

Analysis of the Transient Current in Nanopores Using a Circle-Median Filter

Shen Bin^{1,2}, Hu Zheng-Li³, Gu Zhen¹, Ying Yi-Lun³, Wang Hui-Feng^{1,*}, Long Yi-Tao^{3,*}

¹ Key Laboratory of Advanced Control and Optimization for Chemical Processes Ministry of Education, East China University of Science and Technology, Shanghai 200237, China

² Informatization Office, East China University of Science and Technology, Shanghai 200237, China

³ State Key Laboratory of Analytical Chemistry for Life Science, School of Chemistry and Chemical Engineering, Nanjing University, Nanjing 210023, China

*E-mail: shenbin@ecust.edu.cn , whuifeng@ecust.edu.cn

Received: 29 August 2022 / Accepted: 9 October 2022 / Published: 20 October 2022

To identify and analyze the transient current jump caused by the weak interaction between biomolecules in a nanopore experiment, CM Analyzer software was designed based on the Circle-Median algorithm. The algorithm effectively recognized falling events and outliers in nanopore data. To verify the accuracy of the developed software, a simulation signal with different noise (std: 0.5~2) was generated, and results showed that the software achieves data-processing requirements with a 250-kHz sampling rate and high noise. The algorithm was also used in α -hemolysin biological nanopore-based single DNA molecule detection experiments, and results showed that the software can accurately identify the falling current and rising transit in single nanopore blockages.

Keywords: Nanopore; Data analysis; Single molecule analysis; Circle-Median filtering

1. INTRODUCTION

With the update and development of laboratory nanotechnology in recent years, nanopore single-molecule detection techniques have become rapid, high-throughput and low-cost ultrasensitive detection platforms with many application prospects [1~4]. Currently, nanopore single-molecule detection techniques have been widely used in the detection of nucleic acids [5~13], proteins [14~17], polypeptides [18], organic small molecules [19], metal ions and other substances [20~23]. The basic principle of the nanopore single-molecule detection technique is to apply an external voltage to both ends of the buffer solution in a detection cell so that molecules to be tested can pass through nanopores one by one and produce single-molecule current blockages driven by electric field force (Fig. 1 a). When molecules to be tested pass through nanopores, due to the different physical and chemical properties of various molecules to be tested and their different interactions with nanopores, nanopore weak current

signals (pA level) are output as blockages with different current amplitudes, durations and signal frequencies. On this basis, the extraction and analysis of characteristics of nanopore blockage signals can help describe the nature of single molecules and weak interactions between biomolecules.

The nanopore single-molecule detection platform primarily consists of three parts: a detection cell, a signal acquisition device and data processing software. To record nanopore data accurately, it is necessary to amplify and restore weak currents using a signal acquisition device with high precision, high sampling rate and low noise [24~31].

