

Automated Platform for Microseismic Signal Analysis: Denoising, Detection and Classification in slope stability studies

Jiangfeng Li, Lina Stankovic, *Senior Member*, Stella Pytharouli and Vladimir Stankovic, *Senior Member*

Abstract—Microseismic monitoring has been increasingly used in the past two decades to illuminate (sub)surface processes such as landslides, due to its ability to record small seismic waves generated by soil movement and/or brittle behaviour of rock. Understanding the evolution of landslide processes is of paramount importance in predicting or even avoiding an imminent failure. Microseismic monitoring recordings are often continuous, noisy and consist of signals emitted by various sources. Manually detecting and distinguishing the signals emitted by an unstable slope is challenging. Research on automated end-to-end denoising, detection, and classification of microseismic events, as an early warning system, is still in its infancy. To this effect, our work is focused on jointly evaluating and developing suitable approaches for signal denoising, accurate event detection, non site-specific feature construction, feature selection and event classification. We propose an automated end-to-end system that can process big data sets of continuous seismic recordings fast and demonstrate applicability and robustness to a wide range of events (distant and local earthquakes, slidequakes, anthropogenic noise etc.). Algorithmic contributions lie in novel signal processing and analysis methods with fewer tunable parameters than the state of the art, evaluated on two field datasets and benchmarked against the state of the art.

Index Terms—Microseismic, Graph Based Bilateral Filter, Neyman-Pearson lemma and Graph Laplacian Regularization Classification

I. INTRODUCTION

SLOPE instability is caused by natural sources, e.g., earthquakes, erosion in caves and seepage through soils, and man-made sources, e.g., construction on cuttings or embankments, temporary excavation and mining activities. At the onset of a major slope failure, minor disturbances (evidenced in the form of microseismic events) occur with small magnitude. Microseismic events correspond to brittle failure mainly attributed to the reduction in effective stress, usually have a negative moment (typically between -3 and 0) and are always non-stationary signals [1]. Besides microseismic events, we also detect and classify landslide induced, micro-earthquakes (also referred to as tremors), i.e., earthquakes with Richter Magnitude < 3 [2].

The highly heterogeneous nature of soil makes the derivation of universal parameters for the characterization of such signals very site-specific. Thus, while microseismic signal

monitoring and analysis is critical to effectively predict major geological disturbances in order to take timely action and minimize fatalities and infrastructure damage, it also remains a challenging problem.

With microseismic monitoring for slope stability, seismometers are deployed outside the unstable soil mass, even at several kilometers distance, to offer continuous monitoring over a large area and time scales. This results in extremely big data sets making the analysis a challenging task. Furthermore, recordings include signals generated by other, non landslide related sources (wind, rainfall, human activities, animals, etc.) as well as unstable ground, and robust methods are needed to extract the signals that represent slope instability, from the rest. Therefore, an end-to-end decision support system would comprise signal denoising followed by event detection, feature construction and selection, and classification. Localization, i.e., knowing the hypocentre of the event, is also important, but this is outside the scope of this paper.

A review of the state of the art, as reported in Section II, indicates that detection and classification of seismic events related to (sub)surface processes focus mainly on volcanic activity, earthquakes, avalanches and mining blasting as such signals have distinct temporal and spectral features. Furthermore, so far, automated event detection and classification methods for seismic signals shows that existing methods have limited accuracy [3]. Recent years have seen increased activity in tackling the more challenging microseismic signal analysis problem, but generally tackling signal denoising, event detection, and more recently multi-class event classification, in isolation. While signal detection is often used interchangeably with arrival picking in the literature, we note that we do not focus on arrival picking but rather on seismic signal detection, where precise timing of the seismic event is not as critical. Our aim is to detecting as many seismic event waveforms as possible, then classify them into multiple event types.

Building on our earlier conference abstract [4] where an end-to-end system, including bandpass filtering (BPF), Short Term Average over Long Term Average (STA/LTA) and Newton-Pearson-based thresholding (NP detector) detection followed by feature selection and classification of 4 event types, via Support Vector Machine (SVM) and Random Forest (RF) are proposed, we highlight and attempt to address the gap in the literature of understanding how denoising, detection, feature construction and selection, and classification approaches jointly influence one another and how they can be tuned to the recordings on real sites in order to pro-

J. Li, L. Stankovic and V. Stankovic are with Dept. Electronic and Electrical Engineering, University of Strathclyde, Glasgow G1 1XW UK e-mail: {jiangfeng.li, lina.stankovic, vladimir.stankovic}@strath.ac.uk

S. Pytharouli is with Dept. Civil and Building Engineering, University of Strathclyde, Glasgow G1 1XJ UK e-mail: stella.pytharouli@strath.ac.uk

vide an accurate automated end-to-end denoising, detection, feature construction and selection, and classification system that operates on big sensor data in near real time. Targeting realistic scenarios where a limited number of seismic sensors are deployed, and thus where array signal analysis cannot be performed, we propose an autonomous single-channel recorded microseismic signal analysis system which includes denoising, detection, feature construction and selection, and classification. Starting from raw measurements, the proposed system can accurately detect and classify three main groups of seismic events including earthquakes of various magnitudes and duration and at various distances from the seismometers. Our contributions are as follows:

- Characterization of the noise distribution in two field datasets for the purpose of evaluating event denoising and detection, and their dependencies
- An automated end-to-end system, comprising event denoising, detection, feature construction and selection, and classification, to efficiently and accurately process and classify large datasets of raw seismic data into specific events
- System design robust enough that is able to distinguish a much larger range of events than current literature, namely earthquakes, slidequakes and tremors, and their sub-classes
- Propose and evaluate a landslide induced seismic-signal denoising approach, tuned to dataset distribution, via an adapted Graph Signal Processing-based bilateral filtering
- Detection of events via a low-complexity thresholding method based on Neyman-Pearson Lemma, designed to detect all true events and minimize false event detection
- Determining which features best characterize each type of microseismic event
- Evaluation of which subset of the hundreds of features commonly used in the seismic literature are pertinent to the 9 classes of landslide induced seismic events, in two field datasets, via feature selection
- Classification of events via an adapted Graph Laplacian Regularization (GLR) classifier, which performs well when the training dataset is small
- Evaluation on two different annotated datasets of real events (not synthetic or generated under lab conditions) to jointly understand denoising, detection and classification, benchmarked against state-of-the-art methods for denoising, detection and classification.

This paper is organized as follows. Section II presents a literature review focused on seismic event denoising, detection, feature construction and selection, and classification. Section III describes the proposed methodology. The results and discussion of results are presented in Section IV. The final section concludes the paper.

II. BACKGROUND

A. Denoising

(Sub)surface seismic measurements are affected by: 1) ambient noise, 2) wave propagation related noise (inc. surface

wave and geologic noise), 3) sensor measurement noise and 4) data processing artifacts [5].

A recent approach to seismic signal denoising is stacking multi-channel recorded data or characteristic functions, which is optimal only when noise components are of similar magnitude, normally distributed, stationary, and uncorrelated across all traces [6]. Thus stacking technique is not applicable for single-station or single-channel recordings. Next, we briefly review denoising approaches, based on signal transformation, decomposition and learning.

Signal denoising is commonly performed in a transform domain, e.g., wavelet domain, which has been shown in [7], [8] to effectively remove long duration, non-stationary ambient or instrumental noise and Gaussian noise. A synchrosqueezing continuous wavelet transform is developed for seismic denoising based on general cross-validation thresholding, with poor performance when signal-to-noise ratio (SNR) < 2 dB due to significant signal waveform distortion [9]. [10] proposed an improved denoising approach via block thresholding which is an adaptive, non-diagonal estimator, where signals are grouped into seismic or noise blocks and the latter removed. However, this approach is unsuitable for real-time denoising due to the complexity of multi (inverse) signal transform. [11] proposed a faster denoising approach with recursive averaging noise estimation; the threshold and block size obtained by minimizing Stein's unbiased risk estimate. However, effectiveness is limited by the noise distribution assumption, which affected by the window length [12]. Besides wavelets, other domains such as curvelets, dreamlets and shearlets have been proposed for multichannel seismic signal denoising as reviewed in [13].

