Supporting Material

Exponential sum-fitting of dwell-time distributions from ion channels without specifying starting parameters

David Landowne*, Bin Yuan, and Karl L. Magleby*

*Department of Physiology and Biophysics, University of Miami Miller School of Medicine, Miami, FL 33136 USA.

Correspondence: dlandowne@med.miami.edu or kmagleby@med.miami.edu
Bin Yuan's present address is Pisces Electronic Solutions, binyuana@yahoo.com

Contents

Estimating the number of initial exponentials to use for fitting

We have found that 20 initial exponentials with equal log spacing of the time constants from the briefest to the longest interval in the distribution have been sufficient to find all of the significant exponentials in the examples considered in this study. Although we have not looked at every published dwell-time distribution, for those examined, 20 initial exponentials should be more than sufficient to detect the exponential components reported in those dwell-times. Nevertheless, it would be useful to have a method of estimating the minimal number of initial exponentials required to find the significant exponentials in a distribution. This section presents such a method.

The fitting algorithm needs to start with a sufficient number of initial exponentials with time constants equally spaced on a log axis so that the time constant of every significant exponential in the fitted distribution is approximated by a time constant of one of the initial exponentials. With this approach, the significant exponentials have already been found so that the fitting procedure does not have to search for exponentials, but only eliminate those exponentials that make insignificant contributions to the likelihood while making adjustments to the areas and time constants of the exponentials to maximize the likelihood. If the ratio between the time constants of the adjacent initial exponentials is equal to or less than the minimal ratio of any of the significant adjacent exponentials summing to form the dwell-time distribution, then each significant exponential in the dwell-time distribution would be approximated by at least one of the initial exponentials, so that the significant exponentials should be found. On this basis, if the smallest ratio of the time constants of adjacent initial exponentials (longer/shorter) is R, the time to be spanned by the initial exponentials is T_2-T_1, where T_2 and T_1 are the longest and shortest intervals in the distribution, respectively, and N is the number of initial exponentials required to span the time range, then