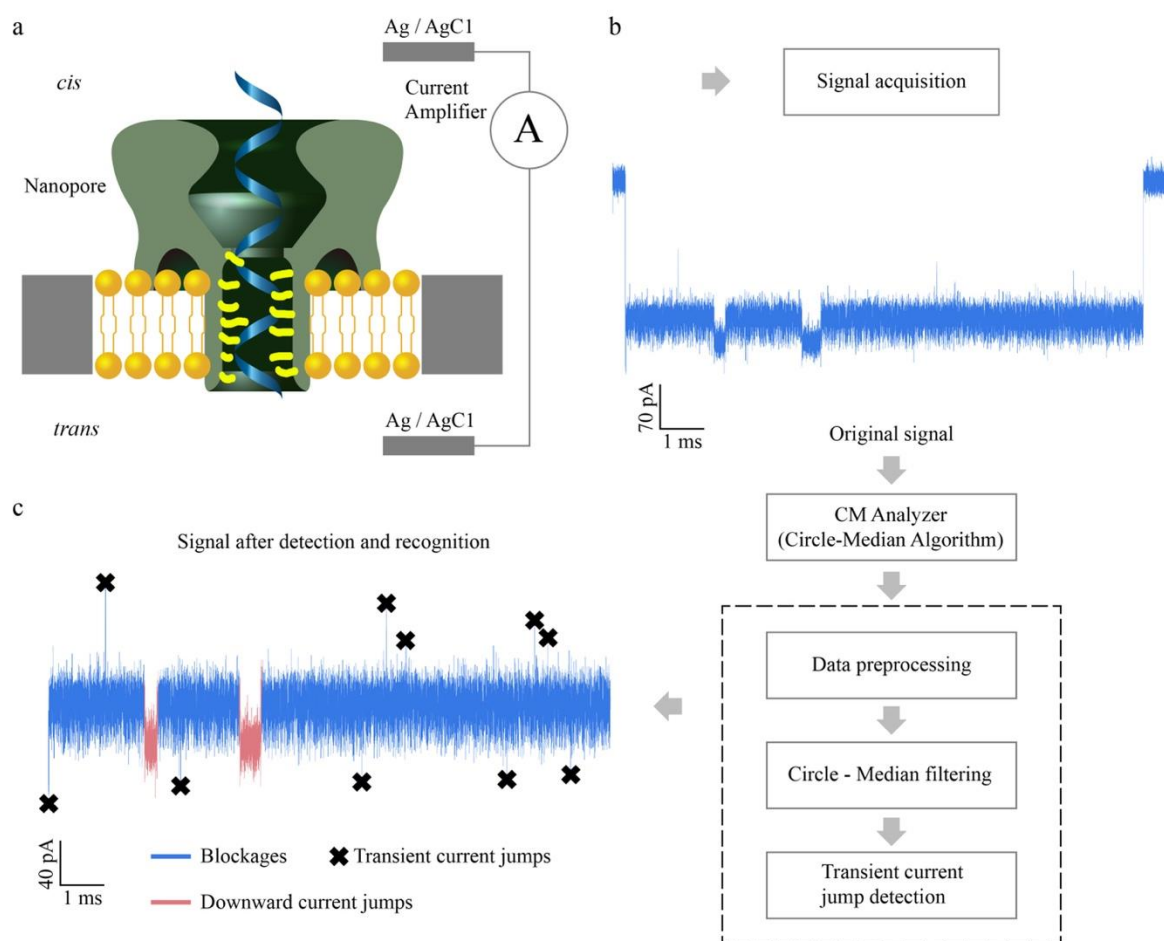


Figure 1. Schematic of nanopore detection platform: (a) illustration of single nucleic acid molecule traversing across an α -hemolysin (α -HL) nanopore; (b) original signal recorded by the signal acquisition device of nanopore detection platform, and schematic diagram of the function module of the data processing algorithm, Circle-Median; (c) signal after detection and recognition by the Circle-Median algorithm composed of blockage event (blue), current falling event (red), and outlier (black X marker).

A data processing system with high efficiency, high adaptability and low error is also required to analyze the recorded data and extract signals of research value. Currently, methods that have been used in nanopore data processing software include moving window [32], CUSUM [33], MOSAIC [34], Nanopore Analysis [35], Hidden Semi Markov Models (HSMM) & K-means hybrid algorithm [36],

Modified Hidden Markov Models [37], real-time event recognition [38], etc. These methods or pieces of software can detect or identify blockages when processing nanopore experimental data but fail to analyze transient current jumps produced by weak interactions between biological molecules.

To extract and analyze transient current jumps in nanopore data, CM Analyzer (Fig. 1b-c), which is a type of data processing software that is based on the Circle-Median algorithm, is designed in this paper and realizes the accurate identification of transient current jumps in collected nanopore data.

2. EXPERIMENT

2.1 Experimental design

2.1.1 Instrument and reagents

α -HL was purchased from Sigma–Aldrich (St. Louis, MO, USA) and were not further purified before use. Both EDTA and decane (anhydrous, $\geq 99\%$) were purchased from Sigma–Aldrich (St. Louis, MO, USA). 1,2-Diphylyl phospholipid (chloroform, $\geq 99\%$) was purchased from Avanti Polar Lipids Inc. (Alabaster, AL, USA). DNA was synthesized by Sangon Biotech (Shanghai) Co., Ltd., and purified with HPLC. Unless specifically noted, all reagents used in this experiment were analytically pure. The water for the experiment was ultrapure water ($18.2 \text{ M}\Omega\cdot\text{cm}$, 5°C) and prepared using a Milli-Q Water Purification System (Billerica, MA, USA). The sequence of DNA molecules to be tested is as follows: 5'-TTTTTTTTTTTTTTTTTTTTTCAACATCAGTCTGATAAGCTATTTTTTTTTTTTTTTTTTTT-3'

2.1.2 Design of simulation signals

In this study, random simulation signals programmed with Python were used. Blockage signals were generated and were collected after being recorded for $T = 100 \text{ ms}$ at a sampling frequency of $f = 250 \text{ kHz}$ in the nanopore single-molecule detection experiment. These signals were composed of four parts:

- 1) the blockage formed by molecules to be tested through nanopores, whose current value was $I_{\text{blockage}} = 20 \text{ pA}$;
- 2) the downward current jump formed by weak interaction between biomolecules, whose number was $3 \leq N_{\text{falling}} \leq 5$, current was $8 \leq I_{\text{falling}} \leq 15$. The data size recorded in each downward current jump was $300 \leq n_{\text{falling}} \leq 2400$;
- 3) the transient current jump formed by weak interaction between biomolecules, whose number was $4 \leq N_{\text{outlier}} \leq 8$, and the current of each transient current jump was $13 \leq |I_{\text{outlier}} - I_{\text{main}}| \leq 15$;
- 4) different noises were added in sequentially.

After the simulation signals were synthesized, the signals were analyzed and processed using the data processing software CM Analyzer.

2.1.3 Nanopore single-molecule detection experiment

The nanopore detection method was drawn from studies reported by references [5], [10]. One milliliter of buffer solution (1.0 M LiCl, 10 mM Tris, 1.0 mM EDTA, pH 8.0) was added to two chambers, cis and trans, of an acetal resin detection cell (Warner Instruments, Hamden, CT, USA). Phospholipid decane solution (30 mg/mL) was applied to a 50- μ m micropore of the detection cell to build a phospholipid bilayer. After a stable phospholipid bilayer membrane was formed, α -HL (10 μ L, 15 μ g/mL) was added to the cis end of the detection cell. α -HL assembled itself into a stable single nanopore on the phospholipid membrane. Then, molecules to be tested were added to the cis end of the detection cell, and a 120-mV bias voltage was applied. The current signals were passed through a 5-kHz low-pass filter and then amplified and acquired at a sampling frequency of 250 kHz using an Axon 200B patch clamp amplifier, a DigiData 1440A D/A converter and Clampex 10.4 data recording software (Molecular Devices, Forest City, CA, USA). The resulting data were processed with CM Analyzer (the data analysis software developed in this study), ClampFit 10.4 (Molecular Devices, Forest City, CA, USA) and OriginLab 8.0 (OriginLab Corporation, Northampton, MA, USA).

