# Data Analysis of Translocation Events in Nanopore Experiments

**Daniel Pedone, Matthias Firnkes, and Ulrich Rant***

*Walter Schottky Institut, Technische Universität München*

**Nanopores have become important tools for single molecule experiments, where information about the properties of DNA/RNA or proteins is inferred from current pulses elicited by individual molecules as they traverse a single pore. However, because of necessary electronic filters employed in the measurement technique, the extraction of meaningful information from short pulses is limited. This restricts the use of nanopores for the investigation of small molecules which cross the pore rapidly. Here we present a method which significantly improves the accuracy of the analysis of noise-filtered current pulses. We introduce improved criteria to measure the pulse width and propose a method to evaluate the pulse height from the falling edge of the pulse, which renders the identification of a pulse plateau unnecessary. The new methods are compared to conventional routines and validated by analyzing representative current pulses as well as experimental protein translocation data. It is demonstrated that the pulse properties can be recovered with satisfying accuracy beyond the usual limitations of Bessel filters, i.e., from pulses featuring a width of merely $0.3f_c^{-1}$ ($f_c$ being the filter cutoff frequency).**



Over the last years nanopores have evolved into powerful and indispensable devices for the investigation of single molecules.[1] In these experiments, biological[2−7] or solid-state[6,8−10] pores with dimensions similar to the size of biological macromolecules (DNA, proteins) are used to scrutinize the state of individual molecules as they travel through a solitary pore which connects two fluid compartments separated by a membrane. The Coulter method[11] is a convenient technique to monitor the passage of individual molecules through the pore by electrical means: In aqueous electrolyte solutions, an ionic current is driven through the pore by applying a voltage between electrodes placed on either side of the pore. When a molecule traverses the pore, it generates a current pulse, that is, it transiently disturbs the ionic current flowing through the pore, which can be detected with external measurement electronics. Figure 1 A shows examples of typical pulses (blockades) measured during the translocation of DNA and protein molecules through artificially engineered pores in $Si_4N_3$ membranes. A well resolved square-shaped pulse is observed for the translocation of DNA, whereas diminutive pulses are usually found in experiments with proteins (similar pulse

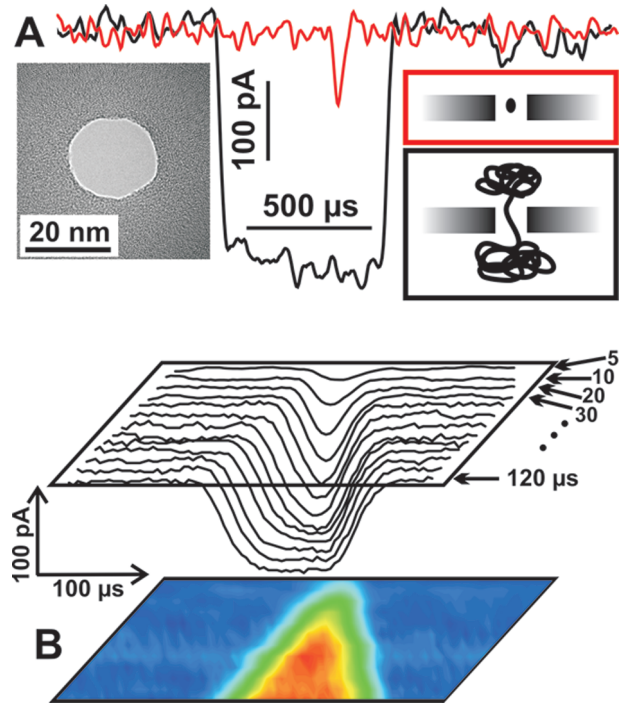


**Figure 1.** DNA and protein translocation through nanopores and the influence of filtering on measured current pulses. (A) Representative current pulses caused by individual avidin (red) and λ-DNA (black) molecules translocating through a solid-state nanopore. The TEM image shows a nanopore in a $Si_3N_4$ membrane used for the avidin translocation. Panel B depicts the attenuation and dilation of square pulses created by a function generator when filtered with a seven-pole $f_c = 10$ kHz Bessel filter. The nominal pulse height was 100 pA, and the nominal widths were varied from 5 to 120 µs.

* To whom correspondence should be addressed. Phone: +49 89 289 12778. Fax: +49 89 320 6620. E-mail: rant@wsi.tum.de.