Combining Ensemble Empirical Mode Decomposition (EEMD) and f-x-domain EEMD thresholding, [14] shows good performance for microseismic and seismic signal denoising. However, the parameters of this thresholding method again need to be set manually, which is time-consuming [15]. Similar to EEMD, [16] discuss two denoising approaches based on morphological decomposition targeting low-frequency noise (ground rolls, low-frequency ambient noise, swell and cable strum noise), characteristic of hydraulic fracturing but not so much for landslide induced microseismic monitoring. However, their applicability for non-stationary signals with non-uniform noise distributions as in naturally occurring landslide induced microseismic signals is uncertain.

Learning-based denoising approaches are emerging in the form of deep neural networks to remove background random noise from earthquake signals [17], [9]. However, these approaches require a large amount of labeled training data in the form of recordings or seismograms.

The main drawbacks of the current denoising approaches are processing and parameter tuning complexity, assumption of a particular noise distribution and insufficient evidence of suitability for field data with microseismic and tremor events with insufficient examples to train the models.

Bilateral filtering with Gaussian kernels has outperformed wavelet transform-based denoising, dictionary learning-based denoising and standard Bilateral filter with noise thresholding for seismic image denoising [18]. Motivated by this and its effectiveness in improving event detection for time-series

non-intrusive load monitoring [19], we adopt and adapt a *graph-based bilateral filter* (GraphBF) with Gaussian kernel [20], which has never been tested for denoising of time-series raw seismic signal recordings. To benchmark against GraphBF, we also implement wavelet-based denoising with block thresholding [10], Short-time Fourier transform (STFT)-based denoising [11], wavelet transform denoising with Matlab software package (Wavelet Signal Denoiser), and EEMD [14].

B. Detection

Energy-based variants of STA/LTA detection based on time series signal mean and variance are the most widely used approaches and can detect a seismic signal with small amplitude. However, window length and threshold selection processes of the STA/LTA detection can be time-consuming and inefficient [21]. Moreover, this method cannot detect events with low spatial and temporal separation [22]. Matched filtering [23] based on template matching requires prior knowledge of representative parent waveforms [24] but are proving highly effective.

Learning-based detection approaches are also emerging. In [25], binary SVM is used to classify events from non-events, but use 191 constructed features to train the classifier for detection of events. Deep learning based seismic detection approaches have emerged for earthquake detection [26], [22], but they have only shown effectiveness for well defined earthquake signals and require a large amount of training data, especially for highly variable landslide induced microseismic and tremor signals.

Recently, a review by [3] highlights that there is no established methodology to detect volcano-seismic events in continuous recordings due to the following challenges: 1) numerous events occurring in a short period of time and signals of different types and amplitudes overlapping in time, 2) emergent signals whose amplitude increases and decreases very slowly, 3) reduced ability to clearly detect start and end points, especially when the analysis is carried out on relatively short temporal windows, 4) high variability in an event's duration, 5) methods such as STA/LTA need to be manually tuned at each application (setting thresholds, window lengths, etc), 6) limited testing on relatively small data sets (a few hundred or fewer than 100 samples in some cases) or on data sets including only a given class of signals. While these findings are focused on volcano-seismic signal detection, they are also very relevant to other microseismic signals, whose low SNR further amplifies the above challenges.

We propose event detection via a very low complexity thresholding method, based on the Neyman-Pearson Lemma, never before demonstrated for seismic event detection, to satisfy the following conditions: no manual tuning, detection of the full range of earthquakes, slidequake and tremor microseismic signals and minimizing false event detection, that will in turn improve accuracy of the following feature selection step. The proposed NP detector only requires limited noise signal information as prior knowledge, and is not limited by fine parameter tuning as in STA/LTA based approaches, a good quality template as in matched filtering approaches and a

large amount of training waveforms as in deep-learning-based approaches.

C. Feature Construction and Selection

Accurate signal representation and feature construction are essential for successful classification. Previous research on endogenous landslide seismic sources provides abundant information of signal patterns both in time and frequency domain, including slopequake, rockfall and granular flow [27]. *Earthquake events* have clearly distinguishable P- and S-wave arrival times, including local and distant earthquakes. The dominant frequency of a local earthquake is around 1–20 Hz, while for distant earthquakes, can be under 5 Hz [28]. *Slidequake events* are characterized by small duration, (around 2 s and up to 10 s for microearthquake) and low dominant frequencies of 5–50 Hz, including distant and nearfield slidequake and nearfield microearthquake. In addition, the microearthquake events usually have clearly distinguished P and S phases [28] [29]. *Tremors* are events with duration as long as several minutes and can be confused with environmental noise caused by vehicles or other moving objects [30], because the frequency content varies from very low value to 100 Hz. Tremor-like events include ETS-like signal (Episodic tremor and slip), rockfalls, harmonic tremors, dispersive and external sources of microseismic noise and tremor-like radiations [28]. The electrocardiogram and speech signals have some similar characteristics to tremor signals, e.g., entropy, as a feature to measure the signal's spectral distribution, indicating that it is a strong feature to distinguish an audio signal from the ambient noise-like signal. Other features include brightness to measure the frequency centroid, and silence ratio to measure the ratio of low energy time window over the entire signal duration [31]. In the recent review paper of [28] for landslide induced events, a comprehensive feature set was presented: temporal features, describing the signal waveform in the time domain; spectral features to capture the signal spectrum information; and cepstral features, popular in speech recognition, to highlight a signal's harmonic properties. Additionally, signal skewness and kurtosis combine temporal, spectral, and cepstrum features [3]. Table I provides a list of the 99 temporal, spectral and cepstrum features from the literature that we consider appropriate for characterising our dataset. Additionally, unlike [3], which uses a full feature set for classification, we perform an additional step, namely, feature selection in order to reduce the number of features that uniquely characterize each event type, to reduce the complexity of the classification task. We leverage on standard machine learning feature selection methods- see Table II - to search for the best distinguishing features of our event types. Note that microseismic feature engineering, comprising feature construction, extraction and selection, is a study of its own, as discussed in [32], where a joint iterative graph-based feature engineering and classification approach is proposed, and not the focus of this paper.

D. Classification

Manual classification is still widely used due to the absence of annotated datasets, e.g., (micro-)earthquake, block fall, rock

TABLE I: List of features drawn from the literature: temporal $s(t)$, power signal $p(t)$, envelope $e(t)$, auto correlation function $ac(t)$, spectral $f(v)$, cepstrum domain $ce(v)$ and envelope $ss(t)$, temporal $es(t)$ and spectral $fs(v)$ with (1–5 Hz), (5–9 Hz), (9–13 Hz), (13–17 Hz) and (17–20 Hz) passband. **PMF** refers to Probability Mass Function.