$$R^{N\text{-}1} = T_2/T_1 \qquad \text{(S1)}$$

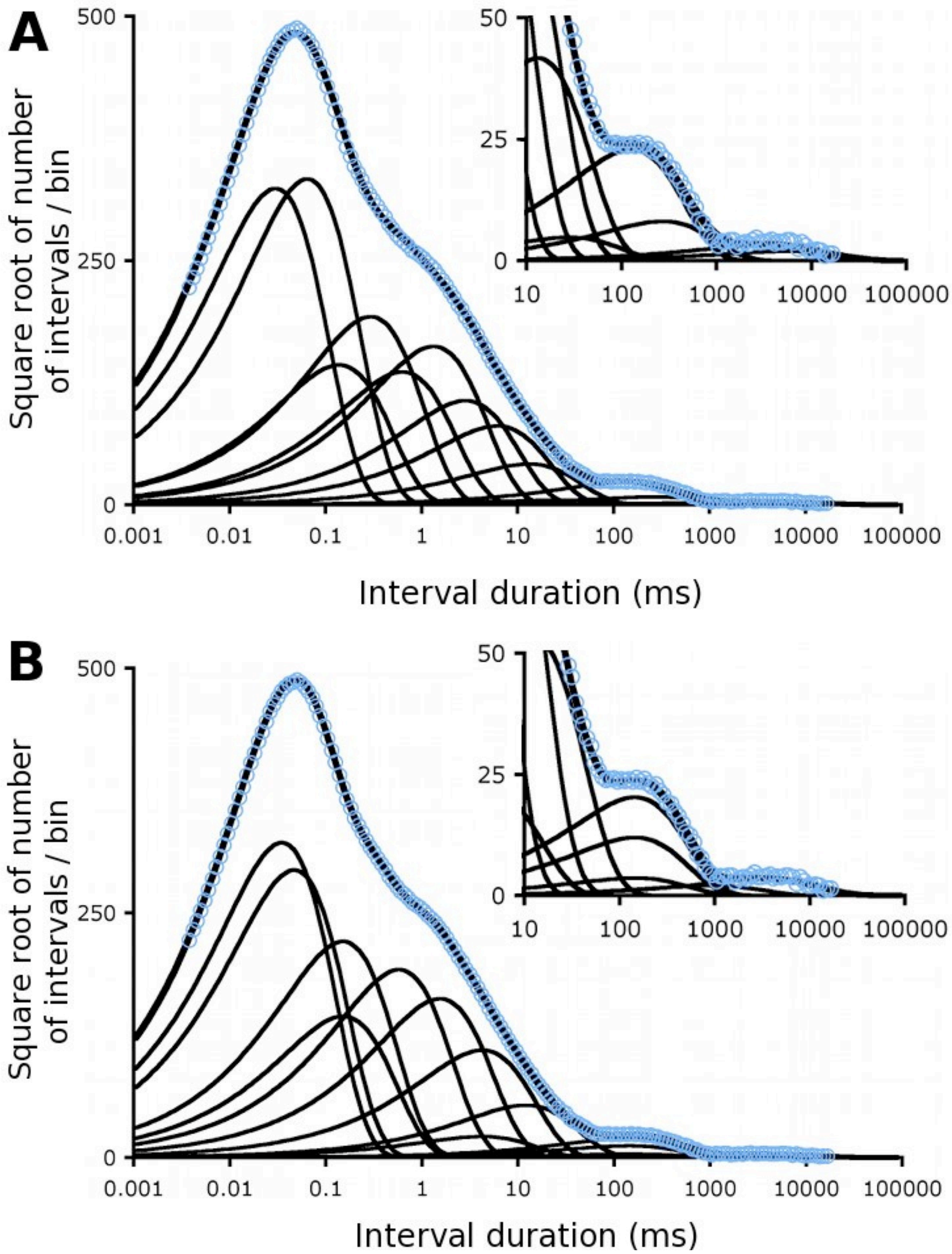
which can be rearranged to

$$N = 1 + \log\,(T_2/T_1)\,/\,\log R \qquad \text{(S2)}$$

As an example, consider the data in Fig. 1 and Table 1, were the minimal ratio between adjacent time constants was ~2.7 (1.48 ms/0.547 ms) and the longest, T_2, and shortest, T_1, intervals included in the fitting of the dwell-time distribution were ~18,000 ms and ~0.004 ms. Substitution into Eq. S2 gives $N = 16.4$ for the minimum number of initial exponentials. We found that all eight significant exponentials were found using from 10 to 40 initial exponentials for the data in Fig. 1 and Table 1. Hence, Eq. S2 should provide sufficient initial exponentials to find the significant exponentials. For experimental data, the minimal ratio between adjacent exponentials is not known, so R could be selected to place an approximate lower limit on the sensitivity of detecting significant exponentials assuming sufficient intervals were available for detection. The only drawback for making R smaller than needed is an increase in the time required for fitting. As with all analysis programs, simulation and fitting of data similar to that being analyzed should be used to test the limits and applicability of the fitting programs for the types of data being analyzed.

Setting the minimal area for detection of an exponential in steps 4 and 5

In steps 4 and 5 the minimal area of an exponential to be detected was set to 10^{-5}, as this value allowed detection of all significant exponentials in the examples used in this paper and also appears suitable to detect the significant exponentials fitted to those dwell-time distributions in the literature that we have examined. Hence, a single setting of 10^{-5} could be used for most (perhaps all) single channel data. Nevertheless, to look for significant exponentials with areas < 10^{-5} in distributions with >100,000 intervals, this value could be reduced. Maximum likelihood fitting allows the detection of a significant slower exponential arising from only a few intervals (see Fig. 8 in McManus and Magleby (19)), or even from a single interval (this can be shown by simulation and will not be presented here), provided that the duration of the interval or intervals from the slower exponential are sufficiently long compared to the time constant of the next slowest exponential. The minimal area that could be detected for an exponential based on a single interval would be approximated by 1/(the number of intervals in the distribution). Hence, the minimal area could be set to $< 2 \times 10^{-6}$ to have the ability to detect an exponential based on a single interval from a distribution with 500,000 intervals. The fitted time constant for the detection of a significant exponential with a single interval would be approximated by the duration of the interval, but the time constant and area would be very poorly defined due to stochastic variation, so additional data giving additional intervals from the slow exponential would be needed to place reasonable limits on the time constant and area.



