SPIKE SORTING BASED ON LOW-RANK AND SPARSE REPRESENTATION

Libo Huang‡, Bingo Wing-Kuen Ling*, Yan Zeng† and Lu Gan‡*

*School of Information Engineering, Guangdong University of Technology, 51006, China
† School of Computers, Guangdong University of Technology, 51006, China
‡Department of Electronic and Electrical Engineering, Brunel Univeristy London, UB8 3PH, UK
www.huanglibo@gmail.com; yongquanling@gdut.edu.cn
yanazeng013@gmail.com; lu.gan@brunel.ac.uk

ABSTRACT

As the first step to study the coding mechanism and synergistic behaviour of neurons, spike sorting plays an important role in the neurosciences research community. Despite many empirical successes in spike sorting models, there are still sufferings from the overlapping and noise corruption problems. To ease these situations, in this paper, we present an efficient and effective method with the help of optimization theory. Firstly, by introducing the low-rank strategy, the global structure underlying the spike data could be discovered. Secondly, by engaging the sparse coding to balance the noise, the proposed model is robust in the overlapping and noise spike sorting scenario. We have conducted experiments on the Wave_clus dataset compared with two state of the art models. The results verify the efficacy of our scheme and confirm the claims above.

***Index Terms*—** Low-rank, overlapping spike, spike sorting, sparse optimization

1. INTRODUCTION

Implanting single- or micro-electrodes array into the brain to record the local electrical activity around the probe tip is one of the most reliable methods for monitoring neural activities and further unveiling the behaviour of neuronal networks [1]. The action potential issued by neurons is defined as spike. Since we cannot guarantee the recorded signal issued only from single neuron rather than multiple ones, spike sorting, which aims at identifying and assigning spikes to the correct neurons based on the features extracted from themselves, is inevitable and important [2]. Especially when it makes tremendous contributions in neural coding [3], diseases treating [2], and home-care surveillance [4].

The conventional spike sorting technology mainly consists of four steps: filtering, spike detection, feature extraction, and cluster analysis, while the spike detection and the cluster analysis occasionally include manual post-verification [5]. With the increasing development of neuroscience, spike sorting technology has always been moving towards the invention of fast, accurate, and non-manual intervention directions [2]. However, the overlapping and noise disruptions in the recorded signals are becoming serious along with the development of the signal acquiring device [1], which severely deteriorates the precision of sorting and requires tremendous computational resources. So how to classify the overlapping and noise corrupted spikes accurately and automatically has attracted more attention from the research community [5, 6].

Modelling these problems into a linear combination of the dominating waveforms plus noise [6, 3] is a widely accepted methodology in the last decades. As one of the most successful linear models, sparse representation (SR) [6] could excavate the data's local column-wise structure. However, it's hard to acquire dominating waveforms with high quality, which will predominately determine the sorting results.

On the other hand, by fully automatic transaminating the spikes to the orthogonal and high variance dimension, Principal Component Analysis (PCA) is also broadly used in the spike sorting [7] due to its possible capability of handling overlapping spikes [5]. Note that when the noises exist, it is severely hard for PCA to distinguish the spikes [5]. This is mainly because the PCA-based methods are under the assumption that the data is approximately drawn from a single low-rank subspace, whereas the spikes are more likely to lie a mixture of several low-rank subspaces [8]. This implies to use the low-rank representation (LRR) to model the spikes data, which is more reasonable and capable [9].

Thus, in this paper, we propose a Low-Rank and Sparse Representation Model, abbreviated as LRSR, to solve the spike sorting problem, especially when the spikes are over-

This work was supported in part by the National Nature Science Foundation of China (no. U1701266, no. 61372173 and no. 61671163), the Team Project of the Education Ministry of the Guangdong Province (no, 2017KCXTD011), the Guangdong Higher Education Engineering Technology Research Center for Big Data on Manufacturing Knowledge Patent (no. 501130144), and Hong Kong Innovation and Technology Commission, Enterprise Support Scheme (no. S/E/070/17), and the China Scholarship Council (CSC).