Parameter	Description	Equation	Ref.
Temporal Feature			
T1	Duration	_____	[29][33][34][35]
T2-4	STD, Mean, Median of $s(t)$	_____	[34][28][35] [36] [25]
T5-7	Max, Mean, Median of $e(t)$	_____	[33][28][36]
T8	Rising duration $s(t)$	_____	[35]
T9	Decreasing duration $s(t)$	_____	[35]
T10	Entropy of $s(t)$	$-\sum \mathbf{PMF}(s(t)) \log \mathbf{PMF}(s(t))$	[3][31][37][25]
T11	Zero Cross Rate of $s(t)$	_____	[28] [25]
T12	STD at the decreasing part of $s(t)$	$\sqrt{\frac{1}{(N-t_{max}-1)} \sum_{t_i=t_{max}}^N (s(t_i) - \mu)^2}$	[33]
T13-15	Skewness of $s(t)$, $p(t)$ & $e(t)$	$\frac{1}{N} \sum_{i=1}^N \left(\frac{x(t_i) - \mu}{\sigma}\right)^3$	[29][33][28][35] [25]
T16-18	Kurtosis of $s(t)$, $p(t)$ & $e(t)$	$\frac{1}{N} \sum_{i=1}^N \left(\frac{x(t_i) - \mu}{\sigma}\right)^4$	[29][33][28][35] [25]
T19&20	Rate of attack of $s(t)$ & $e(t)$	_____	[3]
T21&22	Rate of decay of $s(t)$ & $e(t)$	_____	[3]
T23	Ratio of Max and mean of $e(t)$	$\frac{T_5}{T_6}$	[35]
T24	Ratio of Max and Median of $e(t)$	$\frac{T_5}{T_7}$	[35]
T25	Ratio of Max and STD of $e(t)$	$\frac{T_5}{STD(e(t))}$	[35]
T26	Ratio of (tmax)/(N-tmax)	$\frac{t_{max}}{N-t_{max}}$	[33][27][35]
T27	Ratio of T18 and T20	$\frac{T_{18}}{T_{20}}$	[33]
T28-T32	Energy of $es(t)$	$\sum_{i=1}^N es(t_i)^2$	[33][35]
T33-T37	Average power of $ss(t)$	$\frac{1}{N} \sum_{i=1}^N ss(t_i)^2$	[28]
T38&39	Energy of (1:N/3) and (N/3:N) of $ac(t)$	$\sum_{i=1}^{N/3} ac(t_i)^2$ $\sum_{i=N/3}^N ac(t_i)^2$	[33]
T40	Int-ratio of $ac(t)$	$\frac{T_{38}}{T_{39}}$	[33][35]
T41	Num of peaks of $ac(t)$	_____	[33][35]
T42	duration of $ac(t)$	$\max_t(ac(t) < 0.2 \max(ac(t)))/(T_1)$	[27]
T43	Measure of location	$\sum_{i=1}^N ip(t_i)$	[37]
T44	Measure of dispersion	$\sqrt{\sum_{i=1}^N (i - T_{43})^2 p(t_i)}$	[37]
T45	Measure of asymmetry	$\frac{1}{T_{44}^3} \sum_{i=1}^N (i - T_{43})^3 p(t_i)$	[37]
T46	Measure of concentration around single value	$\frac{1}{T_{44}^4} \sum_{i=1}^N (i - T_{43})^4 p(t_i)$	[37]
Spectral Features			
F1	Absolute value of mean $f(v)$	_____	[29][33][35]
F2	Absolute value of max $f(v)$	_____	[33][35] [36]
F3	Absolute value of median $f(v)$	_____	[33][35]
F4	Variance of $f(v)$	_____	[33][35][28] [36] [25]
F5	Max envelop of $f(v)$	$\max(e(f(v)))$	[27] [36]
F6	Num of peaks >0.75 bandwidth $f(v)$	_____	[33][35]
F7	Dominate frequency	_____	[33][29] [36]
F8	Spectral centroid	_____	[33][35] [37] [36] [25]
F9	Int-ratio of $f(v)$	$\frac{\sum_{i=1}^{N/3} f(v_i)^2}{\sum_{i=N/3}^N f(v_i)^2}$	[31]
F10-14	Kurtosis $fs(v)$	$\frac{1}{N} \sum_{i=1}^N \left(\frac{fs(v_i) - \mu}{\sigma}\right)^4$	[33][35]
F15	Num of peaks $f(v)$	_____	[33]
F16-20	Energy of $fs(v)$	$\sum_{i=1}^N fs(v_i)^2$	[36] [35]
F21	Gamma 1	$\frac{\sum_{i=1}^N v_i f(v_i)^2}{\sum_{i=1}^N f(v_i)^2}$	[35]
F22	Gamma 2	$\sqrt{\frac{\sum_{i=1}^N v_i^2 f(v_i)^2}{\sum_{i=1}^N f(v_i)^2}}$	[35]
F23	Gamma 3	$\sqrt{ F_{21}^2 - F_{22}^2 }$	[35]
F24	Mean frequency	$\frac{\sum_{i=1}^N \mathbf{PSD}(v_i) v_i}{\sum_{i=1}^N \mathbf{PSD}(v_i)}$ PSD is the power spectral density of $f(v)$	[27]
F25	Frequency bandwidth	$2 \sqrt{\frac{\sum_{i=1}^N \mathbf{PSD}(v_i) v_i^2}{\sum_{i=1}^N \mathbf{PSD}(v_i)} - F_{24}^2}$	[27]
F26	Minimal frequency	$\min_v(\mathbf{PSD}(v) < 0.2 \max(\mathbf{PSD}(v)))$	[27]
F27	Maximal frequency	$\max_v(\mathbf{PSD}(v) < 0.2 \max(\mathbf{PSD}(v)))$	[27]
F28	Gyration radius	$\sqrt{\frac{m_3}{m_2}} m_i$ is the i th moment	[33]
F29	Spectral centroid width	$\sqrt{F_8^2 - F_{28}^2}$	[33] [25]

continued on the next page

TABLE I: Literature reviewed features – continued from previous page

Parameter	Description	Equation	Ref.
Cepstrum Features			
C1	STD $\mathbf{ce}(v)$	$\frac{std(\mathbf{ce}(v))}{\sigma}$	[3]
C2	Skewness of $\mathbf{ce}(v)$	$\frac{1}{N} \sum_{i=1}^N \left(\frac{\mathbf{ce}(v_i) - \mu}{\sigma} \right)^3$	[3]
C3	Kurtosis of $\mathbf{ce}(v)$	$\frac{1}{N} \sum_{i=1}^N \left(\frac{\mathbf{ce}(v_i) - \mu}{\sigma} \right)^4$	[3]
C4	Max value of $\mathbf{ce}(v)$	$\max(\mathbf{ce}(v))$	[3]
C5-14	First 10 $\mathbf{ce}(v)$	—	[3] [25]
Acoustics Features(Linear prediction filter coefficients)			
A1-10	10 Linear Prediction Coefficients	—	[27]

TABLE II: Feature Selection methods

Feature Selection Method	
1.	Infinite Latent Feature Selection (ILFS)
2.	Infinite Feature Selection (Inf-FS)
3.	Eigenvector Centrality Feature Selection (ECFS)
4.	Minimum Redundancy Maximum Relevance (mRMR)
5.	Relieff-based feature selection (Relieff)
6.	Mutual Information-based feature selection (mutInfFS)
7.	Feature Selection via Concave Minimization (FSV)
8.	Laplacian Score (LS)
9.	Recursive Feature Elimination (RFE)
10.	Fisher/Correlation Score (Fisher)
11.	Unsupervised Discriminative Feature Selection (UDFS)
12.	Local Learning-Based Clustering (llcfs)
13.	Feature Selection with Adaptive Structure Learning (FSASL)
14.	Dependence Guided Unsupervised Feature Selection (Dgufs)
15.	Unsupervised Feature Selection with Ordinal Locality (UFSOL)
16.	Loss Object - for calculating loss functions (Loss)

fall, quarry blast, and multiple events are manually classified based on frequency range, amplitude, signal shape and duration [38]. SVM is widely used for seismic classification, e.g., classification of Long period (LP) events, Tremor (TR), Volcano Tectonic (VT) events, Tectonic (TC) in [34], automatic classification of earthquake and non-earthquake events in [39] and of volcano-seismic events in [3]. RF is used for classification of landslide seismic events, namely rockfall, quake, earthquake and noise [33]. Deep learning, including Convolution neural network (CNN), is used to classify landslide-based seismic events and human-made events such as mine blasts. The microseismic events and quarry blast events are classified in [40] and [41] respectively, with good accuracy.