(1) Howorka, S.; Siwy, Z. *Chem. Soc. Rev.* **2009**, *38*, 2360–2384.
(2) Kasianowicz, J. J.; Brandin, E.; Branton, D.; Deamer, D. W. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 13770–13773.
(3) Bayley, H.; Braha, O.; Gu, L. Q. *Adv. Mater.* **2000**, *12*, 139–142.
(4) Meller, A. *J. Phys.: Condens. Matter* **2003**, *15*, R581–R607.
(5) Nakane, J. J.; Akeson, M.; Marziali, A. *J. Phys.: Condens. Matter* **2003**, *15*, R1365–R1393.
(6) Zwolak, M.; Di Ventra, M. *Rev. Mod. Phys.* **2008**, *80*, 141–165.
(7) Branton, D.; Deamer, D. W.; Marziali, A.; Bayley, H.; Benner, S. A.; Butler, T.; Di Ventra, M.; Garaj, S.; Hibbs, A.; Huang, X. H.; Jovanovich, S. B.; Krstic, P. S.; Lindsay, S.; Ling, X. S. S.; Mastrangelo, C. H.; Meller, A.; Oliver, J. S.; Pershin, Y. V.; Ramsey, J. M.; Riehn, R.; Soni, G. V.; Tabard-Cossa, V.; Wanunu, M.; Wiggin, M.; Schloss, J. A. *Nat. Biotechnol.* **2008**, *26*, 1146–1153.
(8) Li, J.; Stein, D.; McMullan, C.; Branton, D.; Aziz, M. J.; Golovchenko, J. A. *Nature* **2001**, *412*, 166–169.
(9) Dekker, C. *Nat. Nanotechnol.* **2007**, *2*, 209–215.
(10) Vlassiouk, I.; Kozel, T. R.; Siwy, Z. S. *J. Am. Chem. Soc.* **2009**, *131*, 8211–8220.
(11) Coulter, W. H. U.S. Patent No. 2,656,508, 1953.

shapes have been observed by other authors[12−14]). In order to obtain statistically relevant information from many single molecule events, it is necessary to discriminate pulses from a noisy background and to analyze them with respect to the pulse width ($w$) and pulse height ($h$) in an automated way.

In past experiments, which were predominantly focused on the translocation of DNA and RNA molecules through α-hemolysin, computer-aided data analysis has not posed a substantial problem because of long-lasting events (>100 $\mu$s) and good signal-to-noise ratios.[15−17] Recently, short translocation events have attracted considerable attention because solid-state nanopores are progressively used to study comparably small molecules, e.g., proteins[12−14,18,19] or short DNA and RNA fragments,[20,21] which pass through the pore rapidly (approximately microseconds). However, short pulses are difficult to detect because current noise increases as the measurement bandwidth is enhanced. To suppress high frequency noise, electronic filters have to be employed, in particular when using solid-state nanopores which exhibit a considerably larger noise than biological pores like α-hemolysin.[22] Electronic filtering affects the measurement as it distorts the signal in two ways: (i) short pulses appear dilated and (ii) the pulse height gets reduced.[23,24] This is shown in Figure 1B, which depicts square pulses that were low-pass filtered with a seven-pole Bessel filter featuring a cutoff frequency (−3 dB attenuation) of $f_c = 10$ kHz ($f_c^{-1} = 100\ \mu$s). Note that a 10 kHz filter has been chosen exemplarily in this work, as this filter frequency is commonly used in nanopore experiments; however, the results are generally applicable and may be translated to Bessel filters of arbitrary cutoff frequency. Because of the finite rise time of the filter ($\tau_{rise} \approx 0.33 f_c^{-1} = 33\ \mu$s),[25,26] the pulse edges are not vertical and the pulses are attenuated when $w_{nom} \leq 2\tau_{rise}$ ($w_{nom}$ being the nominal pulse width). The future use of nanopores for studies of biomolecular processes like protein−protein interactions will depend on the ability to measure short translocation events. As the measurement techniques have reached physical limits with respect to achievable time resolution and simultaneous noise suppression,[22] new methods are needed to recover the accurate width and height of short current pulses.[14]

Here we propose two methods to analyze both the width and height of current pulses that enable us to collect meaningful data beyond the time constraints imposed by the filter electronics. The pulse width is analyzed by an improved pulse stop criterion; the pulse height is inferred from the slope of the pulse's falling edge. Since the shape of pulses originating from rapid translocation events cannot be known ab initio, at first artificial square pulses featuring a defined shape are used to validate the introduced methods. For conditions which are representative for typical nanopore translocation experiments, we exemplarily examine pulses of known height $h = 100$ pA and defined widths (varied within the range $5\ \mu s \leq w_{nom} \leq 140\ \mu$s) which are low-pass filtered by a $f_c = 10$ kHz Bessel filter. A total of 500 current pulses are analyzed with the proposed as well as conventional methods, and the evaluated width and height values are compared. Validity ranges are obtained by contrasting the analyzed with the nominal values for both methods. The analysis routines have been implemented in Matlab code, which is gladly provided upon request. Eventually, we apply the introduced methods to the analysis of a protein translocation experiment where rapid as well as long-lasting blockade events were observed. The quality of data analysis is improved significantly by the new methods.

## PULSE WIDTH ANALYSIS

Generally, a pulse is identified when a data point deviates from the current baseline by a value that is a given multiple of the standard deviation ($\sigma$). We chose a wide bound of $I^{base} \pm 6\sigma$ ($I^{base}$ being the baseline level) here in order to keep the probability of falsely identified pulses negligible.[25] The baseline $I^{base}$ is often defined globally by averaging over the complete data set, but we prefer using a locally defined baseline ("moving baseline") by averaging over a limited number of data points (typically 100) preceding a pulse. This latter method is superior in that long-time drifts and moderate low-frequency noise (e.g., $1/f$ noise) hardly affect the data analysis.