2.2 Design of the circle-median algorithm

The function module diagram of the nanopore data processing algorithm, Circle-Median, is shown in Fig. 1b and included three modules: data preprocessing, circle-median filtering and transient current jump detection.

Because the bias voltage applied in each experiment varied, the original signals may produce different baseline currents. The data preprocessing module sliced the original signals according to the baseline current value and captured and saved all blockages. In previous research, the baseline current value might have been given in advance. Because the nanopore data obtained in the experiments are a type of time series data, they should contain rich semantic information, including image and language data. Using a deep learning method, the baseline current values can be predicted and calibrated through a trained deep neural network.

In this study, denoising, baseline tracking and other data preprocessing tasks for nanopore data were converted into deep learning tasks. The data preprocessing module automatically calculates the baseline current value. The workflow of this module is as follows:

First, the nanopore data obtained from the experiments were converted into timestamp embedding and numerical embedding by the method reported in the literature [39]. Such an embedding method can allow the encoder to use both global time and local time information. Second, the input embeddings processed in the previous step were input into the transformer [40] encoder and decoder network for calculation. Finally, a fully connected neural network and a softmax layer were connected after the transformer network. This softmax layer maps the output between (0, 1) to indicate the probability that the predicted value is the baseline value.

The transformer network used in this study had 6 transformer blocks, and the number of multihead attention blocks was 8. The training nanopore data were collected from previous experiments.

The captured blockages were analyzed and processed using the Circle-Median filtering module, which can identify downward current jumps in signals. The principle was to replace the value of a given point in a signal with the median of points in its neighborhood, as expressed in Formula (1):

$$x_{(i)} = \text{Med}[x_{(i-N)}, \dots, x_{(i)}, \dots, x_{(i+N)}] \quad (1)$$

where $x_{(i)}$ is the i^{th} point in the signal. The neighborhood size of median filtering was $2N + 1$. This filtering method effectively suppressed noise and can make the detected value closer to the true value. Compared with other linear filtering methods, median filtering can eliminate isolated noise in signals more effectively. However, when a given point, $x_{(i)}$, was located in the head or the tail of the signal and the set neighborhood boundary exceeded the signal boundary, points that could not be obtained in median filtering were replaced with zero. Thus, a large deviation would occur when median filtering was used in the head or the tail of a signal; Circle-Median filtering reduced this deviation. Inspired by the attention mechanism [40], the nanopore events were regarded and all the neighborhood weights were calculated once for each point in the event. Then, for a given point $x_{(i)}$ in the event and a given neighborhood of this point, the circle-median filter principle can be expressed as in Formula (2):

$$x_{(i)} = \begin{cases} \text{Med}[x_{(n-N+i)}, x_{(n-N+i+1)}, \dots, x_{(n)}, x_{(1)}, \dots, x_{(i)}, \dots, x_{(i+N)}], & (i < N) \\ \text{Med}[x_{(i-N)}, \dots, x_{(i)}, \dots, x_{(i+N)}], & (N \leq i \leq n - N) \\ \text{Med}[x_{(i-N)}, \dots, x_{(i)}, \dots, x_{(n)}, x_{(1)}, x_{(2)}, \dots, x_{(N-n+i)}], & (i > n - N) \end{cases} \quad (2)$$

where $x_{(i)}$ is the i^{th} point in the signal; the data size recorded is n ; and the neighborhood size of circle-median filtering is $2N + 1$. This neighborhood size was determined by the sample size n . After Circle-Median filtering was conducted on the nanopore data, we had to separate blockages from current jumps in the data. A common method in previous studies was K-means cluster computation. However, due to the randomness of nanopore data and inclusion of unpredictable currents in this experiment, it was difficult to estimate K in the proposed algorithm. If the initial value was selected improperly, the final clustering analysis was affected. In this study, the Zscore MAD method was used to match circle-median filtering. Blockages and current jumps in the filtered nanopore data were clustered. The principle was to analyze each value in the signal and compare it with the preset threshold to determine whether the point was in current jumps. The Zscore MAD can be expressed in Formulae (3) and (4):

$$\text{MAD} = \text{Med}[|x_{(i)} - \tilde{x}|] \quad (3)$$

$$\text{Zscore} = \left| \frac{0.6745 \times (x_{(i)} - \tilde{x})}{\text{MAD}} \right| \quad (4)$$

where $x_{(i)}$ is the i^{th} point in the signal. When the Zscore was greater than the preset threshold (4.0), we determined that $x_{(i)}$ was in a current jump.

The principle of the transient current jump detection module derived from box plot and quantile in statistics. In detection, first, all signals in blockages were arranged by current, from small to large. The rearranged signals can be expressed in Formula (5):

$$x = [x_{(0)}, x_{(1)}, \dots, x_{(n-1)}] \quad (5)$$

where the total data size of blockages was $n - 1$, and $x_{(0)} < x_{(1)} < \dots < x_{(n-1)}$. The quantile can be expressed in Formula (6):

$$P_m = \left\lceil \frac{m}{100} \times (n - 1) \right\rceil \quad (6)$$

where P_m is the corresponding subscript of the m -quantile of the data in Formula (5): when $m = 0$, P_m was the minimum. When $m = 100$, P_m was the maximum. When transient current jumps were

screened from blockages, the starting and ending times of transient current jumps should be calculated using the method reported previously [41~43].