978-1-7281-1331-9/20/$31.00 ©2020 IEEE

lapping and noise corrupted. It combines the superior performance of local structure from the SR, which aids in sorting the noise corrupted spikes, and avoids acquiring the dominating waveforms. Furthermore, it possesses the advantages of LRR for its global representation for overlapping spikes.

The remainder of this paper is organized as follows. In Section 2, the preliminary related to our proposed model, LRSR, is outlined. Applying LRSR to solve the spike sorting problems is described in detail in Section 3. Then in Section 4, we conduct several experiments to verify the achievement of LRSR compared with two state of the art works. Finally, conclusions are drawn in Section 5.

2. PRELIMINARY

In this section, two preliminary methods are emphasized, including the low-rank and the sparse optimization models.

2.1. Low-Rank Optimization

Matrix minimization theory [10] demonstrates that a noise corrupted observation matrix $Y = X + E$, where X is an unknown low-rank matrix and E represents the noise, can be formulated as a regularized rank optimization problem,

$$\min_X \tau \,\mathrm{rank}(X) + \frac{1}{2}\|X - Y\|_F^2 \tag{1}$$

where $\mathrm{rank}(X)$ is equal to the rank of matrix X and $\|\cdot\|_F$ refers to the Frobenius norm for the noise [6]. Since Eq.(1) is nonconvex and NP-hard [10], an alternative is relaxing tightly the discrete rank operator to the nuclear norm $\|\cdot\|_*$, which is the sum of the singular values. That is,

$$\min_X \tau\|X\|_* + \frac{1}{2}\|X - Y\|_F^2 \tag{2}$$

Note that Eq.(2) is a convex optimization problem and the optimal solution is given by $\hat{X} = U\mathcal{D}_\tau(\Sigma)V^T$, where U and V are obtained by the singular value decomposition of Y and $\mathcal{D}_\tau(\Sigma)$ is a singular value shrinkage operator [11]. It is worth mentioning that the optimizer $\hat{X}$ is block diagonal under some mild assumptions. Hence, various spectral clustering algorithms [12] then could be applied on $\hat{X}$ instead of directly performing clustering analysis on the noise corrupted observation matrix Y.

2.2. Sparse Optimization

For ease of presentation, we review a simple sparse model here, recovering a matrix Y as follows,

$$\min_X \tau\|X\|_0 + \frac{1}{2}\|X - Y\|_F^2 \tag{3}$$

where X is a sparse matrix pursuing $X = Y$, $\|\cdot\|_0$ denotes the l_0-norm, and $\tau > 0$ is a parameter that tradeoffs $\|X\|_0$ and $\frac{1}{2}\|X - Y\|_F^2$. However, l_0-norm is a discrete count operator, which is non-convex and NP-hard for optimization, and replacing it with l_1-norm is reasonable and solvable [13]. The problem (3) then is recast to,

$$\min_X \tau\|X\|_1 + \frac{1}{2}\|X - Y\|_F^2 \tag{4}$$

where $\|\cdot\|_1$ stands for l_1-norm. Considering the first-order condition of Eq.(4) elementwisely, the optimal solution then is given by the soft threshold (shrinkage) operator [14],

$$\hat{X} = (\mathrm{abs}\,(Y) - \tau)_+ \,\mathrm{sign}(Y) \tag{5}$$

where abs and $sign$ represent the absolute value operator and the sign operator, respectively, and $(\cdot)_+ = \max\{\cdot, 0\}$.

3. MODEL AND METHODS

Detailed descriptions of our proposed methods are presented in this section. Specifically, we demonstrate firstly the derivation of our model, LRSR. The solution of LRSR is then proposed. Finally, implementing LRSR in spike sorting is given.