SVM can solve the high dimensional non-linear classification problem with small training datasets [42]. RF is parallelizable, performs well for high dimensional signals, quick in prediction/training, robust to outliers and non-linear data, handles unbalanced data and low bias [43]. Lara et al. [44] presented a fair comparison of different classifiers (multilayer perceptron neural network (MLP), linear discriminant analysis (LDA), RF, and SVM) for the volcano-seismic signal, where it was shown that SVM outperformed all other classifiers, LDA was the worst performing classifier, and the performance of MLP and RF were approximately the same. Note that RF has lower complexity than MLP. In [45], to classify long-period and volcano-tectonic signals, five classifiers: Naive Bayes (NB), SVM, K-Nearest Neighbors (KNN), feed-forward back-propagation neural network (FFBP) and RF were implemented with 84 features; RF performed the best, following by NB and KNN classifiers with FFBP neural network slightly worse. The SVM classifier did not perform well because it was

implemented with a linear kernel only. Although Deep Neural Network has significant advantages in many different fields, including seismic event classification, it has not been well developed, most likely due to lack of sufficient annotated data to train the models. This is not the case for volcano-seismic monitoring, with datasets incorporating over 20,000 events. We chose to focus on more explainable and well understood classification methods (deep learning is not one of them), where features are hand-crafted instead of depending on deep-learning based feature selection, in order to understand the unique features that best represent all nine of our event classes, as listed in Table III. The classification literature only addresses a few classes with little detail on understanding the importance of features.

We target classifiers shown to have good performance when the training dataset is small and adopt a GLR-based classifier, benchmarked against SVM with a radial basis function kernel and RF. The main idea of GLR-based classification is to smooth the label signal through graph Laplacian regularization [46], and has not been used for seismic event classification. GLR-based classifiers are semi-supervised, which mainly concentrate on performance improvement with a small amount of labeled samples [47]. This is a promising approach for seismic event classification due to the complexity and uncertainty of seismic signal labeling. GLR classifiers have been applied to many 1D, 2D, and 3D signal classification tasks, including image classification and load disaggregation [48], [49]. Here we adapt the GLR-based classification method described in [49] to microseismic event classification.

III. METHODOLOGY

The proposed automated early warning system comprises denoising, detection, feature construction and selection, and classification, as shown in Fig. 1. Each of these components is described in the following subsections.

A. Event Denoising

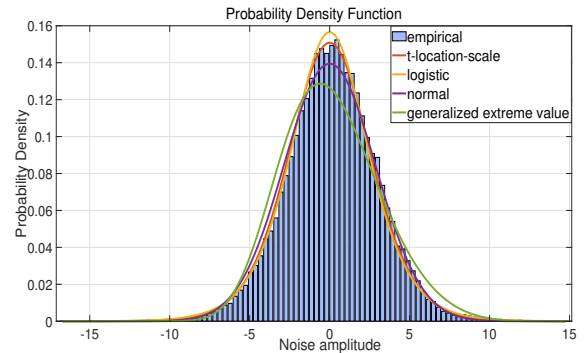


Fig. 2: Probability density function curve fit

1) *Background noise distribution:* While synthetic seismic noise tends to be modeled as Gaussian in the literature, we observed that the noise present in the field dataset (see Section IV-A), after BPF, has a best fit to the t -location-scale distribution, as shown in Fig. 2, which is more prone to outliers or heavy tails than the normal distribution. The t -location-scale distribution is a mixture of normal distributions with gamma

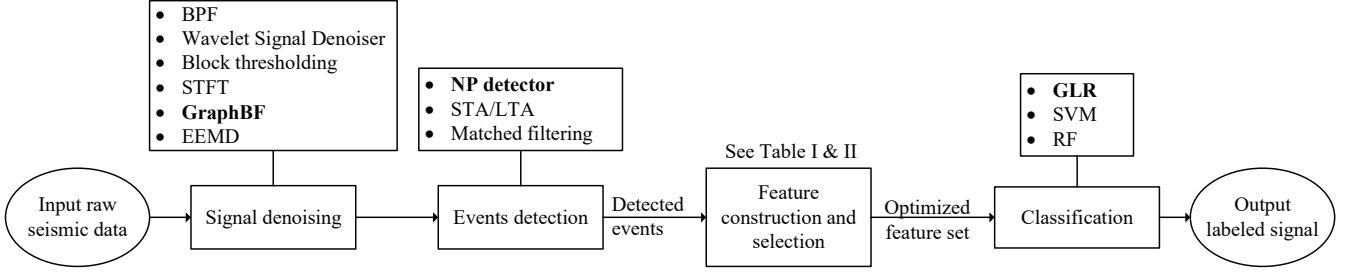


Fig. 1: System Flow Chart.

mixing weights, and therefore the variance follows gamma distribution with parameter $\frac{n}{2}$. The probability density function (PDF) is given by:

$$f_X(x) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{\pi n} \Gamma(\frac{n}{2})} \left(\frac{n}{n+x^2} \right)^{\frac{n+1}{2}} \quad (1)$$

2) *Signal Denoising*: Since the dominant frequencies of microseismic events are relatively small and expected to fall in the (1–20 Hz) range [28], the first step is to bandpass filter the raw signals to remove high-frequency components of coherent noise. We used a Butterworth filter with the aforementioned passing frequency band. Standard BPF does not sufficiently denoise the signal, and therefore additional denoising approaches are needed.

GraphBF denoising: Given a band-pass filtered and sampled signal \mathbf{x} , we design an undirected graph $\mathcal{G} = (\mathcal{V}, \mathbf{A})$, where \mathcal{V} is the set of nodes, each corresponding to one sample in \mathbf{x} , and \mathbf{A} is defined as the weighted adjacency matrix:

$$a_{i,j} = \exp \left\{ -\frac{(x_i - x_j)^2}{2\sigma^2} \right\}, \quad (2)$$

where σ is known as kernel bandwidth. Note that $a_{i,j}$ reflects correlation between samples x_i and x_j of \mathbf{x} . Let \mathbf{D} be a diagonal matrix, given by $d_{i,i} = \sum_j a_{i,j}$. Then the output of GraphBF, for a noisy input \mathbf{x} is given by (3) [19].

$$\arg \min_{\mathbf{s}} \frac{1}{2} \|\mathbf{s} - \mathbf{x}\|_2^2 + \alpha \frac{1}{2} \|\mathbf{s} - \mathbf{D}^{-1} \mathbf{A} \mathbf{s}\|_2^2 \quad (3)$$

Note that the first term of (3) maintains the similarity between the denoised signal \mathbf{s} and the input signal \mathbf{x} , the second term is the smoothness prior that tends to minimize the difference between the restored signal and its neighbours, and α is the trade-off factor. This optimization problem has a closed-form solution [50], [19]:

$$\tilde{\mathbf{s}}_{BF} = (\mathbf{I} + \alpha(\mathbf{I} - \mathbf{D}^{-1} \mathbf{A})^*(\mathbf{I} - \mathbf{D}^{-1} \mathbf{A}))^{-1} \mathbf{x}, \quad (4)$$

where \mathbf{I} is the identity matrix. $\tilde{\mathbf{s}}_{BF}$ is the denoised signal that is passed to the event detection block.

The GraphBF closed-form solution involves two parameters: the trade-off factor α , set heuristically to 300, and kernel bandwidth σ , optimized based on Neyman-Pearson (NP) Hypothesis Testing [51], discussed in Section III-B.

B. Event Detection

Given two hypotheses $H_0 : f_0(x)$ and $H_1 : f_1(x)$, the NP lemma decides on the valid hypothesis based on probabilities of detection P_D and false alarm P_{FA} defined as:

$$P_D = \int_{\mathcal{X}} f_1(x) dx, \quad P_{FA} = \int_{\mathcal{X}} f_0(x) dx. \quad (5)$$

Specifically, we wish to maximize P_D while keeping P_{FA} within a small boundary θ , heuristically set to 0.01. The objective function can then be defined as:

$$\max_{\mathcal{X}} (P_D - g(P_{FA} - \theta)), \quad (6)$$

where g is the Lagrange multiplier, and

$$P_{FA} = \int_{\lambda}^{+\infty} p_r(r|H_0) dr, \quad (7)$$

where $r \sim \mathcal{N}(0, M\beta^2)$ follows the noise distribution shown in Fig. 2 with M and β , representing, respectively, the length of the window and standard deviation of background noise.