3. RESULTS AND DISCUSSION

3.1 Simulation signals

First, we used simulation signals to test the algorithm. The simulation signals generated contained 4 current jumps and 7 transient current jumps. Noises whose mean was 0 pA and standard deviations were 0.5, 1, 1.5 and 2 were added sequentially. After synthesis, the simulation signals generated by the system are shown in Fig. 2a. The noise density diagram with four different standard deviations is shown in Fig. 2b. After entering the high-frequency area, noises were distributed uniformly randomly and presented the characteristics of white noises. With an increasing standard deviation of noise, the energy of noise in the frequency domain also increased gradually.

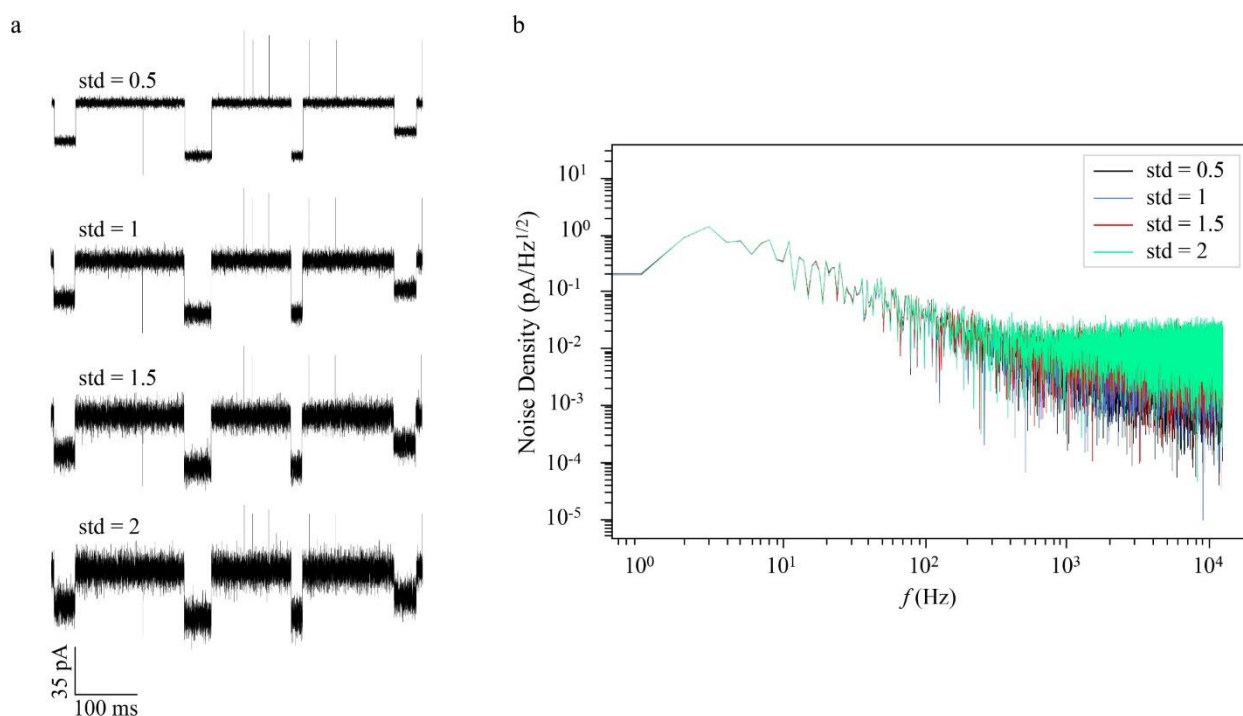


Figure 2. Generation of simulation signals: (a) simulation signals at different noise of std levels: 0.5, 1, 1.5, and 2; the current value of blockade event is 20 pA and contains four current falling events and seven outliers; the current value of current falling event is from 8 to 15 pA; the difference between peak outlier and the current value of blockade event is between 13 and 15 pA.

3.2 Effect of noise on data processing

To test the performance of the algorithm under different noises, the circle-median filtering module in CM Analyzer was first used to test the identification accuracy of the algorithm for current jumps in simulation signals of noises with different standard deviations, as shown in Fig. 2a. The

identification results are shown in Fig. 3a. Then, the transient current jump detection module in CM Analyzer was used to test the identification accuracy of the algorithm for transient current jumps in simulation signals. The results are shown in Fig. 3b.