3.1. LRSR

After filtering and spike detection which are the same as conventional spike sorting [2], the spike matrix X consisting of D rows (features) and N columns (spikes) is acquired. In practical situations, some of the spikes may be corrupted by the noise originated from the remote neuron clusters, probe devices, and other sources. Hence, the detected spikes could be modeled as $X = Z + E$, where Z indicates the pure spike and E presents the noise. In addition, the overlapping phenomenon usually occurs due to the location distribution and electrical discharge characteristics of neurons [5] as well as the intrinsic structure of acquisition device [2]. Hence, similar as an endeavor of SR [3, 6], we also introduce an appropriate dictionary, A, to get a better model as follows,

$$X = AZ + E \tag{6}$$

Note that the dictionary A in [3, 6] is constructed by the "ideal" templates which can not guarantee success to get from the original corrupted signal. By taking the advantage of matrix decomposition [8], we adopt XQ as the dictionary instead. Here, Q is the orthonormal basis of X^T where the superscript T refers to a matrix transposition operator. It can be computed in advance by orthogonalizing the rows of X. If the final solution Z^* is obtained, then QZ^* can be used to recover Z accurately by $X = XQZ^* + E$. For more details, please refer to [8].

To cluster the collected signals X, i.e. to explore the underlying structure of observation data, it is reasonable to restrict the clean data Z to a low-rank matrix [10]. Besides, the noises mentioned above are randomly disturbed over pure signals after filtering and detection. Hence, we constrain the noise with the sparse condition, the sorting problem then is,

$$\begin{aligned} \min_{(Z,E)} \ & \operatorname{rank}(Z) + \lambda \|E\|_0 \\ s.t.\ & X = AZ + E \end{aligned} \tag{7}$$

where λ is a parameter that weight the intensity of noise interference, and while it tends to be infinity, the model does not hold the structure of pure data.

As mentioned in Section 2, Eq. (7) can be reformulated as the following convex optimization problem,

$$\begin{aligned} \min_{(Z,E)} \ & \|Z\|_* + \lambda \|E\|_1 \\ s.t.\ & X = AZ + E \end{aligned} \tag{8}$$

3.2. Solution of LRSR

With the help of Inexact Augmented Lagrange Multipliers (IALM) [14], the convex problem (8) can be solved efficiently as follows. Firstly, we introduce an auxiliary variable J into (8) and get,

$$\begin{aligned} \min_{(J,Z,E)} \ & \|J\|_* + \lambda \|E\|_1 \\ s.t.\ & X = AZ + E, \ \ Z = J \end{aligned} \tag{9}$$

Then, the corresponding unconstrained optimization problem of (9) is defined as,

$$\begin{aligned} \min_{(J,Z,E)} \ & \|J\|_* + \lambda\|E\|_1 + \langle Y_1, X - AZ - E\rangle + \langle Y_2, Z - J\rangle \\ & + \frac{\mu}{2}\left(\|X - AZ - E\|_F^2 + \|Z - J\|_F^2\right) \end{aligned} \tag{10}$$

where Y_1, Y_2 are the Lagrange multipliers, $\mu > 0$ stands for the penalty parameter, and $\langle A, B\rangle$ denotes the inner product between matrices A and B. Now such problem can be minimized by alternating direction IALM algorithm with details as follows.

3.2.1. *Update J*

Given matrices Z and E, J is updated by,

$$J^{t+1} = \arg\min_{J^t} \ \|J^t\|_* + \left\langle Y_2, Z^t - J^t\right\rangle + \frac{\mu}{2}\left\|Z^t - J^t\right\|_F^2 \tag{11}$$

where t is the iterative index. The inner product, the square of Frobenius norm, and the trace operators are equivalent for the matrices, e.g. $\langle A, B\rangle = \operatorname{tr}(A^T B)$ between matrices A and B, and $\langle A, A\rangle = \|A\|_F^2$. Hence, Eq.(11) is expressed as,

$$\begin{aligned} J^{t+1} &= \arg\min_{J^t} \|J^t\|_* + \operatorname{tr}\left(Y_2^T Z^t - Y_2^T J^t\right) \\ &\quad + \frac{\mu}{2}\operatorname{tr}\left({Z^t}^T Z^t - 2{Z^t}^T J^t + {J^t}^T J^t\right) \\ &= \arg\min_{J^t} \frac{1}{\mu}\|J^t\|_* + \frac{1}{2}\left\|J^t - \left(Z^t + \frac{Y_2}{\mu}\right)\right\|_F^2 \end{aligned} \tag{12}$$