Given the PDF of the noise (1), we calculate $p_r(r|H_0)$ and P_{FA} as:

$$p_r(r|H_0) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{\pi n M \beta^2} \Gamma(\frac{n}{2})} \left(\frac{n M \beta^2}{n M \beta^2 + r^2} \right)^{\frac{n+1}{2}}, \quad (8)$$

$$P_{FA} = \int_{\lambda}^{+\infty} \frac{\Gamma(\frac{n+1}{2})}{\sqrt{\pi n M \beta^2} \Gamma(\frac{n}{2})} \left(\frac{n M \beta^2}{n M \beta^2 + r^2} \right)^{\frac{n+1}{2}} dr \quad (9)$$

$$\lambda = \sqrt{M \beta^2} \text{tinu}(1 - P_{FA}, n) \quad (10)$$

where $\text{tinu}()$ represents the t inverse cumulative distribution function, and n is the degree of freedom of the t -location-scale distribution, obtained after the PDF curve fitting.

Note that the detector output is the linear detection threshold λ . All signal amplitudes larger than λ are considered detected events. To exclude detection errors caused by sensor noise or electrical failure, we discard all detected events shorter than 5 samples. This duration was chosen to include short-duration (≤ 2 s) quake-like signals, to a wide variety of longer-duration tremor-like radiations (≥ 2 s to several min). In addition, if the time difference between two consecutive detected events is shorter than 0.5 s, then these events are merged into a single event. This step is especially useful for detecting Class 3.1, i.e., multiple events (Rock fall/ near repeater), correctly.

To design benchmarks, we tune the following adjustable STA/LTA parameters as per [52]: 1) short term window length as 1s and 0.5s, 2) long term window length as 50s, 3) STA/LTA trigger threshold level set to 2, and 4) STA/LTA de-trigger threshold level set to 0.8. For the matched filtering, the trigger threshold is set as five times the standard deviation value of the cross-correlation, as suggested in paper [23]. With simple thresholding and cross-correlation, the matched filtering is relatively fast.

C. Feature Construction, Weighting, and Selection

Since the wave amplitude is highly affected by wave propagation attenuation and we do not have information about seismic source locations, we normalized the filtered signal before features are constructed.

We use MATLAB Toolbox: Feature Selection Library (FS-Lib 2018) [53] to select the best subset of features from Table I to characterize each class. Note that effective feature selection, besides decreasing complexity, can improve accuracy by avoiding overfitting [54].

The subset of features which has the best classification result using 5-fold cross-validation was selected. The feature selection method also returns a soft value score for each feature k , indicating the usefulness of the feature, which, after normalization, is used in the proposed graph Laplacian regularization, as c_k , described next.

D. Event Classification

We break down the classification task as multiple binary classification tasks using a one-against-all approach, where we consider one class at the time, and classify samples as belonging to the class (label +1) or not (label -1). This is the preferred approach compared to a multi-class classification approach because it enables us to distinguish more classes accurately.

For the GLR-based classifier, using the training annotated dataset, we design an undirected graph, $\mathcal{G} = (\mathcal{V}, \mathbf{A})$, where \mathcal{V} is the set of nodes and \mathbf{A} is defined as the weighted adjacency matrix, with a_{ij} being the weight of the edge connecting Nodes i and j . Each event identified by the event detection approach is assigned to a node in \mathcal{V} . Let N be the total number of events/nodes. The weight a_{ij} reflects the level of correlation between the Nodes i and j (that is, between the features constructed in Event i and j). Following [55], we use the Gaussian kernel function and set:

$$a_{i,j} = \exp\left\{-\sum_{k=1}^K c_k \frac{(f_k(i) - f_k(j))^2}{2\sigma^2}\right\}, \quad (11)$$

where K is the number of features after the feature selection, $f_k(i)$ is the value of the k -th feature of Event i , σ is the kernel bandwidth and c_k is the weight of k -th feature obtained from the soft score of the feature selection methods as mentioned before. Then, the combinatorial Laplacian matrix of the graph can be calculated as $\mathbf{L} = \mathbf{D} - \mathbf{A}$, where \mathbf{D} is a diagonal matrix, defined in Section III-A.

Let n be the number of events in the training dataset. Then, we define the graph signal for Event i as:

$$s_i = \begin{cases} +1, & \text{if Event } i \text{ belongs to the Class and } i \leq n \\ -1, & \text{otherwise and } i \leq n \\ 0, & \text{for } n < i \leq N. \end{cases} \quad (12)$$

Signal classification using graph Laplacian regularization is based on the assumption that the nodes that belong to the sample class are connected using high weighted edges; hence, if we look at the classification labels as a graph signal, this signal should change smoothly across the graph. By finding the smoothest signal given constraints in (12), we can estimate the unknown class labels (initially set to 0). That is, we minimize the quadratic form of the global smoothness of the graph, often refer to as, graph Laplacian regularization term [46], which has a closed-form solution:

$$\begin{aligned} \mathbf{s}^* &= \min_{\mathbf{s}} \mathbf{s}^T \mathbf{L} \mathbf{s} \\ &= \mathbf{L}_{n+1:N, n+1:N}^{\#} \left(-\mathbf{s}_{1:n}^T \mathbf{L}_{1:n, n+1:N}\right)^T, \end{aligned} \quad (13)$$

where $\#$ denotes pseudo-inverse matrix.

As in [49], once \mathbf{s}^* is calculated as above, for testing data, it is set to +1 if $s_i^* > 0$, or -1, otherwise.

IV. RESULTS

In this section, first we describe the datasets and used metrics. Then we present denoising and event detection, feature construction and selection, and classification results.

A. Datasets

Two datasets are analyzed - Super-Sauze (SZ10) and Pechgraben (PG) datasets [28]. These two datasets are chosen as they are publicly available, enabling reproducibility of the research, are annotated/labeled and include a variety of different types of microseismic events, including earthquakes, slidequakes, and tremors, as well as their sub-classes.

SZ10 and PG datasets contain recordings from two ongoing mud-based landslides located in the southwestern French Alps and upper Austria, respectively. SZ10 data was collected at a sampling rate of $F_s = 1000$ Hz by DGNSS (Differential Global Navigation Satellite System), which monitored the landslide for 58 days, in 2010. The PG dataset includes 15 days of seismic activities recorded at a sampling rate of $F_s = 250$ Hz over two years (nine days in 2015 and six days in 2016 denoted by PG-15 and PG-16, respectively, in [28]).

Both datasets are pre-processed and manually labeled using expert knowledge as described in [28]. Nine classes are identified in the SZ10 dataset: local earthquake, distant earthquake (teleseismic), distant slidequake, nearfield low-frequency slidequake, nearfield microearthquake, multiple tremors (Rockfall/near repeater), dispersive tremors, undefined sinusoidal events (harmonic), and calibration shots. The PG dataset consists of six classes: local earthquake, distant earthquake (teleseismic), distant slidequake, multiple tremors (Rockfall/near repeater), undefined sinusoidal events (harmonic) and calibration shots. Note that the calibration shots are artificially induced signals used for signal localization. Accurately detecting

and classifying these signals will help assess the performance of our system.

TABLE III: The Events No. in the SZ10 and PG datasets.

Event No.	Class	SZ10	PG
1.1	Local earthquake	50	7
1.2	distant earthquake (Teleseismic)	6	9
2.2	distant slidequake (Moderate distant quake)	39	6
2.3	Nearfield low frequency slidequake	12	0
2.4	Nearfield microearthquake	7	0
3.1	Multiple event (Rock fall/near repeater)	15	5
3.2	Dispersive tremor	7	0
3.4	Undefined sinusoidal event (Harmonic)	15	4
9	Calibration shots	11	5
99	Undefined signal (Environmental noise)	12	6
Total number of events		174	42

Table III shows the number of events per class for the SZ10 and PG datasets. The classes are numbered as in [28].

