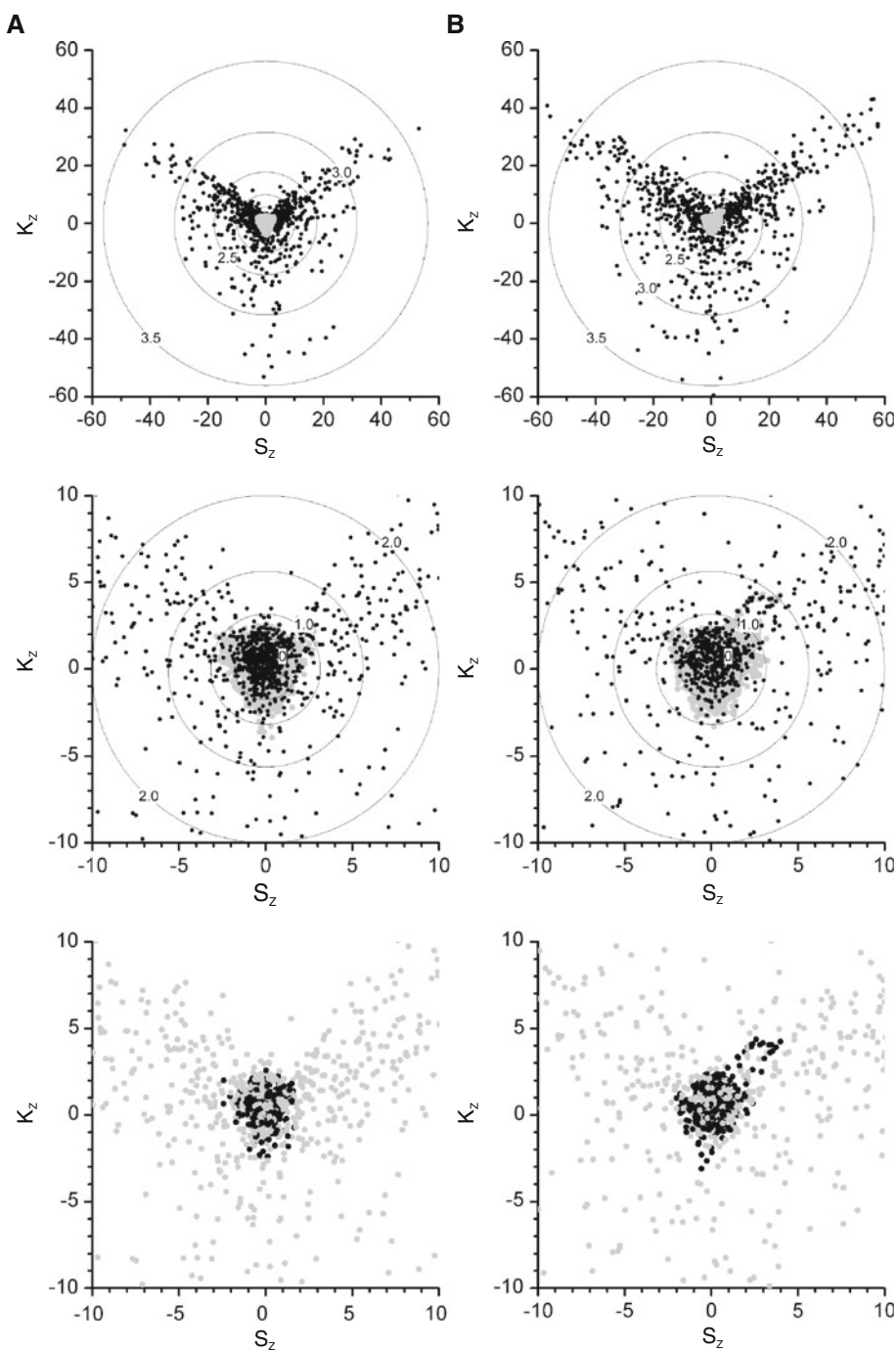
Statistics of DSPs from Real Data

The aim of the previous analysis of noise and simulated data was to define what the statistics can and cannot show and how best to apply them. We could then confidently apply the statistics to real data where we had no idealization to hand. One particular set of real data we were interested in were some patch-clamp recording of single maxi channels. These are large conductance ion channels (hence the name) found in a diverse range of cell types (Parsons et al. 2012; Sabirov and Okada 2009). They have a complex subconductance behavior, there is no obvious model for visual assessment of CPD fitting, making them perfect for our statistical approach.