It is clear to see that J^{t+1} in (12) can be easily determined by the singular value shrinkage operator on $(Z^t + {}^{Y_2}/_{\mu})$ as mentioned in the Section 2.1.

3.2.2. *Update Z*

Given matrices J and E, Z is updated as follows,

$$\begin{aligned} Z^{t+1} &= \arg\min_{Z^t} \left\langle Y_1, X - AZ^t - E^t\right\rangle + \left\langle Y_2, Z^t - J^{t+1}\right\rangle \\ &\quad + \frac{\mu}{2}\left(\|X - AZ^t - E^t\|_F^2 + \|Z^t - J^{t+1}\|_F^2\right) \\ &= \arg\min_{Z^t} \operatorname{tr}\Big[\frac{\mu}{2}{Z^t}^T\left(A^T A + I\right)Z^t + \Big(-Y_1^T A \\ &\quad + Y_2^T - \mu X^T A + \mu {E^t}^T A - \mu {J^{t+1}}^T\Big)Z^t\Big] \end{aligned} \tag{13}$$

where I is the identity matrix with compatible dimension. By engaging the trace operator's properties and the first-order condition, Z^{t+1} can be updated by,

$$\begin{aligned} Z^{t+1} = \left(A^T A + I\right)^{-1}\Big(&\frac{A^T Y_1 - Y_2}{\mu} + A^T X - A^T E^t \\ &+ J^{t+1}\Big) \end{aligned} \tag{14}$$

3.2.3. *Update E*

Given matrices J and Z, E is updated as follows,

$$\begin{aligned} E^{t+1} &= \arg\min_{E^t} \lambda\|E^t\|_1 + \left\langle Y_1, X - AZ^{t+1} - E^t\right\rangle \\ &\quad + \frac{\mu}{2}\left\|X - AZ^{t+1} - E^t\right\|_F^2 \end{aligned} \tag{15}$$

Similar to the derivation of (13), problem (15) is reformulated as,

$$E^{t+1} = \arg\min_{E^t} \frac{\lambda}{\mu}\|E^t\|_1 + \left\|E^t - \left(\frac{Y_1}{\mu} + X - AZ^{t+1}\right)\right\|_F^2 \tag{16}$$

This problem is a typical sparse optimization described in Section 2.2. Thus the solution of it can be obtained by applying soft-threshold rule as follows,

$$\begin{aligned} E^{t+1} = &\left(\operatorname{abs}\left(\frac{Y_1}{\mu} + X - AZ^{t+1}\right) - \frac{\lambda}{\mu}\right)_+ \\ &\cdot \operatorname{sign}\left(\frac{Y_1}{\mu} + X - AZ^{t+1}\right) \end{aligned} \tag{17}$$

Finally, the overall procedures for solving the LRSR problem are summarized in Algorithm 1.

3.3. Spike Sorting with LRSR

We perform the same first two processions as [15], that is filtering with the four-pole Butterworth band-pass filter to 300-6000Hz, and spike detection with automatic threshold.

Algorithm 1 The LRSR based algorithm

Input: The spikes matrix, $X \in \Re^{D\times N}$, and the parameter λ.

Output: The block-diagonal matrix, $Z^* \in \Re^{N\times N}$, and the noise matrix, $E^* \in \Re^{D\times N}$.