B. Metric

For the overall end-to-end system testing we use accuracy metrics that are the norm for machine learning classification tasks. These are precision, recall, and F1 score, defined as:

$$\text{Precision} = \frac{TP}{TP + FP}, \quad \text{Recall} = \frac{TP}{TP + FN}, \quad (14)$$

$$\text{F1 score} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}, \quad (15)$$

where True Positive (TP) is the number of correctly detected and classified events or classes; False Positive (FP) is the number of wrongly detected events or classes; and False Negative (FN) is the number of missed events or classes. We consider the labels provided by the dataset authors [28] as ground-truth.

C. Denoising and Event Detection Results

1) *Signal denoising result:* To provide a statistically sound measure of the denoising ability of our denoising approaches, we use signal to noise ratio of [56]:

$$\text{SNR} = \sqrt{\frac{\sum_{i=1}^n (\text{Signal}_i)^2}{\sum_{i=1}^n (\text{Noise}_i)^2}}, \quad (16)$$

where n is the number of samples in the window, Signal_i is measured signal during the event, while Noise_i is the measured signal with noise present only (no event).

Table IV presents the robustness of the proposed GraphBF and state-of-the-art benchmark denoising approaches, where it can be seen that the GraphBF outperforms all other approaches for denoising all types of events. As expected, this is closely followed by the best known wavelet denoising approach. The remaining approaches performed poorly.

2) *Signal detection result:* The performance of the proposed (NP detector) and benchmark (STA/LTA and Matched filtering) detection approaches for the case of BPF only, without additional denoising, and two cases of BPF with the GraphBF and Wavelet Signal Denoiser, are shown in Tables V and VI, for the SZ10 and PG datasets, respectively. As discussed in Section II-B, our aim is to detect all labeled events, without missing any (i.e., a recall of 1) and minimize the number of false positives (the larger the precision the better).

As expected, the improvement due to additional denoising, besides standard BPF, can be observed through the consistent increase in number of events detected (TPs and FPs), and improvement of the recall value for all 3 detection approaches. However, the increase in false alarms (FPs) is mitigated most effectively with the proposed GraphBF denoising approach in conjunction with the NP detector, as evidenced by the improved precision value. In general, the proposed NP detector has the best performance, closely followed by the matched filtering. The STA/LTA provides lower detection accuracy, irrespective of the denoising method because it tends to detect too many FPs.

Since GraphBF denoising, together with the proposed NP detector approach, provides the highest detection accuracy for both datasets, without missing an event, they are therefore used for the next step: classification.

D. Feature Construction and Selection Result

Since different classes or event types are characterized by temporal, spectral and cepstral features, we attempt to narrow down the most pertinent or unique features for each class through feature selection methods of Table II on all 99 feature candidates of Table I, and using GLR classifier with 50% split training and testing data with five-fold cross-validation, resulting in Table VII.

For example, as mentioned in Section IV-A, Class 1.2, distant earthquake signal, has a low, dominant frequency, hence the feature selection method highly ranks feature F3 median value of the Discrete Fourier Transform (DFT) coefficient. Class 2.2, distant slidequake, has a consistent beginning, i.e., characterized by the following 2 features: rate of attack T19 and decay T20. Classes 2.3 and 3.1 have low-frequency content and short duration which leads to the signal shape-based features T21 to be highly ranked. Class 2.4, nearfield microearthquake, is the only signal with distinguishable P and S phase, hence the signal waveform-based features, feature T12, T13, and T14, are highly ranked. Class 9 calibration shot has a unique waveform, and therefore waveform-based features T25 and T38 are picked by the feature selection method.

E. Classification Results

It can be seen from Table III that the classes are very unbalanced, which would have a negative impact on training the classification models. Indeed, some classes, such as distant earthquake, nearfield microearthquake and dispersive tremors, have too small number of events to train the model. To

TABLE IV: Denoising result (SNR dB)

Denoising approach	class 1.1	class 1.2	class 2.2	class 2.3	class 2.4	class 3.1	class 3.2	class 3.4	class 9	average
BPF	2.68	3.24	4.54	5.96	9.57	9.81	7.23	5.02	35.13	9.24
Wavelet Signal Denoiser	12.34	7.79	17.77	23.72	15.93	30.18	12.49	14.54	37.21	19.11
Block thresholding [10]	9.18	9.55	14.48	18.24	22.44	24.13	18.73	13.92	39.94	18.96
STFT [11]	10.04	7.29	26.29	11.47	12.65	14.97	14.15	9.61	35.33	15.76
GraphBF [20]	13.45	12.33	19.13	27.06	24.27	30.76	25.24	22.39	47.59	24.69
EEMD [14]	9.13	6.45	-2.68	8.05	9.34	13.22	8.01	12.87	35.13	11.06

TABLE V: Event detection results for the SZ10 dataset.

Detection methods	TP	FP	Recall	Precision	F1 score
BPF					
NP detector	174	69	1	0.72	0.83
STA/LTA	157	45	0.9	0.78	0.83
Matched filtering [23]	169	91	0.97	0.65	0.78
BPF + GraphBF					
NP detector	174	45	1	0.79	0.89
STA/LTA	165	71	0.95	0.70	0.80
Matched filtering [23]	170	52	0.98	0.77	0.86
BPF + Wavelet Signal Denoiser					
NP detector	170	78	0.98	0.69	0.81
STA/LTA	171	339	0.98	0.34	0.50
Matched filtering [23]	172	84	0.99	0.67	0.80

TABLE VI: Event detection results for the PG dataset.

Detection methods	TP	FP	Recall	Precision	F1 score
BPF					
NP detector	42	10	1	0.81	0.89
STA/LTA	36	35	0.86	0.51	0.64
Matched filtering [23]	40	21	0.95	0.66	0.78
BPF + GraphBF					
NP detector	42	5	1	0.89	0.94
STA/LTA	38	44	0.90	0.46	0.61
Matched filtering [23]	40	7	0.95	0.85	0.90
BPF + Wavelet Signal Denoiser					
NP detector	42	48	1	0.47	0.64
STA/LTA	41	56	0.98	0.42	0.59
Matched filtering [23]	40	34	0.95	0.54	0.69

TABLE VII: Feature selection result. Column 2 shows the best features, numbered as in Table I, for each event type.

class	features selected
1.1	T44,T43,T27,T40,T24,F5,T22,F10,T13,F2 A8,T20,A9,T9,F11,F16,F13,T21,A7,T15 T12,T33,T2,T14,F14,T6,T1,T10,T23
1.2	C3,F3,F15,F6,F21,F24,T1,T9
2.2	F11,C4,T22,C6,T17,T24,T21,T12,T19,T11 T20,T25,T37,C11,A6
2.3	C8,T21,F28,C14,C10,A8,T3,T25,C2,T34
2.4	C8,T14,T13,T12
3.1	T21, F7,F4,F3,T22,C11,C7,C4,A2,T12 F13,T11,F23,F25,T25,A1,T10,A3
3.2	F27,T2,T20,T34,T8,F3,C5,F28,C1,C14 T10,F21,T42,C12,T6,A7,T50,F23,T5,T12 F7,C2,T7,C13,T25,F24,T38
3.4	T1,F29,F8,T9,T8,F28,T35,T41,T12,T51
9	F25,T38,C6,C10,F26,T25

overcome this, we generated synthetic data using Synthetic Minority Over-Sampling Technique (SMOTE) [57], directly in the feature domain. SMOTE preprocessing is a standard tool for mitigating imbalanced data model learning by first confirming the oversampling number N , obtained with 1:1 class distribution, then selecting the instance of the minority class and the K nearest neighbors, and finally with N of the

K instances, new samples are generated by interpolation [58]. The synthetic features are generated based on the features constructed described in Section II-C, and only used for training the classifiers.