As for the simulated data (Fig. 6), Welch's *t*-test produced many short DSP compared to the derivative and Cochrane's test (Fig. 7). However, unlike the simulated data, the DSP statistic distributions (whatever the CPD) were much more like those from the finitely bad CPD. D'Agostino's *K*-squared statistic increased with length (compare Fig. 6 to Fig. 3) and the $K_Z(S_Z)$ plots had definite wings (Fig. 7).

It may be noticed that in the simulated data analyzed by finitely bad CPD (Figs. 4, 5b), the $K_Z(S_Z)$ tail is comparable in extent to the wings, whereas with real data analyzed by nonartificial CPD (Fig. 7) the wings and tail are of comparable length. There is a simple explanation for this. A nonartificial CPD is extra prone to nondetection of short deflections (wing errors) due to the finite window length of the detector, in comparison to longer deflections (tail errors) of the same amplitude. Thus, the wings are bigger than the tail. In comparison the finitely bad CPD does not

Fig. 4 Performance of the infinitely good CPD and finitely bad CPD with simulated SPs. As in Fig. 3, 4010 SPs of random length and amplitude were simulated and then added to either **a** in silico simulated standard normal white noise (4.5 kHz digital low pass) or **b** noise recorded from an inside–out patch without ion channel activity (10 kHz sampling frequency, 5 kHz amplifier low pass). Scatter plots of DSP normalized kurtosis (K_Z) as a function of normalized skew (S_Z). *Upper four panels* performance of infinitely good against finitely bad CPD. *Gray points* infinitely good CPD. *Black points* finitely bad CPD. Contours indicate log of D'Agostino–Pearson omnibus statistic. *Lower two panels* DSP designation with finitely bad CPD. *Gray points* DSP(f−); *black points* DSP(t)

distinguish between short and long deflections, so wings and tail are of similar extent.

Discussion

Assumption of Normal Noise

It was surprising to us that filtering degraded the normality of noise. We had no expectation that the distribution of

noise would be anything other than Gaussian. That this seems to be a common assumption is reflected by three facts. First, this is often stated as an assumption (Pastushenko and Schindler 1997; Schultze and Draber 1993). Second, the simulations used to test CPDs have often (though probably not in the majority) used unfiltered Gaussian white noise (Moghaddamjoo 1988; Pastushenko and Schindler 1997; Schultze and Draber 1993). Third, many CPDs are parametric—that is, they are based on the assumption of a particular distribution, in this case
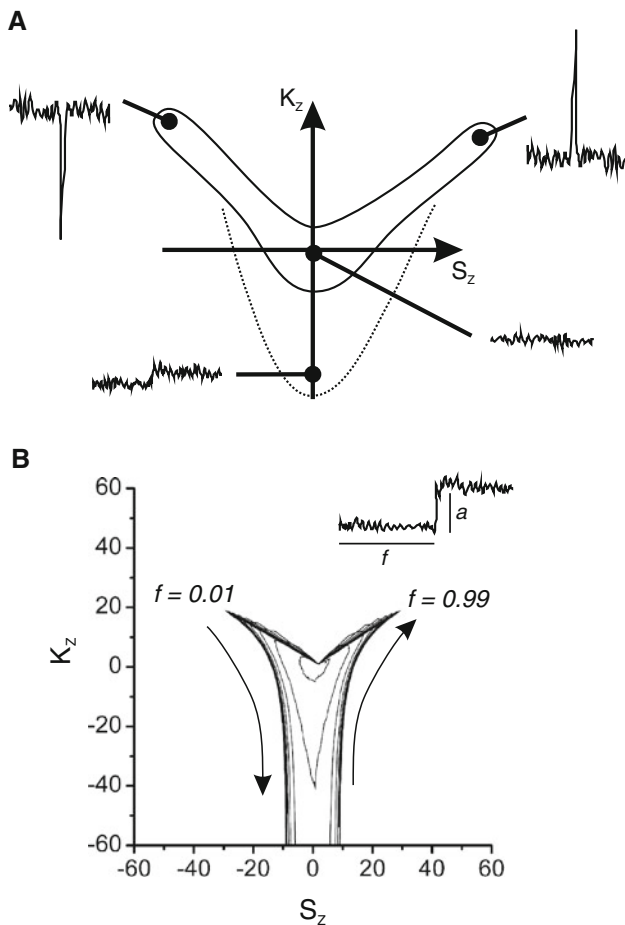
Fig. 5 Interpretation of $K_Z(S_Z)$ plots. **a** Simulation. To simulate a DSP, a thousand samples were taken from noise recorded from an inside–out patch without ion channel activity (10 kHz sampling frequency, 5 kHz amplifier low pass). To simulate a DSP(f−) a CP was added a fraction of time (*f*) into the DSP with a particular increase in amplitude (*a*). In the $K_Z(S_Z)$ plot each contour line represents a particular CP amplitude (increasing from 1 to 10, inner to outer) with the fraction increasing along the line (*arrows*). **b** Schematic. The two limbs at positive kurtosis represent DSP(f−) with short deflections (relative to the length of the DSP). This deflection increases the length of either the left tail of the DSP's distribution (negative deflection, negative skew limb) or right tail (positive deflection, positive skew limb) and so increases kurtosis (the DSP is leptokurtic). As the size of the deflection decreases, both skew and kurtosis decrease at the same rate. The bough at negative kurtosis represent DSP(f−) with a small amplitude CP(f−). This causes a fat (platykurtic) DSP distribution