1: Initialization: $Q = \mathrm{orth}(X^T) \in \Re^{N\times M}$, $A = XQ \in \Re^{D\times M}$, $Z^0 = J^0 = Y_2^0 = 0_{N\times N}$, $E^0 = Y_1^0 = 0_{D\times N}$, $\mu^0 = 10^{-6}$ and $t = 0$.

2: Update J, Z, and E based on $\mathcal{D}_{1/\mu}\left(Z^t + Y_2/\mu\right)$, (14), and (17), respectively.

3: Update other parameters:
$Y_1^{t+1} = Y_1^t + \mu^t\left(X - AZ^{t+1} - E^{t+1}\right)$,
$Y_2^{t+1} = Y_2^t + \mu^t\left(Z^{t+1} - J^{t+1}\right)$,
$\mu^{t+1} = \min\left(1.1\times\mu^t, 10^6\right)$.

4: Quit the loop if the convergence criterion is met; Otherwise, set $t = t+1$ and go back to step 2.

5: Return $Z^* = QZ^{t+1}$ and $E^* = X - AZ^{t+1}$.

Then, by taking advantage of LRSR method, the structures including the global low-rank of pure spikes and the sparse underlying noises of raw data X are exploited. Afterwards, feature extraction and clustering of the spike will be performed simultaneously with the help of spectral clustering theory. Specifically, Ng-Jordan-Weiss [12] spectral clustering algorithm is engaged in this paper.

Besides, applying the skinny SVD [8] on the block diagonal matrix Z to reduce the computational complexity further improves clustering performance [8]. Overall, the spike sorting based on LRSR method is outlined in Algorithm 2.

Algorithm 2 spike sorting based on LRSR

Input: The raw data signal, the cluster number k, and the parameter λ.

Output: The cluster membership.

1: Apply the Butterworth bandpass filter with passband located at 300-6000Hz to raw data signal and perform the automatic threshold detection to get the spikes matrix X.

2: Normalize each feature vector in X such that the values of its elements are between 0.2 and 0.8.

3: Execute Algorithm 1 with matrix X and parameter λ to get the approximate block diagonal matrix Z.

4: Construct the intermediate matrix P with normalized rows by $U\Sigma^{1/2}$, where U and Σ are got from performing skinny SVD on Z, i.e. $Z = U\Sigma V^T$.

5: Define the affinity matrix W with its (i,j)-element as $\left(\left[PP^T\right]_{i,j}\right)^2$.

6: Perform Ng-Jordan-Weiss algorithm on matrix W with cluster number k.

Table 1: Wave_clus Dataset

DS	NL	SN(O)	Num/Clusters (Overlapped)
D1	005	3383(767)	1115(244) 1113(256) 1155(267)
	010	3448(810)	1164(260) 1155(269) 1129(281)
	015	3472(812)	1159(275) 1172(260) 1141(277)
	020	3414(790)	1136(267) 1099(257) 1179(266)
D2	005	3364(829)	1120(271) 1109(274) 1135(284)
	010	3462(720)	1187(230) 1136(238) 1139(252)
	015	3440(809)	1142(284) 1113(262) 1185(263)
	020	3493(777)	1151(260) 1195(277) 1147(240)

4. EXPERIMENTS AND ANALYSIS

We will conduct experiments in this section to corroborate the efficiency and effectiveness of LRSR.

4.1. Wave_clus Dataset and Benchmark Methods

Wave_clus dataset is an online access[1] and widely-used dataset since it not only provides ground truth but also contains different signals to noise ratios (inverse to be noise level) and overlapping spikes. The waveforms in it were recorded from the neocortex and basal ganglia and the noise was constructed to duplicate the realistic background activity, with details shown in the Tabel 1. Here, “DS”, “NL”, “SN(O)”, and “Num/Clusters” stand for the name, the noise levels, the number of spikes (overlapped spikes), and the number of spikes in each cluster, respectively, of dataset. For ease expression, we denote, for example, “D1_noise005” as a dataset with 5% noise level in the rest of paper. Note that three neurons were simulated in each recording and more information can be found in [15].