While the individual criteria to identify the pulse-start and pulse-stop points are defined slightly differently by various authors, a common practice is to use the last data point before the current drops below the baseline as the start point and the first data point when the current returns to the baseline value as the stop point.[13,27−30] These points are indicated in Figure 2A, as open circles and triangles, respectively. We analyzed current pulses of varying durations using this conventional method and plotted the evaluated pulse widths in Figure 2B (blue line). Evidently, the method overestimates the nominal pulse width by $\Delta w = w - w_{nom} \approx 60\ \mu$s, as can be seen from the constant offset of the blue line from the gray "zero error" (slope = 1) line.

We propose a modified way to measure the pulse width: using the same start point, the stop point is chosen to be the last (or only) local minimum of the pulse before the signal starts to return


(12) Fologea, D.; Ledden, B.; McNabb, D. S.; Li, J. L. *Appl. Phys. Lett.* **2007**, 91.
(13) Han, A.; Creus, M.; Schurmann, G.; Linder, V.; Ward, T. R.; de Rooij, N. F.; Staufer, U. *Anal. Chem.* **2008**, *80*, 4651–4658.
(14) Talaga, D. S.; Li, J. L. *J. Am. Chem. Soc.* **2009**, *131*, 9287–9297.
(15) Bates, M.; Burns, M.; Meller, A. *Biophys. J.* **2003**, *84*, 2366–2372.
(16) Meller, A.; Branton, D. *Electrophoresis* **2002**, *23*, 2583–2591.
(17) Akeson, M.; Branton, D.; Kasianowicz, J. J.; Brandin, E.; Deamer, D. W. *Biophys. J.* **1999**, *77*, 3227–3233.
(18) Movileanu, L. *Trends Biotechnol.* **2009**, *27*, 333–341.
(19) Oukhaled, G.; Mathe, J.; Biance, A. L.; Bacri, L.; Betton, J. M.; Lairez, D.; Pelta, J.; Auvray, L. *Phys. Rev. Lett.* **2007**, 98.
(20) Meller, A.; Nivon, L.; Brandin, E.; Golovchenko, J.; Branton, D. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 1079–1084.
(21) Storm, A. J.; Storm, C.; Chen, J. H.; Zandbergen, H.; Joanny, J. F.; Dekker, C. *Nano Lett.* **2005**, *5*, 1193–1197.
(22) Tabard-Cossa, V.; Trivedi, D.; Wiggin, M.; Jetha, N. N.; Marziali, A. *Nanotechnology* **2007**, 18.
(23) Uram, J. D.; Ke, K.; Mayer, M. *ACS Nano* **2008**, *2*, 857–872.
(24) Wonderlin, W. F.; French, R. J.; Arispe, N. J. In *Neurophysiological Techniques*; Boulton, A. A., Baker, G. B., Vanderwolf, C. H., Eds.; Humana Press: Clifton, NJ, 1990; pp 35−142.
(25) Colquhoun, D.; Sigworth, F. J. In *Single Channel Recording*, 2nd ed.; Sakmann, B. Neher, E., Eds.; Plenum Press: New York, 1995; pp 483−587.
(26) Colquhoun, D. In *Microelectrode Techniques (The Plymouth Workshop Handbook)*; Standen, N. B., Gray, P. T. A., Whitaker, M. J., Eds.; Company of Biologists: Cambridge, U.K., 1987; pp 83−104.
(27) Iqbal, S. M.; Akin, D.; Bashir, R. *Nat. Nanotechnol.* **2007**, *2*, 243–248.
(28) Sexton, L. T.; Horne, L. P.; Sherrill, S. A.; Bishop, G. W.; Baker, L. A.; Martin, C. R. *J. Am. Chem. Soc.* **2007**, *129*, 13144–13152.
(29) Smeets, R. M. M.; Keyser, U. F.; Krapf, D.; Wu, M. Y.; Dekker, N. H.; Dekker, C. *Nano Lett.* **2006**, *6*, 89–95.
(30) Heng, J. B.; Ho, C.; Kim, T.; Timp, R.; Aksimentiev, A.; Grinkova, Y. V.; Sligar, S.; Schulten, K.; Timp, G. *Biophys. J.* **2004**, *87*, 2905–2911.