Gaussian. Although the cumulative sum can be used directly as a CPD signal, most commonly it is put into the probabilistic form of a likelihood ratio test based on Gaussian distributions (Basseville 1988). Also both Welch's *t*-test and Cochrane's test assume a Gaussian distribution. This perhaps explains an observation of Gross and colleagues (Carter et al. 2008) that Welch's *t*-test performed variably depending on the filtering of the data, in comparison to other nonparametric CPDs they examined.

There are a multitude of examples of "non-Gaussian noise" in electronics, optics, acoustics, communication and other realms of applied physics. Also there is significant recognition of nonnormal noise in the field of digital signal processing (Kassam and Thomas 1988; Webster 1993). There are a number of possible causes for nonnormality of noise. Firstly step changes in amplitude, what we regard here as "signal," is called "impulsive" noise in the fields of electronics and communication. This is a rather obvious source of nonnormality, causing a lengthening of tails when the impulses are short or a bimodal distribution when they are longer (Fig. 5). It is the basis of our assessment of DSPs. However, there are two less obvious other sources of nonnormality. The assumption of Gaussian noise has its mathematical basis in the central limit theorem. Here each noise sample is assumed to represent the average of a large number of random processes. For example in a patch-clamp recording, different components of the amplifier's circuits, electromagnetic radiation, movement of the patch and thermal fluctuations of the channel itself. According to the central limit theorem if these processes all have the same distribution, or in some cases just the same variance, then their average (i.e., the measured noise) will have a normal distribution. However, this may not be the case. As was stated by Webster (1993), this condition (identical distribution or same variance) is not likely to be met in a natural environment. The second, less obvious cause of nonnormality, as we found here, is filtering. That filtering of white noise (that is uncorrelated noise with a flat frequency spectrum) will color that noise (make it correlated with a nonflat frequency spectrum), seems to be a well appreciated fact. However, it seems to be less well appreciated that filtering of Gaussian white noise will also induce nonnormality. This is perhaps because of the relative sparsity and technical nature of the literature and that there is no simple relation between filter and distribution of output (Rice 1945; Webster 1994; Wolff 1967).

DSP(f+), DSP(f−) and Filtering Versus Refinement

By design our method does not detect DSP(f+). These will increase in CPDs with a low CPD signal to noise ratio (Fig. 6, Table 1). Therefore, much can be done by just observing the CPD signal and selecting a CPD with a the highest signal to noise ratio. However, there is probably no CPD, which can completely avoid DCP(f+) and to "solve" the problem a more thorough analysis will be needed of what causes DCP(f+) in individual CPD.