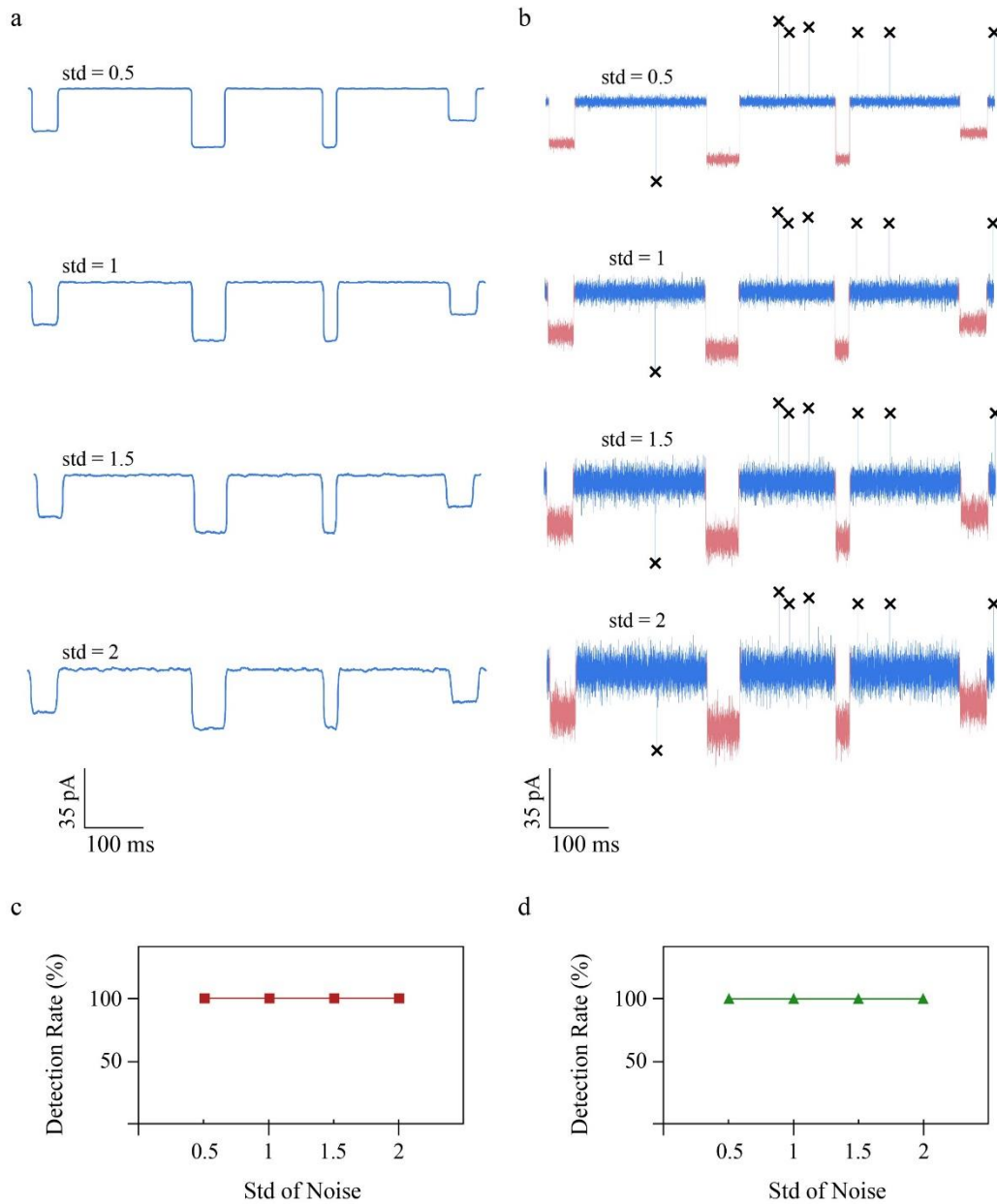


Figure 3. Effects of noise on data analysis: (a) performance of filtering simulation signals with different noise by CM Analyzer; (b) identification results of simulation signals with different noise by CM Analyzer; (c) detection rate of current falling events; and (d) detection rate of outliers.

Concurrently, to determine the accuracy of the identification results and obtain the detection rates of the CM Analyzer for current jumps and transient current jumps under the impact of noise with different standard deviations, we expressed the detection rates using Formula (7):

$$\text{Detection Rate} = \left| \frac{\text{Value}_{\text{Detected}}}{\text{Value}_{\text{True}}} \right| \times 100\% \quad (7)$$

where $Value_{True}$ is the true or transient current jump contained in the blockages; and $Value_{Detected}$ is the detected or transient current jump identified by CM Analyzer.

Fig. 3a shows that with an increasing standard deviation, only marginal fluctuation was added to the signal waveform processed by the circle-median filtering module; thus, no marked impact was observed. The same situation was also found in Fig. 3b. Thus, an increasing standard deviation did not affect the identification of transient current jumps. The statistical summary charts of the detection rate of the CM Analyzer for the simulation signals are shown in Figs. 3c and 3d. The red curve in Fig. 3c shows the detection rate of the algorithm for current jumps. The green curve in Fig. 3d shows the detection rate of the algorithm for transient current jumps. When the standard deviation of noise increased from 0.5 to 2, the detection rates of the CM Analyzer for current jumps and transient current jumps remained 100%. The data processing of the CM Analyzer maintained high identification accuracy under high noise.

3.3 Application in the analysis of nanopore experimental data

To test the performance of the algorithm in more detail, data collected in the nanopore experiment were processed in CM Analyzer to verify the identification accuracy of the algorithm for transient current jumps produced by weak interactions between biomolecules in the nanopore single-molecule detection experiment.