SUPPORTING FIGURE S1 Intermediate steps in fitting for Fig. 1 based on Table 1. Fig. 1 *B* in the Discussion and Results section plots the distribution of initial exponentials after step 2 of the fitting algorithm, with 20 initial exponentials equally log spaced with an area of 0.05 each. This figure then shows the results after step 5 and step 6.

(*A*) A plot of the fit after step 5, which removed six exponentials with areas $<10^{-5}$, leaving 14 of the 20 initial exponentials. The fitted exponentials (black lines) were summed to obtain the predicted distribution (continuous black line) which gives a good approximation to the data (blue circles), but with small differences between the sum of the exponentials and the data. Note that this close approximation to the data was obtained even though the time constants of the exponentials were fixed to those of the initial exponentials.

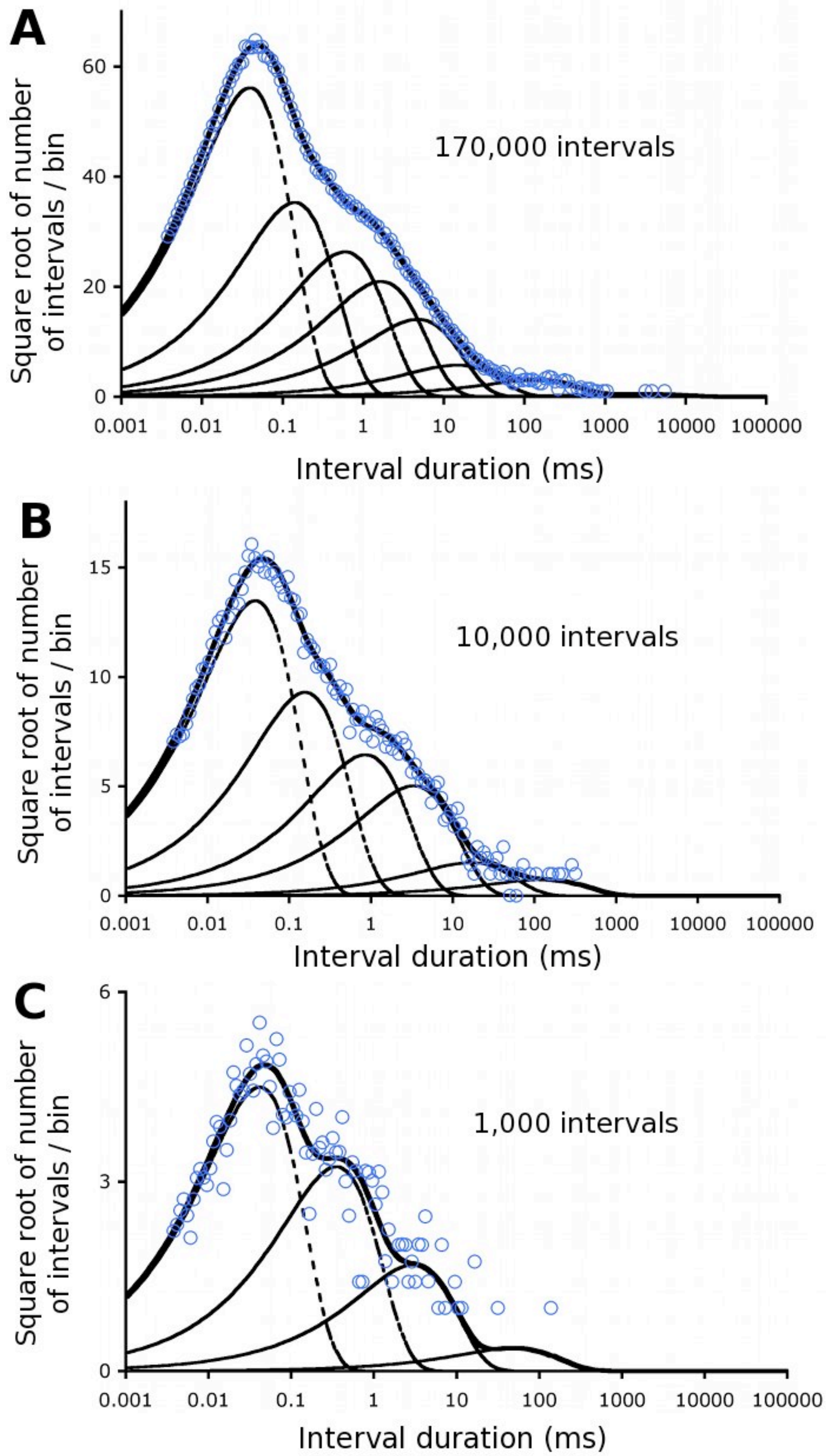
(*B*) The fit after step 6. Starting with the 14 exponentials after step 5 in part *A*, the data were refit with both the areas and the time constants as free parameters. The fitted exponentials (black lines) were summed to obtain the predicted distribution which gave an excellent description of the data (black line through the blue circles). This is the best (most likely) fit that is obtained to the data during the fitting process, but with extra exponentials that have little if any effect on the likelihood.