Where the CP(f−) is small enough in amplitude, the DSP(f−) can be undistinguishable from a DCP(*t*) based on the statistical measures of normality we have used (Figs. 4, 6). It may be possible that a more sensitive test of normality could identify these DSP(f−). Sensitivity might be increased with a parametric test, i.e., based on comparison to an actual
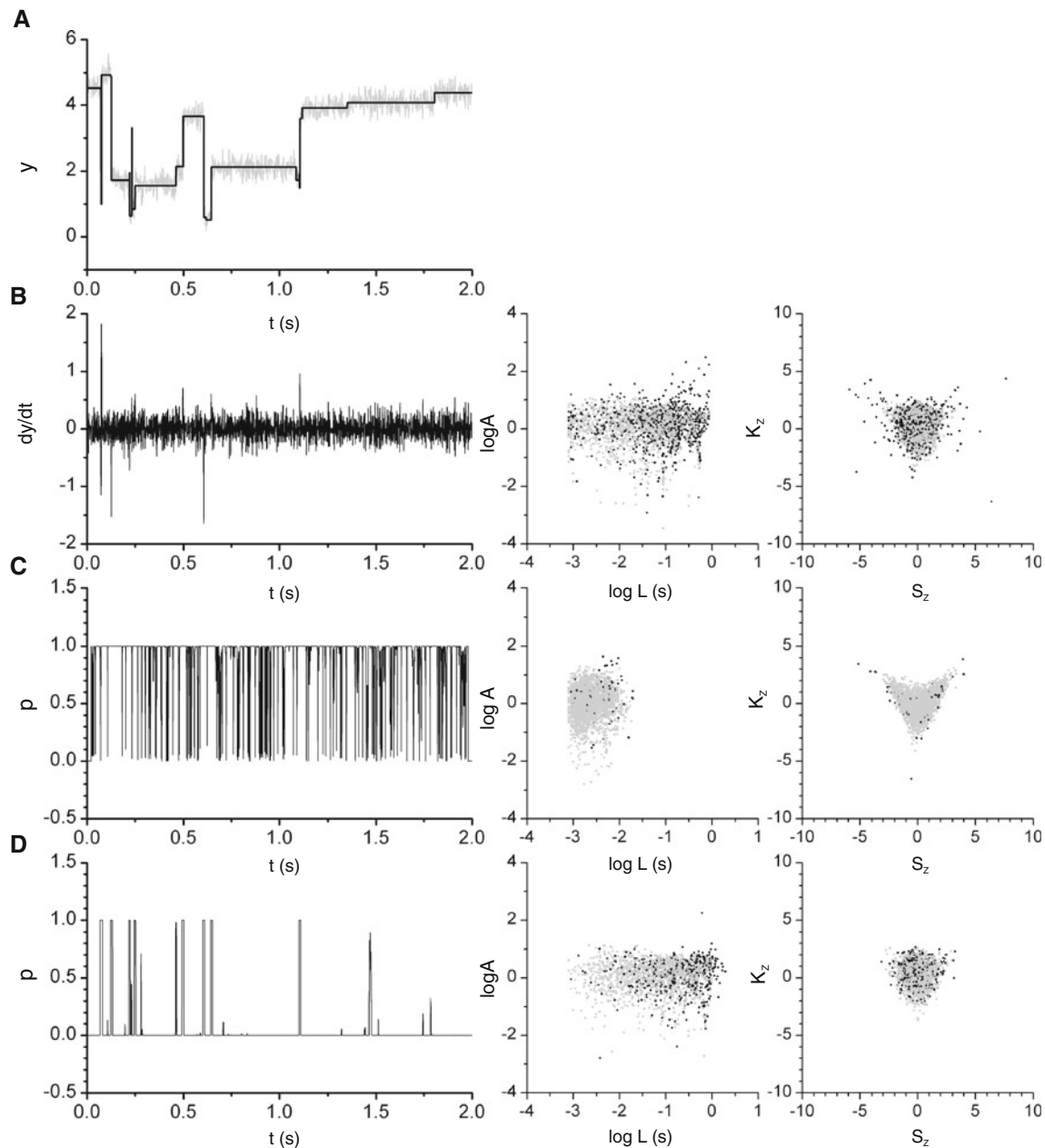
Fig. 6 Performance of nonartificial CPDs with simulated SPs. As in Figs. 3 and 4, 4010 SPs of random length and amplitude were simulated and then added to in silico simulated standard normal white noise. The series was then digitally low pass filtered twice (4.5 kHz and 250 Hz cutoff) the first to simulate filtering by a recording instrument. CPD was performed on the doubly filtered data, but DSP statistics were calculated from the singly filtered series. **a** Part of the simulated series with (*gray*) and without (*black*) doubly filtered noise. **b–d** CPD signal and DSP statistics for derivative CPD (**b**), Welch's *t*-test CPD (**c**) and Cochrane's test CPD (**d**). In statistics plots, *black points* are DSP(f−) and *gray points* all other DSP

Gaussian distribution rather than some value that just indicates normality, like skew and kurtosis. There are several parametric tests of normality based on binning the data so its distribution can be compared directly to a normal (e.g., Pearson's χ^2 test, the Kolmogorov–Smirnov test, Cramér–von Mises statistic). Binning may limit their applicability to short DSP, but it is a direction worth exploring.