A total of 81460736 data points were recorded in the experiment, and the experiment duration was approximately 325 s. Data from 92-102 s were selected for analysis. The original data are shown in Fig. 4a. Comparing the noise density curve of the experimental data in Fig. 4b and the noise density curve of the simulation signals in Fig. 2b, the energy of the noise at high frequencies in the experimental data must be smaller than that in the simulation signals so that the CM Analyzer software can satisfy the data analysis in the nanopore experiment. Blockages recorded throughout the experiment are shown in Fig. 4c. The scatter diagram shows statistics about the degree (the ratio between blocked current value and open current value) and duration of blockage. After analyzing the original data using CM Analyzer, the analysis results of four blockages (red marks in Fig. 4a) are shown in Figs. 4d, 4e, 4f and 4 g. In these figures, the blue curve represents blockage, the red curve represents a current jump in this blockage, and the black mark represents a transient current jump in this blockage.

Fig. 4 shows that the CM Analyzer can accurately identify transient current jumps produced by weak interactions between biomolecules in nanopore single-molecule detection experiments.

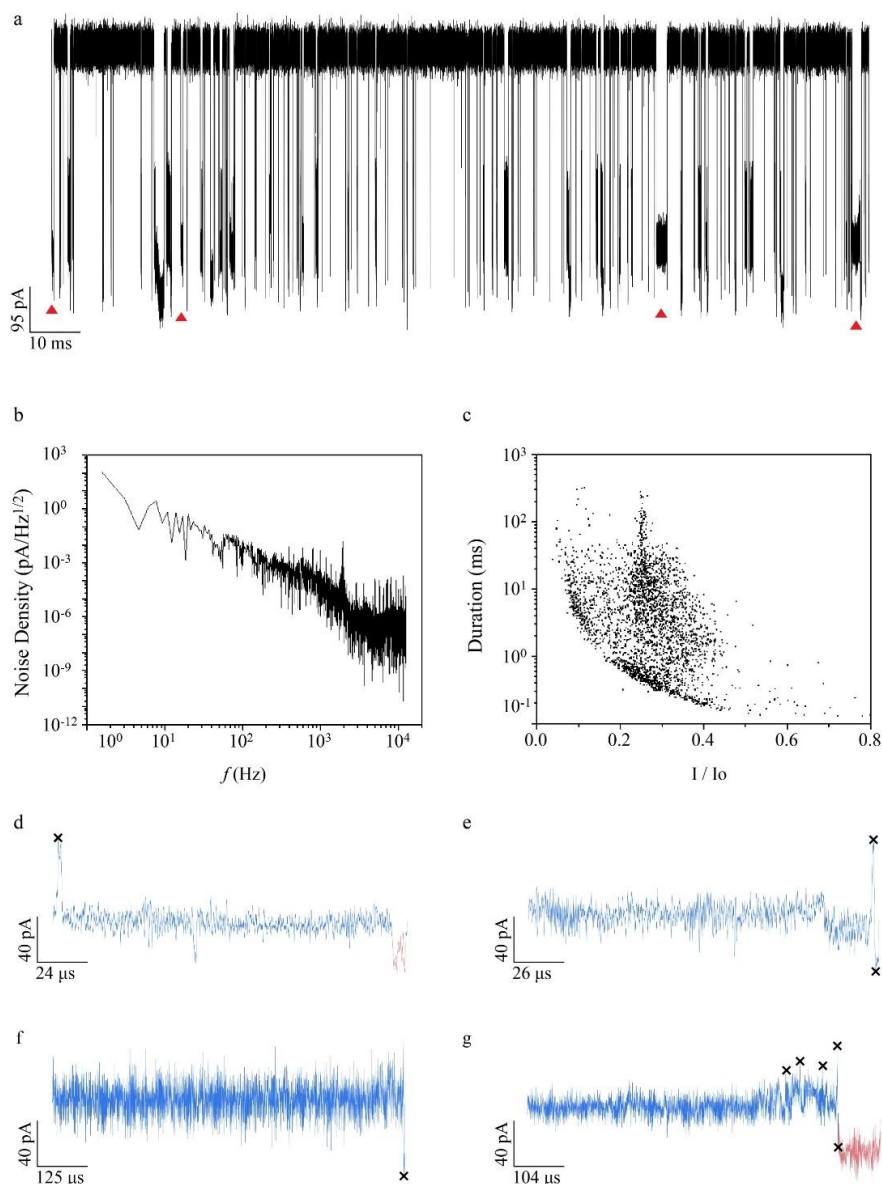


Figure 4. Real experimental data records: (a) original data of the nanopore experiment; (b) noise power spectral density curve of real experimental data; (c) I/I_0 –duration scatter plot, where I/I_0 is the average residual current blockade, I_0 is the average current of an empty pore, and I is the average residual current of a blocked pore; (d–g) analysis of four blockade events, where the blue curve represents blockade events, the red curve represents current falling events, and black x markers represent outliers.

The nanopore data processing process can be divided into four steps [31]: denoising raw data from experiments; identification of current events; extraction features of events; and analysis of features to infer the properties of analytes. Low-pass filters, Kalman filters, and consensus filters are used to denoise the experimental raw data [32–33, 44]. Cumulative Sums, MOSAIC, DBC and other methods are used to process steps 2 or 3 [34–35, 38]. In recent years, with the development of machine learning, approaches such as hidden Markov models, fuzzy-c means, K-means, and AdaBoost have also been used to process nanopore data with good results [36–37, 45]. Currently, deep learning is also being applied

more often to nanopore data processing. For example, [46~48] investigated feature extraction and inferring classifiers using Bi-Path, U-Net, and ResNet in deep neural networks.