The Sparse Coding[2] (SC) [6] and High Accuracy Spike Sorting (HASS) [16] are employed as our benchmark methods. Reasons are as follows, 1) SC succeeds in promoting LRR strategy [9], which is a comparable choice for LRSR based method. 2) HASS reveals the nonlinear relationships between the spike features formed by the wavelet packet decomposition coefficients and the corresponding label vector by engaging mutual information, which has demonstrated a remarkable classification performance.

4.2. Parameter Analysis

In this subsection, we will investigate the impact of parameter λ in the LRSR model. Generally, this parameter depends on the prior knowledge of the noise level of signals, which could be estimated by many existing methods [17]. When the noise is heavy, we use relatively small λ, and vice versa. Experimental settings are as follows. We picked four datasets,

[1] https://www2.le.ac.uk/centres/csn/research-2/spike-sorting
[2] Source code access at https://github.com/yangkai12/SpikeSorting

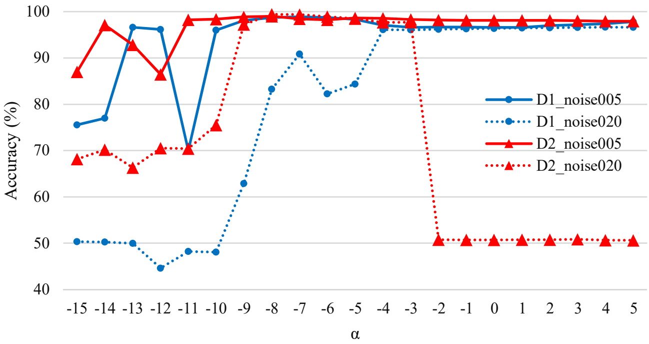

Fig. 1: Effects of the parameter $\lambda = 2^{\alpha}$ on the accuracy of LRSR.

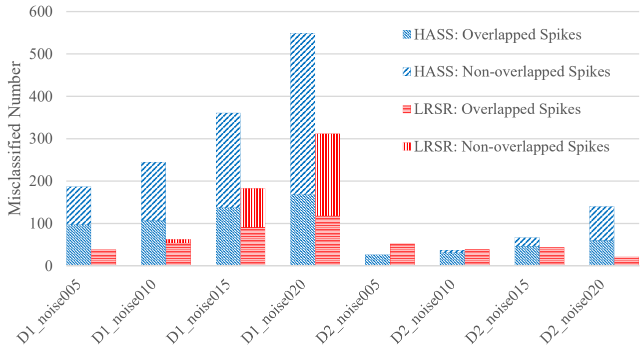

Fig. 2: Misclassified numbers of the overlapped and non-overlapped spikes of HASS and LRSR.

Table 2: Overall comparisons against benchmark methods

DS	NL	HASS	SC	LRSR
D1	005	94.47%	92.99%	**98.99%**
	010	92.89%	63.69%	**98.17%**
	015	89.60%	37.30%	**94.73%**
	020	83.92%	34.68%	**90.86%**
D2	005	**99.23%**	98.34%	98.45%
	010	**98.93%**	65.77%	98.87%
	015	98.05%	70.99%	**98.72%**
	020	95.99%	58.98%	**99.40%**
average time(s)		398.67	141.39	**29.42**

including two lowest and two highest noise level datasets, i.e. D1_noise005, D1_noise020, D2_noise005, and D2_noise020. We set $\lambda = 2^{\alpha}$ where α's range is $[-15, -14, \ldots, 5]$ and other settings are the same as Algorithm 2.

As shown in Figure 1, for the high noise level (20%) disrupt dataset, though LRSR could perform well in D1, it failed in D2 when λ is larger than 2^{-2}. But when setting a smaller λ, such two high noise level disrupted datasets, i.e. D1_noise020 and D2_noise020, obtain satisfactory performances. Similar conclusions are drawn for the small noise corrupted dataset. For ease of presentation, we fix λ to be 2^{-7} in the following simulations.