Once DSP(f−) are identified there is then the question of what to do with them. One approach is to discard them from further analysis. However, this could leave a large gap, especially as DSP(f−) tend to be long (Fig. 3). A better approach than this filtering of DSP would be to take each DSP(f−) and try to identify the CP(f−) within it by a CPD that is more sensitive than the CPD used to first segment the

Table 1 Properties of DSPs from different CPDs with simulated data[a]

CPD	N	t (%)	f− (%)	f+ (%)	f−, f+ (%)
Infinitely good	2,642	100.0	0.0	0.0	0.0
Finitely bad	1,274	24.0	76.0	0.0	0.0
Derivative	3,407	29.9	26.8	59.6	16.3
Welch's t-statistic	27,806	2.0	0.4	97.9	0.3
Cochrane's statistic	1,963	73.9	24.1	3.8	1.8

CPD change point detector

[a] All CPDs were given the same simulated data. Derivative: $h = 7$, $d = 2$ threshold $(dy/dt) = 0.5$. Welch's t-statistic: $w = 200$, threshold $(p) = 0.95$; Cochrane's statistic: $w = 100$, $\sigma^2 = 0.2$, threshold $(p) = 0.95$

series—either a different CPD altogether or the same CPD but with adjusted parameters. Again this CPD would have to be not too sensitive at the risk of multiplying DSP(f+).

Previous Distribution Techniques

A study by Hansen and colleagues came very close to the method we have presented here (Schroder et al. 2004). Their initial reasoning was exactly the same, that SP distributions should be Gaussian, but DSPs may not be Gaussian if CPs are missed. However, their approach was somewhat different. Instead of looking at the distributions of individual DSPs they looked at the distributions of levels, that is sets of DSPs with the same amplitude (determined in their case with cumulative sum CPD algorithms). There are (usually) far fewer levels in a series, than DSP; thus, the level distributions could be assessed by directly graphing them and fitting with Gaussian functions. Their method might be described as aggregated (in respect of the DSPs) as opposed to our unaggregated method. The un-aggregated approach has three advantages in this respect. From a modeling point of view it does not assume a model of DSPs aggregated into levels. This may be an appropriate position to take when dealing with apparently complex kinetics such as for the maxi channel or from a more theoretical stance (Frauenfelder et al. 1988). More practically, most CPDs do not aggregate DSPs into levels. Cumulative sum CPDs only do this because the level amplitudes are input parameters to the algorithm. This makes the unaggregated approach a better fit with most CPDs. Lastly once a level distribution is judged as nonnormal, it will probably not be possible to filter out the "wrong" samples. If a nonnormal level distribution can be thought of as the sum of a normal distribution of "correct" samples and a nonnormal distribution of wrong samples (i.e., a mixture model) than if those two distribution overlap there is no way of assigning a particular sample in that overlap to a particular distribution. The only option is to start over CPD with different parameters.

Another precedent for our method has been the application of beta distributions to single ion channel data (Fitzhugh 1983; Yellen 1984). These can be used to model sample distributions during flicker—i.e., a two level series with time spent in one level very short compared to the other. This is the situation illustrated in Fig. 5b at the wing tips of the $K_Z(S_Z)$ plot. The short transitions to the second level causes a long tail in the overall distribution, most of which reflects samples of the first level. By mathematical analysis it was shown that such a distribution should be a beta distribution (Fitzhugh 1983) and that an analysis based on this could be used to calculate kinetic parameters that might be inaccurately determined by other methods, due to filtering of short transitions. The moments of the beta distribution, including skew and kurtosis, have been derived and so it might be an interesting direction to make a analytical comparison with our method.

A Practical Guide to Statistical Assessment

Test Assumptions

Our assessment method is based on assumption of a model of the data where

(i) The data can be described entirely with CPs, which are near instantaneous, and SPs, which are stationary in the strict sense—i.e., the distribution is time invariant.
(ii) The SP distribution is Gaussian.