Steps 7-10 then removed six exponentials that had insignificant effects on the likelihood after refitting, giving the results shown in Fig. 1. The two fitted exponentials at 0.1503 and 0.1504 ms had essentially the same time constants and were combined into one exponential in step 7 with summed area and weighted time constant with refitting after the combination. In a similar manner, the two exponentials at 4.2363 and 4.2608 ms were combined, and the three exponentials at 145.111, 145.120, and 145.122 ms were combined, with refitting after each combination. This combination of exponentials with time constants that differed by <2% then eliminated four exponentials, leaving 10. Exponentials were then removed one by one and those that had insignificant effects on the likelihood after their removal and refitting were deleted. In this manner exponentials at 0.03487 and 0.04693 ms ended up as a single exponential with a time constant of 0.03885 ms and exponentials at 927.14 and 3562.8 ms ended up as a single exponential with a time constant of 3,311 ms, leaving the eight significant exponentials shown in Fig. 1.

SUPPORTING TABLE S1 Consistent detection of eight significant exponentials

Parameter	(ms)	Time constants from fitting data for six separate simulations						Mean	SD
τ1	0.039	0.038922	0.038850	0.038827	0.038825	0.038822	0.038869	0.038853	0.000039
τ2	0.139	0.139433	0.138776	0.136384	0.137493	0.137394	0.138869	0.138058	0.001151
τ3	0.547	0.537204	0.548794	0.505645	0.522916	0.539900	0.547145	0.533601	0.016508
τ4	1.48	1.397209	1.450843	1.342277	1.380270	1.419130	1.496087	1.414303	0.054194
τ5	4.12	3.843757	3.957510	3.901060	3.935880	3.949595	4.173018	3.960136	0.112303
τ6	11.9	11.37811	11.72684	11.81652	11.63800	11.55447	12.00000	11.68566	0.215267
τ7	146	148.3488	147.6353	145.5389	146.2227	145.4270	147.7733	146.8243	1.252920
τ8	3390	3177.264	3218.838	3347.565	3609.777	3221.034	3293.315	3311.299	158.3603
	(Fractional area)	Areas from fitting data for six separate simulations							
a1	0.539	0.539110	0.537705	0.536206	0.537197	0.536693	0.538219	0.537522	0.001055
a2	0.220	0.220169	0.221607	0.219107	0.220108	0.221668	0.220916	0.220596	0.000991
a3	0.111	0.106624	0.110599	0.104212	0.105768	0.108966	0.111256	0.107904	0.002809
a4	0.0815	0.081694	0.078931	0.087190	0.084999	0.081441	0.082134	0.082732	0.002918
a5	0.0379	0.040686	0.040213	0.042387	0.040793	0.040033	0.037154	0.040211	0.001713
a6	0.0088	0.009982	0.009196	0.009145	0.009385	0.009446	0.008588	0.009290	0.000455
a7	0.0017	0.001689	0.001703	0.001708	0.001705	0.001707	0.001688	0.001700	0.000009
a8	0.000046	0.000046	0.000045	0.000045	0.000046	0.000047	0.000044	0.000046	0.000001