4.3. Overall Comparisons with Benchmark Methods

In this section, we will conduct experiments over eight datasets compared with two benchmark methods. We choose two measurements, the accuracy and the average time. Accuracy denotes the hit targets' percentage[3] of total spikes' number while the average time is counting for the average processing time of per dataset over all datasets.

[3]Specifically, when the experimental result is the same as the ground truth, it will be treated as a hit target.

As shown in Table 2, LRSR outperforms HASS and SC in general, and more robust than other methods in the accuracy. This is mainly because the low-rank restriction in LRSR can capture the global structure underlying the whole spike data, and the sparse constraint can model the noise well. In particular, accuracies from LRSR are more than 98% on most datasets except for D1_noise015 and D1_noise020, whose accuracies are not lower than 90%. However, the HASS and the SC own worse results with 83.92% and even 34.68%, respectively. Further, since SC takes no attention in the sparse representation, it results in an overfitting problem [6]. Although HASS has performed well by employing nonlinear strategy on two datasets, its processing time is the longest in that the computational complexity of the WPD precessing is much high and HASS has to perform WPD on every spike. By using the IALM and optimization method, LRSR conduced the least average time of all three methods. In summary, LRSR is an efficient and considerable approach for spike sorting.

4.4. Effects from Overlapping and Noise Interferences

To further demonstrate clearly the robustness of LRSR in overlapping and noise disruption spike sorting, the misclassification spike number over all datasets and error rates at different noise levels are studied. For comparison, HASS based method is engaged while SC is omitted for the unsatisfactory performance shown in Table 2.

As demonstrated in Figure 2, the total misclassification spikes number obtained by LRSR is less than HASS for most datasets, except D2_noise005. Besides, LRSR outperforms HASS for the overlapping spike sorting. Particularly, LRSR correctly classifies all non-overlapped spikes in five datasets while HASS only succeeds in dataset D2_noise005. In Figure 3, we see that overall the error rates are lower than 0.1 for LRSR while for HASS it exceeds 0.16. As the noise becomes severe, HASS and LRSR show an identical increasing trend in error rate, but LRSR gets a more relatively stable outcome. Thus, it is verified that LRSR gets a more robust performance and has more satisfactory anti-noise characteristics.

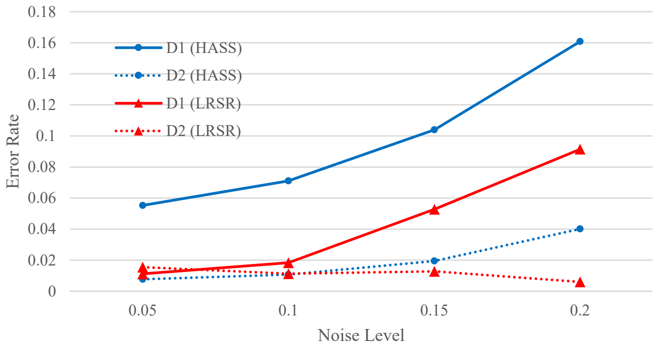

Fig. 3: Error rates of HASS and LRSR.

5. CONCLUSION

In this paper, low-rank and sparse strategies for spike sorting have been proposed. By employing the low-rank strategy to capture the global structure underlying spikes, LRSR recognized spikes well. Further, the sparse representation was integrated into the LRSR to address the noise corruption issue. Thus, the LRSR considerably improved the spike sorting performance, processing the low rank matrix and the sparse matrix with the IALM algorithm. Numerical simulations demonstrated that the LRSR method yielded an outstanding performance compared with other state of the art methods.