Therefore, the first priority is to test whether these assumptions hold. The first assumption should be tested first by visual inspection of the data. Any obvious non-step-like changes in current? Such changes are common enough in ion channel recordings. "Flicker" may result from fast bandwidth limited between multiple conductance states. Slower changes may result from "diffusional" changes in occupancy of states with similar conductance (Millhauser et al. 1988; Vaccaro 2007) or rearrangements of the patch membrane (Chui and Fyles 2012; Fyles et al. 1998). If such changes are common, it may well do to exclude this data from analysis. The second assumption requires two consecutive tests:

(i) Test channel-free patch recordings, which have not been digitally filtered, for normality.
(ii) It has been established that channels contribute their own "noise" to SPs through tiny movements of the channel, tiny variations in their ion flux or current due to movement of charged residues. I put "noise" in quotation marks because of course there may be a continuum between noise and bona fide CP—it could be seen as a matter of definition. However, such "open channel noise" was extensively studied by Sigworth and colleagues in a classic series of papers in the Biophysical Journal (Crouzy and Sigworth
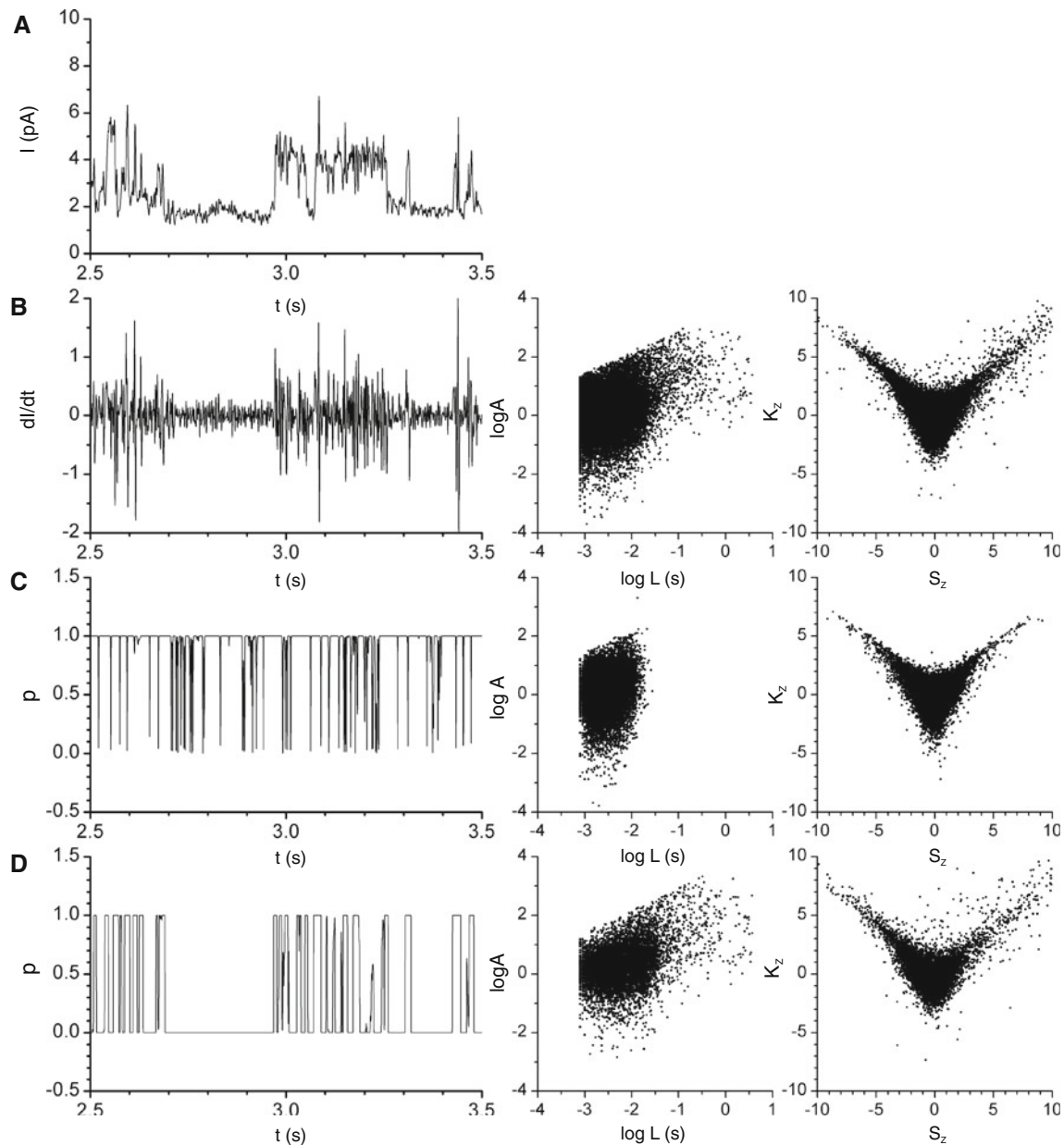
Fig. 7 Performance of CPDs with a patch-clamp recording of a maxi channel. The recording was 6 min long. It was low pass filtered by the amplifier at 5 kHz and low pass filtered digitally at 250 Hz. CPD was performed on the doubly filtered data, but DSP statistics were calculated from the singly filtered series. **a** Part of the doubly filtered series. **b–d** CPD signal and DSP statistics for derivative CPD (**b**), Welch's *t*-test CPD (**c**) and Cochrane's test CPD (**d**)