The time constants (τ1-τ8) and areas (a1-a8) used for the simulation (left two columns) were from McManus and Magleby (1988). The relative error (Eq. 4) calculated from the six estimates of the 16 parameters was 0.023 ± 0.028 (mean ± SD). Estimates are based on 10^7 simulated dwell-times. Fitting was from 0.00389 ms to infinity. The areas used for the fitting and error calculation were: 0.53903, 0.220012, 0.111006, 0.081504, 0.037902, 0.0088, 0.0017, and 0.000046, summing to 1.0.



SUPPORTING FIGURE S2 Detecting the significant exponentials requires fitting a sufficient number of intervals. The results from fitting dwell-time distributions comprised of a decreasing number of simulated intervals based on Table 1 are shown in *A-C*, where the simulated data are plotted as blue circles, the fitted exponential components as black dashed lines, and the sum of the exponential components as a continuous black line through the data points. The plots can be compared with Fig. 1*A* which is based on 10^7 simulated intervals. The relative error for the areas and time constants of the eight significant exponential components for fitting 10^7 intervals was 0.023 ± 0.028 (mean ± SD) calculated from Table 1.

(*A*) With 170,000 simulated intervals, all eight exponentials were detected with an increased relative error for the parameters of 0.10 ± 0.093. The time constants and (areas) of the fitted exponentials were: 0.0393 ms (0.546), 0.141 ms (0.216), 0.591 ms (0.120), 1.66 ms (0.0760), 4.77 ms, 0.0343, 14.4 ms, (0.00574), 140 ms (0.00160), and 2623 ms (0.000043);

(B) With 10,000 simulated intervals only six of the eight significant exponentials were detected. The slowest exponential of 3,390 ms with an area of 0.000046 (Table 1) was not detected, as might be expected, because there would be, on average, be only 0.46 intervals from the slowest component, and the components at 0.547 ms and 1.48 ms were combined into a single component at 0.85 ms with combined area. The time constants and (areas) of the fitted exponentials were: 0.0384 ms (0.538), 0.157 ms (0.256), 0.851 (0.1223), 3.33 ms (0.0746), and 16.6 ms 0.00680), 125 ms (0.00163);

(*C*) For the extreme case of 1,000 simulated intervals, only four of the eight significant exponentials were found. The slowest exponential of 3,390 ms was missing as expected, the next two slower exponentials of 148 ms and 11.9 ms were combined into an exponential at 48.5 ms with a fitted area of 0.00392 (about four intervals), the exponentials at 4.12 ms and 1.48 ms were combined into an exponential at 2.94 ms, and the exponentials at 0.547 ms and 0.139 ms were combined into an exponential at 0.35 ms. The time constants and (areas) of the fitted exponentials were: 0.0413 ms (0.601), 0.351 ms (0.311), 2.94 ms (0.0840), and 48.5 ms (0.00392).

Thus, as the number of intervals in the fitted dwell-time distribution was decreased, the errors in the estimated parameters first increased, and then with further reduction in the number of simulated intervals, the number of significant exponentials decreased (19) through combinations of exponentials. The errors will depend on the number of exponentials, the areas and time constants of the exponentials, and the number of simulated intervals. The deterioration of the fit as the number of fitted intervals is reduced will occur for all methods of exponential sum fitting.