6. REFERENCES

[1] Fernando J Chaure, Hernan G Rey, and Rodrigo Quian Quiroga, "A novel and fully automatic spike-sorting implementation with variable number of features," *Journal of neurophysiology*, vol. 120, no. 4, pp. 1859–1871, 2018.

[2] Hernan Gonzalo Rey, Carlos Pedreira, and Rodrigo Quian Quiroga, "Past, present and future of spike sorting techniques," *Brain research bulletin*, vol. 119, pp. 106–117, 2015.

[3] Chaitanya Ekanadham, Daniel Tranchina, and Eero P Simoncelli, "A unified framework and method for automatic neural spike identification," *Journal of neuroscience methods*, vol. 222, pp. 47–55, 2014.

[4] Yun-Yu Chen, Yi-Min Tsai, and Liang-Gee Chen, "Algorithm and implementation of multi-channel spike sorting using gpu in a home-care surveillance system," in *2011 IEEE International Conference on Multimedia and Expo*. IEEE, 2011, pp. 1–6.

[5] Libo Huang, Bingo Wing-Kuen Ling, Ruichu Cai, Yan Zeng, Jiong He, and Yao Chen, "Wmsorting: Wavelet packets decomposition and mutual information based spike sorting method," *IEEE Transactions on NanoBioscience*, 2019.

[6] Haifeng Wu, Kai Yang, and Yu Zeng, "Sparse coding and compressive sensing for overlapping neural spike sorting," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, no. 8, pp. 1516–1525, 2018.

[7] Cyrille Rossant, Shabnam N Kadir, Dan FM Goodman, John Schulman, Maximilian LD Hunter, Aman B Saleem, Andres Grosmark, Mariano Belluscio, George H Denfield, Alexander S Ecker, et al., "Spike sorting for large, dense electrode arrays," *Nature neuroscience*, vol. 19, no. 4, pp. 634, 2016.

[8] Guangcan Liu, Zhouchen Lin, Shuicheng Yan, Ju Sun, Yong Yu, and Yi Ma, "Robust recovery of subspace structures by low-rank representation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 1, pp. 171–184, 2012.

[9] Zhouchen Lin and Hongyang Zhang, *Low-rank models in visual analysis: Theories, algorithms, and applications*, Academic Press, 2017.

[10] Emmanuel J Candès and Benjamin Recht, "Exact matrix completion via convex optimization," *Foundations of Computational mathematics*, vol. 9, no. 6, pp. 717, 2009.

[11] Jian-Feng Cai, Emmanuel J Candès, and Zuowei Shen, "A singular value thresholding algorithm for matrix completion," *SIAM Journal on optimization*, vol. 20, no. 4, pp. 1956–1982, 2010.

[12] Andrew Y Ng, Michael I Jordan, and Yair Weiss, "On spectral clustering: Analysis and an algorithm," in *Advances in neural information processing systems*, 2002, pp. 849–856.

[13] Ehsan Elhamifar and Rene Vidal, "Sparse subspace clustering: Algorithm, theory, and applications," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 11, pp. 2765–2781, 2013.

[14] Zhouchen Lin, Minming Chen, Leqin Wu, and Yi Ma, "The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices," *Coordinated Science Laboratory Report no. UILU-ENG-09-2215, DC-247*, 2009.

[15] R Quian Quiroga, Zoltan Nadasdy, and Yoram Ben-Shaul, "Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering," *Neural computation*, vol. 16, no. 8, pp. 1661–1687, 2004.

[16] Yao Chen, Libo Huang, Jiong He, Kunyao Zhao, Ruichu Cai, and Zhifeng Hao, "Hass: High accuracy spike sorting with wavelet package decomposition and mutual information," in *2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE, 2018, pp. 831–838.

[17] Jiri Wild, Zoltan Prekopcsak, Tomas Sieger, Daniel Novak, and Robert Jech, "Performance comparison of extracellular spike sorting algorithms for single-channel recordings," *Journal of neuroscience methods*, vol. 203, no. 2, pp. 369–376, 2012.