In previous studies, the 4 steps of nanopore data processing were usually performed independently. In this study, a transformer in deep neural networks was incorporated into nanopore data processing, and the user can use one network to complete data processing steps 1 and 2 without intervention and input. The advantage of splitting the nanopore data processing task into four subtasks to perform is that it allows a more focused effort, while the disadvantage is also marked. The split tasks lose their entirety, and each task must correspond to one network, model, and different model parameters, which increases the complexity of the mission by a factor of 4, while the model parameters increase by much more than a factor of 4. Therefore, we want to use one model to solve all four subtasks.

There are already many multitask models in the fields of computer vision and natural language processing that can solve multiple subtasks with a unified network. Multitask models can use fewer parameters to solve more problems, which is one of the trends of deep learning. Currently, there is no well-designed multitask processing model to manage time-series data, which is a primary direction of future research. We explained the 4 steps of the nanopore data processing process as 4 subtasks that must be solved by a deep learning network and intend to build an end-to-end nanopore data processing model that uses a unified deep neural network to process these four subtasks simultaneously.

4. CONCLUSION

In this paper, a nanopore single-molecule detection algorithm called CM Analyzer is designed to provide a new method to study data processing systems in nanopore single-molecule detection experiments. Compared with other nanopore data processing algorithms, the proposed algorithm can identify weak interactions between biomolecules and accurately analyze current jumps and transient current jumps in experiments. Through tests on simulation signals and in real experiments, the detection rates of the algorithm for current jumps and transient current jumps remain stable under the impact of high noise, satisfy the precision demand of data processing and promote the development of nanopore single-molecule detection techniques.

ACKNOWLEDGMENTS

This study was supported by the National Natural Science Foundation of China (Nos. 21327807, 21421004), Program of Shanghai Subject Chief Scientist (No. 15XD1501200) and the Fundamental Research Funds for the Central Universities (Nos. 222201718001, 222201717003).

References

1. Y.L. Ying, J. Zhang, R. Gao, Y.T. Long, *Angew. Chem. Int. Ed.*, 52 (2013) 13154.
2. Y.T. Long, J.J. Gooding, *ACS Sens.*, 1 (2016) 963.
3. J. Wang, M.Y. Li, J. Yang, Y.Q. Wang, X.Y. Wu, J. Huang, Y.L. Ying, Y.T. Long, *ACS Cent. Sci.*, 6 (2020) 76.
4. Y.L. Ying, J. Wang, A.R. Leach, Y. Jiang, R. Gao, C. Xu, M.A. Edwards, A.D. Pendergast, H.