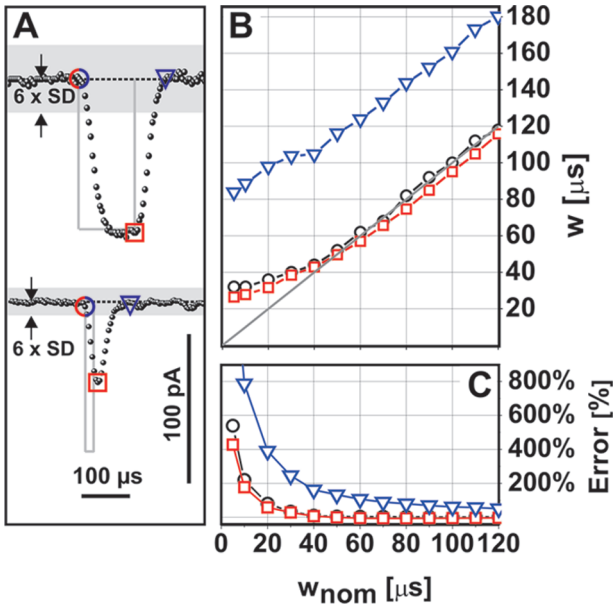


**Figure 2.** Pulse width analysis. Panel A shows ideal square wave current pulses with $h_{nom}$ = 100 pA and $w_{nom}$ = 120 (top) and 20 $\mu$s (bottom) in gray and representative Bessel-filtered current pulses as black solid lines. The start and stop points for the pulse width analysis are indicated as open symbols. Start points are circles; stop points are blue triangles (conventional method) and red squares (new method). (B) Pulse width $w$ analyzed with the conventional method (blue triangles), the fwhm method (black circles), and the new method (red squares) plotted vs the nominal pulse width $w_{nom}$. The solid gray line denotes "perfect" analysis, i.e., $w = w_{nom}$. (C) Relative error ($|w - w_{nom}|/w_{nom}$) of the conventional and new methods.

to the baseline; cf. the open squares in Figure 2A. This criterion is based on the fact that the onset of the rising edge actually marks the end of the real pulse: When the real pulse ends, it creates an impetus for the system to return to the baseline level, to which it will react on a time-scale related to its intrinsic response time, i.e., the filter rise time. With the choice of the onset of the rising edge as the stop point, a systematic overestimation of the pulse width due to the finite filter rise time can be avoided. When analyzing the same data as before using this method, we find that down to $w_{nom}$ = 40 $\mu$s, the evaluated pulse width corresponds almost perfectly to the nominal pulse width (red line in Figure 2B). The observed deviation, $\Delta w < -5$ $\mu$s, is comparable to the used data sampling time which was chosen to be 4 $\mu$s (equivalent to typical sampling rates of commercial patch-clamp amplifiers). Only for $w_{nom} < 20$ $\mu$s, $w$ starts to diverge considerably from $w_{nom}$.

Panel C of Figure 2 shows a comparison of the relative errors $\Delta w/w_{nom}$ brought about by the conventional and new methods, respectively. The conventional method overestimates the real pulse duration significantly; even for comparably long pulse durations equivalent to the inverse of the filter cutoff frequency ($w_{nom} = f_c^{-1} = 100$ $\mu$s), it results in a considerable error of roughly 50%. The error of the new method in that regime is negligible. For short pulses, e.g., $w_{nom}$ = 30 $\mu$s, the conventional method cannot be used anymore since it produces errors of 350%, whereas the new method causes tolerable errors of merely 20%. For very short pulses, $w_{nom} < 20$ $\mu$s, the new method starts to fail as well and results in errors exceeding 50%. The results show that the improved pulse stop point criterion allows us to recover meaningful pulse width data beyond the limits discussed in the literature, which are usually given as $2\tau_{rise} \approx 0.66 f_c^{-1}$.[24−26,31,32] In principle, our method is not affected by the filter rise time as long as the pulse reaches a plateau and the stop point does not lie on the falling edge of the pulse. Thus, the method can be used to analyze pulses with $w_{nom} > 0.2 f_c^{-1}$.

Additionally, we analyzed the data set by the full-width-half-maximum (fwhm) method, which is also used by some authors to determine the pulse width. Although the analysis does not rely on a physically motivated definition of the pulse start and stop points, it is interesting to note that the quality of the fwhm analysis is superior to the conventional and comparable to the proposed method. This results from the fact that during the fwhm analysis half of the pulse fall time and half of the rise time are not taken into account, which favorably partly compensates the pulse dilation due to the finite filter rise time.

## PULSE HEIGHT ANALYSIS

Two methods are mostly used in the literature to evaluate the pulse height: some authors average over the current values within the pulse plateau to obtain a mean value,[20,33] while others simply take the extremal value between the start and stop points as the pulse plateaus.[13,27,28] The pulse height is then obtained by computing $h = I^{base} - I^{plateau}$. We analyzed current pulses of varying widths ($w_{nom}$) and constant height ($h_{nom}$ = 100 pA) using the former method and plotted the height values $h$ obtained from individual pulses as histograms in Figure 3A.