1993; Heinemann and Sigworth 1988, 1990, 1991; Sigworth 1985, 1986; Sigworth et al. 1987). In one of these, they looked at noise distributions (Heinemann and Sigworth 1991) and showed that noise could be non-Gaussian under certain ionic conditions. Therefore, even if instrumental noise has been established as Gaussian (i), SPs may still be nonnormal. Probably this can only be tested by plotting $K_Z(S_Z)$ and if no DSPs appear normal, then perhaps this is more than just a case of poor CPD and the ion channel contributes non-Gaussian noise.

CPD and Statistics

Once the assumptions in (i) are tested, CPD and statistical analysis should be carried out with the following guidance.

(i) CPD can be performed on filtered data, but DSP statistics should be calculated from the unfiltered data.

(ii) Only nonparametric CPDs should be used with filtered data.

(iii) The choice of specific CPD and its parameters should be informed either by considerations of the

raw data (e.g., faster kinetics would recommend a smaller window or a derivative CPD) or empirically by comparison of DSPs from several CPD. Use whatever CPD gives the least presumed DSP(f−) (nonnormal DSP) as a proportion of all DSP.

What to Do with the Statistics

Once CPD has been performed and $K_Z(S_Z)$ plots made, the proportion of presumed DSP(f−) and their type (wing or tail) can be assessed. Three things then seem possible with presumed DSP(f−).

(i)   If they are wing DSP(f−), suggesting fast deflections, the window size might be reduced.
(ii) Exclude all of them from further analysis.
(iii) Try to identify the CP(f−) in them by more sensitive analysis or by reapplying CPD with more sensitive parameters.

We haven't looked at these possibilities. It is probably rather heuristic and would require a much more detailed analysis of different CPD. Here we have only attempted to outline a general method of assessment.

Reference

Basseville M (1988) Detecting changes in signals and systems—a general survey. Automatica 24:309–326

Carter BC, Vershinin M, Gross SP (2008) A comparison of step-detection methods: how well can you do? Biophys J 94:306–319

Chui JKW, Fyles TM (2012) Ionic conductance of synthetic channels: analysis, lessons, and recommendations. Chem Soc Rev 41:148–175

Cochrane WG (1934) The distribution of quadratic forms in a normal system with applications to the analysis of covariance. Math Proc Camb Philos Soc 30:178–191

Cochrane WG (1954) Methods for strengthening the common χ^2 tests. Biometrics 10:417–451

Crouzy SC, Sigworth FJ (1993) Fluctuations in ion channel gating currents. Analysis of nonstationary shot noise. Biophys J 64:68–76

D'Agostino RB, Belanger A, D'Agostino RBJ (1990) A suggestion for using powerful and informative tests of normality. Am Stat 44:316–321

Fitzhugh R (1983) Statistical properties of the asymmetric random telegraph signal, with applications to single-channel analysis. Math Biosci 64:75–89

Frauenfelder H, Parak F, Young RD (1988) Conformational substates in proteins. Annu Rev Biophys Biophys Chem 17:451–479

Fyles TM, Loock D, Zhou X (1998) Ion channels based on bis-macrocyclic bolaamphiphiles: effects of hydrophobic substitutions. Can J Chem 76:1015–1026

Heinemann SH, Sigworth FJ (1988) Open channel noise. IV. Estimation of rapid kinetics of formamide block in gramicidin A channels. Biophys J 54:757–764

Heinemann SH, Sigworth FJ (1990) Open channel noise. V. Fluctuating barriers to ion entry in gramicidin A channels. Biophys J 57:499–514

Heinemann SH, Sigworth FJ (1991) Open channel noise. VI. Analysis of amplitude histograms to determine rapid kinetic parameters. Biophys J 60:577–587

Jarque CM, Bera AK (1987) A test for normality of observations and regression residuals. Int Stat Rev 55:163–172

Kassam SA, Thomas JB (1988) Signal detection in non-gaussian noise. Springer-Verlag, New York