Our classification strategy is one-against-all with 5 fold cross-validation. We start with Class 1.1, and split the data into Class 1.1 events and non-Class 1.1 events. To train the model, we use real non-Class 1.1 data with (different ratios from 10% to 70%) and equal amount of synthetic Class 1.1 data. This way, the ‘binary’ classifier will be balanced. For example: after denoising with GraphBF and detection with NP detector, 219 events (174 true events + 45 false alarms) are detected for the SZ10 dataset with 50 Class 1.1 events and 169 non-Class 1.1 events. Thus, 118 (i.e., 70% of 169) detected non-Class 1.1 events are used for training and 118×99 synthetic features are generated for Class 1.1. The remaining $169 - 118 = 30\%$ of Non-Class 1.1 events together with all detected 50 Class 1.1 events are used for testing. Note that for testing, we use the events detected by the proposed detection approach, which also contain false positives.

The classification results, for the SZ10 and PG datasets, are shown in Tables VIII and IX, respectively. Three classification methods are compared: the proposed GLR-based classifier, SVM and RF classifiers, as described in Section II-D and III-D. All three classifiers are fed with the same events detected by the proposed GraphBF+NP-detector. It can be seen that the GLR-based classifier is, for most classes, the most accurate classifier. This is expected since there is a high intra-class variability that GLR handles well. Table VIII shows that only Classes 2.2 and 3.1 have a relative low F1 score for all three classifiers. The poor performance of Class 2.2 is explained by its small duration and false alarms during detection. Class 3.1 belongs to tremor-like signals, which are usually observed at strike-slip faults [28], having similar characteristics with environmental noise [28], and hence more difficult to distinguish. The results for the PG dataset, Table IX, are worse for Classes 1.1, 3.1, and 3.4. This is expected, since the PG dataset is smaller and contains recordings from two different monitoring periods (9 days in 2015 and 6 days in 2016) which are considered together. This is impacting multi-event classes, such as Classes 3.1 and 3.4. Overall, the best classification result is obtained by GLR classifier with balance split (50%) for training and testing for SZ10 dataset, while for PG dataset, the split ratio of 40% with GLR obtained the highest F1 score.

F. Complexity Analysis

Tables X and XI show the execution time in seconds (s) for each approach to process a signal from the SZ10 dataset of 5 min duration. Note that the training in the classification step

TABLE VIII: SZ10 classification result (F1 score)

class	1.1	1.2	2.2	2.3	2.4	3.1	3.2	3.4	9	mean
GLR	10%	0.53	1.00	0.58	0.87	0.39	0.43	0.71	0.62	0.65
	20%	0.68	1.00	0.75	0.80	0.93	0.50	1.00	0.86	0.79
	30%	0.84	1.00	0.82	1.00	0.93	0.58	1.00	0.89	0.88
	40%	0.88	1.00	0.79	1.00	0.93	0.72	0.92	0.97	0.91
	50%	0.85	1.00	0.93	1.00	1.00	1.00	0.82	1.00	0.96
	60%	0.88	1.00	0.80	0.92	0.93	1.00	0.93	0.79	0.91
	70%	0.95	1.00	0.93	1.00	0.93	0.83	1.00	0.93	0.95
SVM	10%	0.56	0.60	0.57	0.67	0.44	0.21	0.60	0.55	0.53
	20%	0.74	1.00	0.75	0.67	0.93	0.75	0.93	0.82	0.81
	30%	0.83	1.00	0.80	0.83	1.00	0.93	0.82	0.71	0.85
	40%	0.87	1.00	0.76	0.89	1.00	0.82	0.88	0.82	0.87
	50%	0.87	1.00	0.90	0.80	0.88	0.83	0.82	0.88	1.00
	60%	0.91	0.92	0.80	0.92	0.93	0.97	0.93	0.81	0.90
	70%	0.93	1.00	0.93	0.96	0.93	0.86	1.00	0.97	0.90
RF	10%	0.70	0.71	0.69	0.32	0.56	0.61	0.48	0.48	0.57
	20%	0.84	0.91	0.72	0.49	0.70	0.60	0.52	0.86	0.68
	30%	0.84	0.80	0.83	0.76	0.78	0.65	0.56	0.72	0.75
	40%	0.92	1.00	0.74	0.69	0.93	0.70	0.64	0.80	0.80
	50%	0.90	1.00	0.80	0.73	0.82	0.93	0.80	0.85	0.83
	60%	0.95	0.75	0.78	0.96	0.93	0.94	0.78	0.90	0.85
	70%	0.95	1.00	0.88	0.96	1.00	0.94	0.86	0.97	0.90

TABLE IX: PG classification result (F1 score)

class	1.1	1.2	2.2	3.1	3.4	9	mean	
GLR	40%	0.88	0.94	1.00	0.91	1.00	0.91	0.94
	50%	0.53	0.88	0.86	1.00	0.86	0.67	0.80
	60%	0.86	0.95	1.00	0.83	0.86	0.83	0.89
	70%	0.93	0.78	1.00	0.83	1.00	1.00	0.92
SVM	40%	0.92	0.88	1.00	0.77	0.86	1.00	0.90
	50%	0.80	0.95	1.00	0.91	0.80	0.83	0.88
	60%	0.80	0.84	1.00	0.83	1.00	0.83	0.89
	70%	0.92	0.84	1.00	0.91	0.89	1.00	0.93
RF	40%	0.88	0.95	0.92	0.67	0.67	1.00	0.85
	50%	0.88	0.84	0.80	0.91	0.55	0.83	0.80
	60%	0.82	0.90	1.00	0.80	0.75	1.00	0.88
	70%	1.00	0.90	1.00	0.80	0.75	1.00	0.91

TABLE X: The execution time for each approach.

SZ10 (5 min) Fs=1000 Hz 300,000 samples	Approach	Time (s)
BPF + Denoising	BPF	0.04
	Matlab Wavelet app	0.11
	Block thresholding	254.65
	STFT	16.93
	EEMD	2.85
Detection	GraphBF	36.62
	NP detector	0.12
	Match filter	0.12
	STA/LTA	2.25

TABLE XI: Feature selection and classification execution time in seconds

class	1.1	50% train (168)	test (135)	Time (s)
work flow		SVM	RF	GLR
feature selection		761.76	9195.8	959.04
No. optimized features		59	76	29
classification train		0.64	1.91	
classification test		0.06	0.13	0.06

includes feature selection, and the synthetic data generation steps before training the models. The execution time was tested based on class 1.1. All experiments were run on a desktop with i7-7700K CPU and 16GB RAM using Matlab 2019b. From Table X, we observe that although our denoising approach GraphBF is slower than benchmarks, the overall processing time of our system is low. Indeed, the system (BPF + GraphBF + NP detector + GLR classifier) needs roughly 37s to process 5 min of data, hence it can be used for real-time processing. Note one-off feature selection and classification training are

carried out offline, and only testing execution time provided.

V. CONCLUSION

We proposed a novel automated end-to-end system that can denoise a raw seismic signal, detect seismic events, then construct and select best features characterizing each event type and finally classify events into specific classes. Having characterized the noise distribution in our field dataset, we adopt graph signal processing based denoising and classification approaches, and a low-complexity Neyman-Pearson based hypothesis testing approach for detection. Our proposed approaches are benchmarked against the state-of-the-art denoising, detection and classification approaches in the seismic literature, showing consistently better performance for each step in terms of improved denoising capability via SNR measure, improved detection capability by ensuring all true events are detected while minimizing false alarms, and improved classification via F1 score. The system is shown to be of low-complexity via run-time results. Despite demonstration on landslide induced events, the methodology can be tuned to suit other data sets. Future work will focus on feature engineering, including feature embedding, to better understand and explain feature importance on classification of specific seismic events.