Panel B depicts the center height values ($\bar{h}$) as a function of $w_{nom}$ which were obtained by fitting Gaussian bell curves to the histograms. As $w_{nom}$ decreases below ≈50 $\mu$s, we observe that $\bar{h}$ begins to deviate from $h_{nom}$. For short pulses, the evaluated $\bar{h}$ values clearly underestimate the real pulse height (for instance, $[|\bar{h} - h_{nom}|/h_{nom}] \approx 50\%$ for $w_{nom}$ = 20 $\mu$s). The error is again caused by the finite filter rise time. Pulses with $w_{nom} < 2\tau_{rise}$ do not reach their full height, i.e., a plateau does not develop (see Figure 1B). To avoid an underestimation of the pulse height, those pulses have been rejected in earlier works.[24−26,31,32]

We devised a method to analyze the pulse height from the slope of the falling edge to circumvent this problem. The principle of the slope method is shown in panel C of Figure 3, which depicts the falling edges of three pulses with different pulse heights ($h_{nom}$ = 40, 60, and 80 pA). Evidently, the slope becomes steeper with increasing pulse height. We analyzed the slopes of pulses with varying heights and widths and plotted the dependence of the slope on $h_{nom}$ and $w_{nom}$ in Figure 4. A linear relationship between the height and the slope $s$ of the falling edge is found:

$$h_{nom} = \alpha s \qquad (1)$$

The factor α remains constant for $w_{nom} > 30$ $\mu$s ($\alpha \approx 0.027$ $\mu s^{-1}$) as all lines coincide. For shorter pulses, α starts to decrease which is reflected by a deviation of the lines for 20,

(31) McManus, O. B.; Blatz, A. L.; Magleby, K. L. *Pfluegers Arch.* **1987**, *410*, 530–553.
(32) Skinner, G. M.; van den Hout, M.; Broekmans, O.; Dekker, C.; Dekker, N. H. *Nano Lett.* **2009**, *9*, 2953–2960.
(33) Meller, A.; Nivon, L.; Branton, D. *Phys. Rev. Lett.* **2001**, *86*, 3435–3438.

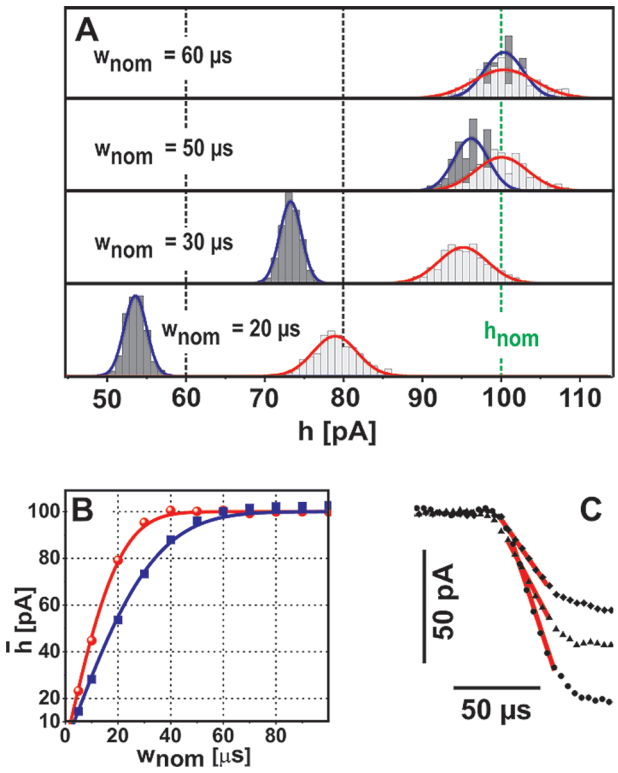


**Figure 3.** Pulse height analysis: (A) Pulse height histograms of 500 artificial pulses ($h_{nom}$ = 100 pA, $w_{nom}$ = 60, 50, 30, and 20 $\mu$s) which were analyzed with the conventional (dark gray columns, blue lines) and slope method (light gray columns, red lines), respectively. Solid lines are Gaussian fits to the data. (B) Center height values ($\bar{h}$) of the bell curves in panel A as a function of $w_{nom}$. Blue squares, conventional analysis; red spheres, analysis with the slope method. Solid lines are fits to the data using eq 2. (C) Magnified view of the falling edges of three pulses with different pulse heights ($h_{nom}$ = 40, 60, and 80 pA; $w_{nom}$ = 500 $\mu$s). Straight red lines are linear fits used for the slope analysis.

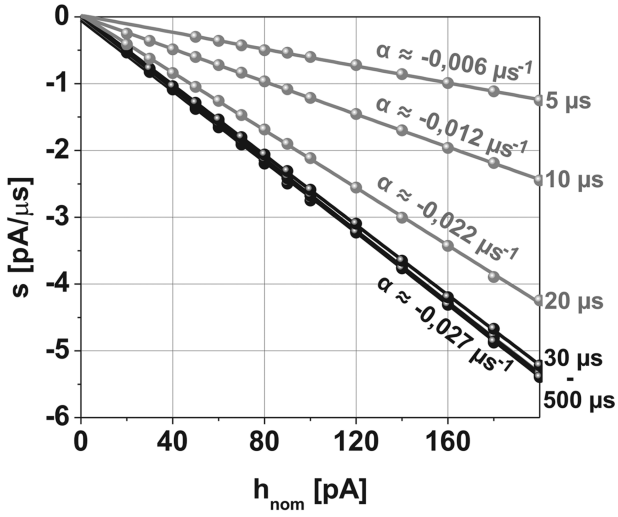


**Figure 4.** Validation of the slope method. Slope $s$ evaluated from falling edges of filtered square pulses are plotted against the nominal pulse height $h_{nom}$ for varying $w_{nom}$ (given on the right). For 30 $\mu$s < $w_{nom}$ < 500 $\mu$s, data coincide. Lines are linear fits to the data (eq 1).