Millhauser GL, Salpeter EE, Oswald RE (1988) Diffusion models of ion-channel gating and the origin of power-law distributions from single-channel recording. Proc Natl Acad Sci USA 85:1503–1507

Moghaddamjoo A (1988) Step-like signal processing with distinct finite number of levels. IEEE Trans Ind Electron 35:489–493

Parsons SP, Kunze WA, Huizinga JD (2012) Maxi-channels recorded in situ from ICC and pericytes associated with the mouse myenteric plexus. Am J Physiol Cell Physiol 302:C1055–C1069

Pastushenko VP, Schindler H (1997) Level detection in ion channel records via idealization by statistical filtering and likelihood optimization. Philos Trans R Soc Lond B Biol Sci 352:39–51

Patlak JB (1988) Sodium channel subconductance levels measured with a new variance–mean analysis. J Gen Physiol 92:413–430

Patlak JB (1993) Measuring kinetics of complex single ion channel data using mean–variance histograms. Biophys J 65:29–42

Pearson ES (1931) Note on tests for normality. Biometrika 22:423–424

Pebay P (2008) Formulas for robust, one-pass parallel computation of covariances and arbitrary-order statistical moments. Sandia National Laboratories, Albuquerque

Rice SO (1945) Statistical properties of random noise currents. Bell Syst Tech J 24:46–156

Riessner T, Woelk F, Abshagen-Keunecke M, Caliebe A, Hansen UP (2002) A new level detector for ion channel analysis. J Membr Biol 189:105–118

Sabirov RZ, Okada Y (2009) The maxi-anion channel: a classical channel playing novel roles through an unidentified molecular entity. J Physiol Sci 59:3–21

Sansom MSP, Ball FG, Kerry CJ, Mcgee R, Ramsey RL, Usherwood PNR (1989) Markov, fractal, diffusion, and related models of ion channel gating—a comparison with experimental-data from 2 ion channels. Biophys J 56:1229–1243

Savitzky A, Golay MJE (1964) Smoothing and differentiation of data by simplified least squares procedures. Anal Chem 36:1627–1639

Schroder I, Huth T, Suitchmezian V, Jarosik J, Schnell S, Hansen UP (2004) Distributions-per-level: a means of testing level detectors and models of patch-clamp data. J Membr Biol 197:49–58

Schultze R, Draber S (1993) A nonlinear filter algorithm for the detection of jumps in patch-clamp data. J Membr Biol 132:41–52

Sigworth FJ (1985) Open channel noise. I. Noise in acetylcholine receptor currents suggests conformational fluctuations. Biophys J 47:709–720

Sigworth FJ (1986) Open channel noise. II. A test for coupling between current fluctuations and conformational transitions in the acetylcholine receptor. Biophys J 49:1041–1046

Sigworth FJ, Sine SM (1987) Data transformations for improved display and fitting of single-channel dwell time histograms. Biophys J 52:1047–1054

Sigworth FJ, Urry DW, Prasad KU (1987) Open channel noise. III. High-resolution recordings show rapid current fluctuations in gramicidin A and four chemical analogues. Biophys J 52:1055–1064

Steinier J, Termonia Y, Deltour J (1972) Comments on smoothing and differentiation of data by simplified least square procedure. Anal Chem 44:1906–1909

Thompson RJ, Nordeen MH, Howell KE, Caldwell JH (2002) A large-conductance anion channel of the Golgi complex. Biophys J 83:278–289

Traynelis SF, Jaramillo F (1998) Getting the most out of noise in the central nervous system. Trends Neurosci 21:137–145

Tyerman SD, Terry BR, Findlay GP (1992) Multiple conductances in the large K^+ channel from Chara corallina shown by a transient analysis method. Biophys J 61:736–749

Vaccaro SR (2007) Nonlinear drift–diffusion model of gating in the fast Cl channel. Phys Rev E Stat Nonlin Soft Matter Phys 76:011923

VanDongen AM (1996) A new algorithm for idealizing single ion channel data containing multiple unknown conductance levels. Biophys J 70:1303–1315

Webster RJ (1993) Ambient noise statistics. IEEE Trans Signal Proc 41:2249–2253

Webster RJ (1994) A random number generator for ocean noise statistics. IEEE J Ocean Eng 19:134–137

Wolff SS (1967) On probability distributions for filtered white noise. IEEE Trans Inf Theory 13:481–484

Yellen G (1984) Ionic permeation and blockade in Ca^{2+}-activated K^+ channels of bovine chromaffin cells. J Gen Physiol 84:157–186