- Ren, C.K.T. Weatherly, W. Wang, P. Actis, L. Mao, H.S. White, Y.T. Long, *Sci. China Chem.*, 63 (2020) 589.
5. Q. Li, Y. Lin, Y.L. Ying, S.C. Liu, Y.T. Long, *Sci. Sin.: Chim.*, 47 (2017) 971.
 6. C. Cao, D. Liao, Y. Ying, Y. Long, *Acta Chim. Sinica*, 74 (2016) 734.
 7. Z. Hu, J. Du, Y. Ying, Y. Peng, C. Cao, Y.T. Long, *Acta Chim. Sinica*, 75 (2017) 1087.
 8. W.L. Hsu, H. Daiguji, *Anal. Chem.*, 88 (2016) 9251.
 9. X. He, P. Wang, L. Shi, T. Zhou, L. Wen, *Electrophoresis*, 42 (2021) 2197.
 10. L. Xue, H. Yamazaki, R. Ren, M. Wanunu, A.P. Ivanov, J.B. Edel, *Nat. Rev. Mater.*, 5 (2020) 952.
 11. T. Ding, J. Yang, V. Pan, N. Zhao, Z. Lu, Y. Ke, C. Zhang, *Nucleic Acids Res.*, 48 (2020) 2791.
 12. I. Djurišić, M.S. Dražić, A.Ž. Tomović, M. Spasenović, Ž. Šljivančanin, V.P. Jovanović, R. Zikic, *ACS Appl. Nano Mater.*, 3 (2020) 3034.
 13. S.L. Amarasinghe, S. Su, X. Dong, L. Zappia, M.E. Ritchie, Q. Gouil, *Genome Bio.*, 21 (2020) 30.
 14. B. Cressiot, E. Braselmann, A. Oukhaled, A.H. Elcock, J. Pelta, P.L. Clark, *ACS Nano*, 9 (2015) 9050.
 15. T.C. Sutherland, Y.T. Long, R.I. Stefureac, I. Bediako-Amoa, H.B. Kraatz, J.S. Lee, *Nano Lett.*, 4 (2004) 1273.
 16. P. Waduge, R. He, P. Bandarkar, H. Yamazaki, B. Cressiot, Q. Zhao, P.C. Whitford, M. Wanunu, *ACS Nano*, 11 (2017) 5706.
 17. Y. Ge, M. Cui, Q. Zhang, Y. Wang, D. Xi, *Nanoscale Adv.*, 4 (2022) 3883.
 18. A. Asandei, I. Schiopu, M. Chinappi, C.H. Seo, Y. Park, T. Luchian, *ACS Appl. Mater. Interfaces*, 8 (2016) 13166.
 19. Y.L. Ying, J. Zhang, F.N. Meng, C. Cao, X. Yao, I. Willner, H. Tian, Y.T. Long, *Sci. Rep.*, 3 (2013) 1662.
 20. S. Wen, T. Zeng, L. Liu, K. Zhao, Y. Zhao, X. Liu, H.C. Wu, *J. Am. Chem. Soc.*, 133 (2011) 18312.
 21. C. Yang, L. Liu, T. Zeng, D. Yang, Z. Yao, Y. Zhao, H.C. Wu, *Anal. Chem.*, 85 (2013) 7302.
 22. H. Che, Y. Li, S. Zhang, W. Chen, X. Tian, C. Yang, L. Lu, Z. Zhou, Y. Nie, *Sens. Actuators, B*, 324 (2020) 128641.
 23. G.M. Roozbahani, Y. Zhang, X. Chen, M.H. Soflaee, X. Guan, *Analyst*, 144 (2019) 7432.
 24. B.Y. Yan, Z. Gu, R. Gao, C. Cao, Y.L. Ying, W. Ma, Y.T. Long, *Chin. J. Anal. Chem.*, 43 (2015) 971.
 25. J.D. Uram, K. Ke, M. Mayer, *ACS Nano*, 2 (2008) 857.
 26. A. Balijepalli, J. Etteedgui, A.T. Cornio, J.W.F. Robertson, K.P. Cheung, J.J. Kasianowicz, C. Vaz, *ACS Nano*, 8 (2014) 1547.
 27. A. Balijepalli, J. Etteedgui, A.T. Comb, J.W.F. Robertson, K.P. Cheung, J.J. Kasianowicz, C. Vaz, *ACS Nano*, 9 (2015) 12583.
 28. Z. Gu, H. Wang, Y.L. Ying, Y.T. Long, *Sci. Bull.*, 62 (2017) 1245.
 29. F.L.R. Lucas, K. Willems, M.J. Tadema, K.M. Tych, G. Maglia, C. Wloka, *ACS Omega*, 7 (2022) 26040.
 30. S.C. Liu, B.K. Xie, C.B. Zhong, J. Wang, Y.L. Ying, Y.T. Long, *Rev. Sci. Instrum.*, 92 (2021) 121301.
 31. C. Wen, D. Dematties, S.L. Zhang, *ACS Sens.*, 6 (2021) 3536.
 32. D. Pedone, M. Firnkies, U. Rant, *Anal. Chem.*, 81 (2009) 9689.
 33. C. Raillon, P. Granjon, M. Graf, L.J. Steinbock, A. Radenovic, *Nanoscale*, 4 (2012) 4916.
 34. J.H. Forstater, K. Briggs, J.W.F. Robertson, J. Etteedgui, O. Marie-Rose, C. Vaz, J.J. Kasianowicz, V. Tabard-Cossa, A. Balijepalli, *Anal. Chem.*, 88 (2016) 11900.
 35. N. Zhang, Y.X. Hu, Z. Gu, Y.L. Ying, P.G. He, Y.T. Long, *Chin. Sci. Bull.*, 59 (2014) 4942.
 36. Y.H. Lai, B.Y. Yan, H.F. Wang, *Journal of East China University of Science and Technology (Natural Science Edition)*, 43 (2017) 220.

37. J. Zhang, X. Liu, Y.L. Ying, Z. Gu, F.N. Meng, Y.T. Long, *Nanoscale*, 9 (2017) 3458.
38. H.F. Wang, F. Huang, Z. Gu, Z.L. Hu, Y.L. Ying, B.Y. Yan, Y.T. Long, *Chin. J. Anal. Chem.*, 46 (2018) 843.
39. H. Zhou, S. Zhang, J. Peng, S. Zhang, J. Li, H. Xiong, W. Zhang, Informer: Beyond Efficient Transformer for Long Sequence Time-Series Forecasting, 35th AAAI Conference on Artificial Intelligence, USA, 2021, 11106.
40. A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A.N. Gomez, L. Kaiser, I. Polosukhin, I. Guyon, U.V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, R. Garnett, Attention Is All You Need, 31st Conference on Neural Information Processing Systems (NIPS), USA, 2017.
41. Z. Gu, Y.L. Ying, C. Cao, P. He, Y.T. Long, *Anal. Chem.*, 87 (2015) 907.
42. W.B. Dunbar, *Anal. Chem.*, 87 (2015) 10650.
43. Z. Gu, Y.L. Ying, C. Cao, P. He, Y.T. Long, *Anal. Chem.*, 87 (2015) 10653.
44. B. Yan, H. Cui, J. Zhou, H. Wang, *Quim. Nova*, 43 (2020) 837.
45. L. Loeff, J.W.J. Kerssemakers, C. Joo, C. Dekker, *Patterns*, 2 (2021) 100256.
46. D. Dematties, C. Wen, M.D. Pérez, D. Zhou, S.L. Zhang, *ACS Nano*, 15 (2021) 14419.
47. M. Tsutsui, T. Takaai, K. Yokota, T. Kawai, T. Washio, *Small Methods*, 5 (2021) 2100191.
48. X. Fu, Y. Wan, X. Li, Y. Ying, Y. Long, Analysis and classification of nanopore data based on feature-level multi-modality, 13th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), China, 2020, 692.