10, and 5 $\mu$s from the $\alpha \approx 0.027\ \mu s^{-1}$ cohort. Hence, we may infer a calibrated range, i.e., where $\alpha$ = const, for which the true pulse height can be calculated from the measured slope $s$ using eq 1. This is a remarkable result, because it means that the pulse height can be deduced from the edge slope even if the pulse is significantly shorter than $f_c^{-1}$ and the pulse has not reached its full height (developed a plateau). Moreover, the pulse height is determined independently from the pulse width.

The benefit of the slope method becomes obvious in Figure 3A,B. As $w_{nom}$ decreases, the deviation of $\bar{h}$ from $h_{nom}$ proceeds substantially slower for the slope method than for the conventional method. Thus, the slope method allows us to recover $\bar{h}$ correctly for pulses as short as $w_{nom} \approx 30\ \mu s\ (= 0.3 f_c^{-1})$. However, we also note that the standard deviation in the histogram is slightly higher for the slope method; hence, the uncertainty in the determination of the height of a single, long ($w_{nom} > f_c^{-1}$) pulse is higher.

Next we compare the performance of the conventional and slope method by evaluating their apparent cutoff frequencies. It has been shown that Bessel filters (for a large number of poles) can be theoretically modeled by a Gaussian filter.[25] The pulse height may then be expressed as

$$h = h_{nom}\,\mathrm{erf}(2.668\, f_c^{eff}\, w_{nom}) \qquad (2)$$

where $f_c^{eff}$ represents an effective filter cutoff frequency. The solid lines in Figure 3B were computed with eq 2 using $f_c^{eff}$ as the only free fit parameter. As expected, we find $f_c^{eff} \cong 10$ kHz $= f_c$ for the conventional method. In contrast, the fit-analysis for the slope method yields $f_c^{eff} \cong 16.4$ kHz $= 1.64 f_c$. This denotes an increase in the effective bandwidth of the system due to the improved pulse analysis method.

After finishing this article, a recent paper by Talaga and Li came to our attention, who had also noticed the necessity to improve data analysis and reported an elegant procedure to correct attenuated protein translocation events.[14] Employing different methods to measure the pulse width and height, Talaga and Li obtained a calibration curve similar to the blue line depicted in Figure 3B.

Finally, we apply the slope method for the analysis of real protein translocation data and demonstrate its advantage over the conventional method. Figure 5 shows representative current traces of avidin molecules in the SiN pore depicted in Figure 1A (avidin concentration, ~1 $\mu$M; bias potential, 150 mV). We characterized avidin in solution by dynamic light scattering and obtained a hydrodynamic diameter of 8.5 nm. The diameter and length of the used nanopore were 20 and 30 nm, respectively.

We observe two types of events: Rapid blockades with durations of several microseconds and comparably small pulse heights appear very frequently (type 1 events). Blockades of this kind have been reported also by other authors studying protein translocations.[12−14] Another type of blockades (type 2) appears less frequently but lasts considerably longer than the rapid events. The duration of type 2 events is very diverse. Some events last a couple of ten microseconds (cf. magnified view between 4.2 and 4.4 s in Figure 5), some last a couple of milliseconds (cf. magnified view between 5.4 and 5.6 s as well as the dips at 5.2 and 5.5 s in Figure 5), and sometimes we observe extremely long-lasting events which manifest themselves as apparent shifts of the current baseline (cf.

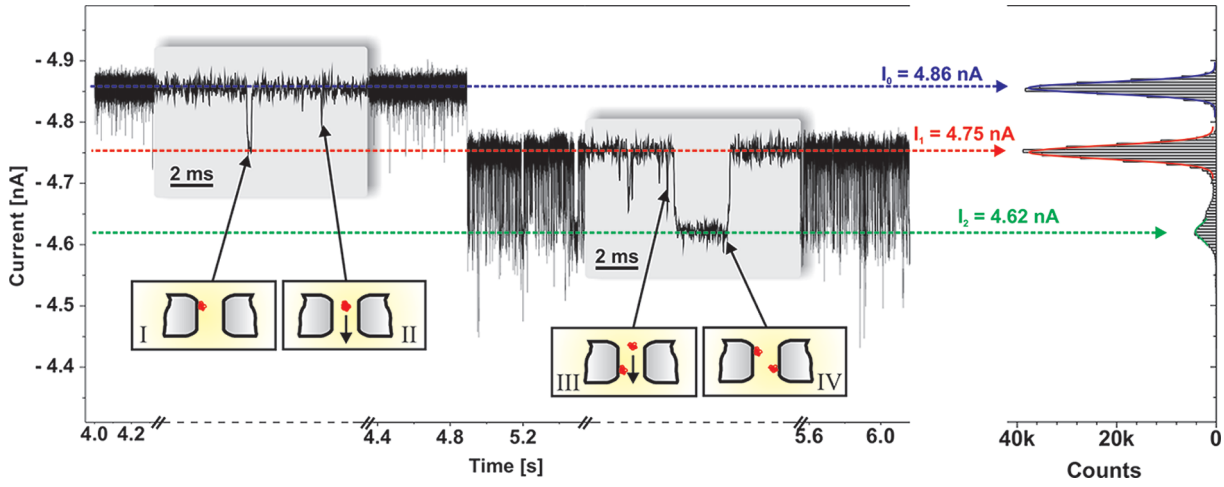


**Figure 5.** Transpore ion current measurements in the presence of avidin. Avidin was added to the cis side of the nanopore while a voltage of −150 mV was applied to the trans side. Left, current time trace including magnified views and, right, histogram of current values. The schematic insets illustrate avidin molecules inside the pore which transiently adsorb to the wall (I) or cross the pore in ballistic flight (II).



4.9 s in Figure 5). In contrast to type 1 events, the height of type 2 events is well-defined and, reminiscent of DNA translocations, the pulse shape is rectangular. Negative control measurements are presented in the Supporting Information as Figure S1.

We attribute the observed current blockades to the presence of avidin molecules within the pore. When inside the pore, the proteins narrow the effective pore diameter and thus suppress the transpore current. The nature of type 1 and type 2 events is different; we suppose that type 1 events arise from "ballistic" translocations where avidin molecules cross the pore without significant interactions, whereas type 2 events come about by the transient, or quasi-permanent, adsorption of avidin on the inner pore walls. This notion is supported by the known tendency of avidin to adsorb nonspecifically onto silicon surfaces.[34] In addition, the diffusive time-of-flight of proteins in nanopores is expected to be much shorter than the long durations observed for type 2 events.[12] Using this interpretation, we can assign the observed pulses in Figure 5 to one of four different states which arise from ballistic and/or adsorption blocking and are schematically depicted as insets in Figure 5.

Following the geometric argument that avidin acts as a plug that partially blocks the pore volume, type 1 and type 2 events should feature the same current blockade magnitude. This, however, is not observed in the measurement where short pulses, in accordance with the previous discussion, appear to be attenuated. When analyzing the pulse height of ~30 000 ballistic translocations through the "open" pore (scheme II in Figure 5) using the conventional method, we obtain a histogram depicted in Figure 6. The histogram peaks at a blockade current level of 50 pA which deviates significantly from the expected value of $h_{ref} = I_0 - I_1 = 105$ pA. Alternatively, when we use the slope method to analyze type 1 events, we obtain a distribution of current blockades which features a maximum at 91 pA, i.e., close to the expected $h_{ref}$ value. Thus, we conclude that the slope method is more suitable to recover the real height of attenuated current pulses. It enables the researcher to identify correlations and similarities among pulses in a data set which, in case of the protein measurement at hand, could have been falsely assigned to originate from distinct phenomena otherwise.

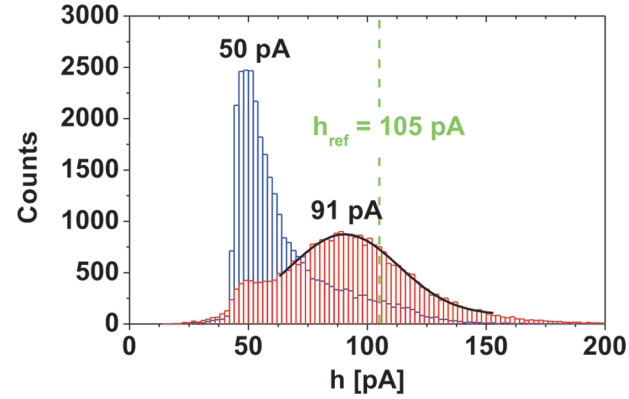


**Figure 6.** Pulse height histograms of rapid blockade (type 1) events of an avidin translocation experiment. The same data set was analyzed with the conventional (blue columns) and slope method (red columns). The solid line represents a Gaussian fit to the data. The vertical dashed line indicates the expected pulse height $h_{ref}$ as inferred from adsorption (type 2) events

## METHODS

**Pulse Generation and Acquisition.** A function generator (Agilent 33250A) is used to generate a square wave signal. The pulses are fed into the stimulus input (10 GΩ) of a HEKA EPC 8 patch-clamp amplifier (test mode) with a seven-pole Bessel filter ($f_c$ = 10 kHz) and a sampling rate of 250 kHz. The filtered pulses are recorded by a LeCroy oscilloscope (waveRunner 44MXi) from the EPC8 current monitor output.

**Data Processing.** Pulses are analyzed by a homemade routine implemented in Matlab 7.1. Protocol of the Matlab code: (1) For every data point, a baseline is computed by averaging a user defined number (typically 100) of preceding data points (moving baseline). (2) The pulse analyzing algorithm is triggered by a sample point that deviates from its baseline by a given multiple of the standard deviation ($\sigma$), which is the trigger level. (3) Tracking back from this trigger point, the start point of the pulse (cf. open circle in Figure 2A) is determined by the data point prior to the first local maximum above the trigger level (or ultimately

(34) Clare, T. L.; Clare, B. H.; Nichols, B. M.; Abbott, N. L.; Hamers, R. J. *Langmuir* **2005**, *21*, 6344–6355.

when the baseline is reached). (4) The positions of local minima below the trigger level are evaluated. The last local minimum is set as the pulse stop point (cf. open square in Figure 2A). (5) The pulse width is computed from the start and stop points. (6) Option (not used for the method validation presented in Figures 3 and 4): The code features the possibility to discriminate those pulses assumed to be attenuated from those which have developed to their full height: If there is more than one local minimum present, the data points in between the first and last minimum are averaged. The pulse height is then defined by the difference of this average value and the baseline value of the start point. (7) The pulse height is determined from the pulse's falling edge (cf. Figure 3C): (a) The data point closest to the mean value of the falling edge is determined (slope point). (b) The slope is calculated by a linear regression of the slope point ± 3 adjacent data points. (c) If a pulse exhibits a local maximum or minimum within this region it is rejected, which is of minor statistical relevance (only approximately 2% of real single molecule translocation data). (d) With the use of the linear relation $h = \alpha s$ between the height ($h$) and the slope ($s$), the pulse height is derived ($\alpha$) and has to be determined by a calibration before (cf. Figure 4). (8) The pulse width and height are output as a text file.

**Nanopore Fabrication.** We employed two methods to fabricate single nanopores. For the DNA translocation experiment, single nanopores are processed in silicon nitride membranes by e-beam lithography and wet chemical etching, followed by TEM induced shrinking.[35,36] Starting with a (100) silicon chip of 200 $\mu$m thickness, which features 50 nm silicon nitride coatings on both sides, we use optical lithography to form an etch mask on the chip's back side for anisotropic etching of the silicon with KOH. Subsequently we utilize e-beam lithography on the front side to open holes of 40−50 nm in the silicon nitride. In the next step, the silicon is etched by KOH resulting in a freely suspended silicon nitride membrane incorporating a single nanopore. To fabricate a desired pore size of 12 nm, we shrink the pores using a Jeol 100CX TEM. For the avidin translocation experiment, the nanopore was directly drilled into a silicon nitride membrane[35,36] of 30 nm thickness with a 300 keV FEI Titan TEM.

(35) Storm, A. J.; Chen, J. H.; Ling, X. S.; Zandbergen, H. W.; Dekker, C. *Nat. Mater.* **2003**, *2*, 537–540.

(36) Wu, M. Y.; Krapf, D.; Zandbergen, M.; Zandbergen, H.; Batson, P. E. *Appl. Phys. Lett.* **2005**, 87.

**Translocation Experiments.** The nanopore chip is passivated by PDMS[22] and placed in between two compartments containing the electrolyte solution and sealed by O-rings. In each compartment, an Ag/AgCl electrode is immersed and 150 mV are applied across the nanopore with a HEKA EPC 8 patch-clamp amplifier. The current is filtered ($f_c$ = 10 kHz) and monitored with the EPC 8 and recorded with a HEKA LIH 1600 data acquisition board set to 200 kHz sampling rate. For the DNA translocation experiment, 8 nM $\lambda$-DNA (Fermentas Life Sciences) solution is used (1 M KCl, 10 mM Tris, 1 mM EDTA, pH 7.6. For the protein translocation, ∼1 $\mu$M avidin (ThermoFisher) was dissolved in 400 mM KCl (pH 10).

## CONCLUSIONS

The proposed methods significantly improve the data analysis accuracy of filtered current pulses. With the use of the suggested pulse-stop criterion, meaningful information on the pulse width may be inferred from pulses as short as $w_{nom} \geq 0.3 f_c^{-1}$. The proposed slope method enables us to recover the real pulse height from pulses with $w_{nom} \geq 0.2 f_c^{-1}$, even if the pulses are too short to reach their full height. Hence, the introduced methods allow for the analysis of short current pulses and will be particularly useful in single molecule experiments where small molecules rapidly pass through nanopores.


## ACKNOWLEDGMENT

We are very thankful to Prof. Gerhard Abstreiter for his support. We also thank Markus Döblinger and Prof. Thomas Bein for their support (TEM) within the Nanosystems Initiative Munich (NIM) and Stephan Renner for discussions. Financial support is gratefully acknowledged from the Bundesministerium für Bildung und Forschung (BMBF Grant 0312031), the TUM Institute for Advanced Study (IAS), and the TUM International Graduate School of Science and Engineering (IGSSE).


## SUPPORTING INFORMATION AVAILABLE

Control measurement in the absence of avidin. This material is available free of charge via the Internet at http://pubs.acs.org.


Received for review August 19, 2009. Accepted October 19, 2009.

AC901877Z