ACKNOWLEDGMENT

The authors would like to thank Dr Naomi Vouillamoz for her help in acquiring and interpreting the SZ10 and PG datasets. This work was supported in part by EPSRC Prosperity Partnership research and innovation programme under grant agreement EP/S005560/1.

REFERENCES

- [1] S. C. Maxwell and T. I. Urbancic, "The role of passive microseismic monitoring in the instrumented oil field," *Lead. Edge*, vol. 20, no. 6, pp. 636–639, 2001.
- [2] V. I. Grechka and W. M. Heigl, *Microseismic monitoring*. Society of Exploration Geophysicists Tulsa, Oklahoma, USA, 2017.
- [3] M. Malfante, M. Dalla Mura, J.-P. Métaxian, J. I. Mars, O. Macedo, and A. Inza, "Machine learning for volcano-seismic signals: Challenges and perspectives," *IEEE Signal Process Mag*, vol. 35, pp. 20–30, 2018.
- [4] J. Li, L. Stankovic, S. Pytharouli, and V. Stankovic, "High-accuracy real-time microseismic analysis platform: case study based on the super-sauze mud-based landslide," in *Geoconvention*, 2020.
- [5] D. Kumar and I. Ahmed, "Seismic noise," *Encyclopedia of Solid Earth Geophysics*, pp. 1–6, 2020.
- [6] G. Liu, S. Fomel, L. Jin, and X. Chen, "Stacking seismic data using local correlation," *Geophysics*, vol. 74, no. 3, pp. V43–V48, 2009.
- [7] A. Ansari, A. Noorzad, H. Zafarani, and H. Vahidifard, "Correction of highly noisy strong motion records using a modified wavelet de-noising method," *Soil Dyn. Earthq. Eng.*, vol. 30, no. 11, pp. 1168–1181, 2010.
- [8] M. Beenamol, S. Prabavathy, and J. Mohanalini, "Wavelet based seismic signal de-noising using Shannon and Tsallis entropy," *Comput. Math. with Appl.*, vol. 64, no. 11, pp. 3580–3593, 2012.
- [9] W. Zhu, S. M. Mousavi, and G. C. Beroza, "Seismic signal denoising and decomposition using deep neural networks," *IEEE Trans Geosci Remote Sens*, vol. 57, no. 11, pp. 9476–9488, 2019.
- [10] S. M. Mousavi and C. A. Langston, "Hybrid seismic denoising using higher-order statistics and improved wavelet block thresholding," *Bull. Seismol. Soc. Am.*, vol. 106, no. 4, pp. 1380–1393, 2016.
- [11] —, "Adaptive noise estimation and suppression for improving microseismic event detection," *J Appl Geophys*, vol. 132, pp. 116–124, 2016.
- [12] C. A. Langston and S. M. Mousavi, "Separating signal from noise and from other signal using nonlinear thresholding and scale-time windowing of continuous wavelet transforms separating signal from noise and from other signal," *Bull. Seismol. Soc. Am.*, vol. 109, pp. 1691–1700, 2019.

- [13] Y. Chen, "Non-stationary least-squares complex decomposition for microseismic noise attenuation," *Geophys J Int*, vol. 213, no. 3, pp. 1572–1585, 2018.
- [14] J. Han and M. van der Baan, "Microseismic and seismic denoising via ensemble empirical mode decomposition and adaptive thresholding," *Geophysics*, vol. 80, no. 6, pp. KS69–KS80, 2015.
- [15] N. Iqbal, E. Liu, J. H. McClellan, A. Al-Shuhail, S. I. Kaka, and A. Zerguine, "Detection and denoising of microseismic events using time–frequency representation and tensor decomposition," *IEEE Access*, vol. 6, pp. 22993–23006, 2018.
- [16] Y. Chen, "Automatic microseismic event picking via unsupervised machine learning," *Geophys. J. Int.*, vol. 212, no. 1, pp. 88–102, 2018.
- [17] R. Kimiaefar, H. Siahkoochi, A. Hajian, and A. Kalhor, "Seismic random noise attenuation using artificial neural network and wavelet packet analysis," *Arab. J. Geosci*, vol. 9, no. 3, p. 234, 2016.
- [18] H. Nuha, M. Deriche, and M. Mohandes, "Bilateral filters with elliptical gaussian kernels for seismic surveys denoising," in *Int. Conf. Front. Signal Process., ICFSP*. IEEE, 2019, pp. 141–144.
- [19] B. Zhao, K. He, L. Stankovic, and V. Stankovic, "Improving event-based non-intrusive load monitoring using graph signal processing," *IEEE Access*, vol. 6, pp. 53944–53959, 2018.
- [20] A. Gadde, S. K. Narang, and A. Ortega, "Bilateral filter: Graph spectral interpretation and extensions," in *IEEE Int. Conf. Image Inf. Process., IEEE ICIIP*. IEEE, 2013, pp. 1222–1226.
- [21] J. Akram and D. Eaton, "Adaptive microseismic event detection and automatic time picking," in *Soc. Explor. Geophys. Int. Expo. Annu. Meet., SEG: Expand. Geophys. Front.*, vol. 15, 2012.
- [22] S. M. Mousavi, W. Zhu, Y. Sheng, and G. C. Beroza, "Cred: A deep residual network of convolutional and recurrent units for earthquake signal detection," *Sci. Rep.*, vol. 9, no. 1, pp. 1–14, 2019.
- [23] E. Caffagni, D. W. Eaton, J. P. Jones, and M. van der Baan, "Detection and analysis of microseismic events using a matched filtering algorithm (mfa)," *Geophys. J. Int.*, vol. 206, no. 1, pp. 644–658, 2016.
- [24] L. Feng, V. Pazzi, E. Intrieri, T. Gracchi, and G. Gigli, "Joint detection and classification of rockfalls in a microseismic monitoring network," *Geophys. J. Int.*, 2020.
- [25] S. Qu, Z. Guan, E. Verschuur, and Y. Chen, "Automatic high-resolution microseismic event detection via supervised machine learning," *Geophys. J. Int.*, vol. 218, no. 3, pp. 2106–2121, 2019.
- [26] Y. Wu, Y. Lin, Z. Zhou, D. C. Bolton, J. Liu, and P. Johnson, "Deepdetect: A cascaded region-based densely connected network for seismic event detection," *IEEE Trans Geosci Remote Sens*, vol. 57, no. 1, pp. 62–75, 2018.
- [27] F. Provost, J.-P. Malet, C. Hibert, C. Abanco Martínez de Arenzana, and M. Hurlimann Ziegler, "Towards a standard typology of endogenous landslide seismic sources," *Earth Surf. Dyn.*, vol. 6, no. 4, pp. 1059–1088, 2018.
- [28] N. Vouillamoz, S. Rothmund, and M. Joswig, "Characterizing the complexity of microseismic signals at slow-moving clay-rich debris slides: the super-sauze (southeastern france) and pechgrabe (upper austria) case studies," *Earth Surf. Dyn.*, vol. 6, no. 2, 2018.
- [29] A. Tonnellier, A. Helmstetter, J.-P. Malet, J. Schmittbuhl, A. Corsini, and M. Joswig, "Seismic monitoring of soft-rock landslides: the super-sauze and valoria case studies," *Geophys. J. Int.*, vol. 193, no. 3, pp. 1515–1536, 2013.
- [30] B. Biescas, F. Dufour, G. Furdada, G. Khazaradze, and E. Suriñach, "Frequency content evolution of snow avalanche seismic signals," *Surv Geophys*, vol. 24, no. 5-6, pp. 447–464, 2003.
- [31] S. Esmaili, S. Krishnan, and K. Raahemifar, "Content based audio classification and retrieval using joint time-frequency analysis," in *ICASSP IEEE Int Conf Acoust Speech Signal Process Proc*, vol. 5. IEEE, 2004, pp. V–665.
- [32] J. Li, C. Yang, V. Stankovic, L. Stankovic, and S. Pytharouli, "Graph-based micro-seismic signal classification with an optimised feature space," in *Dig Int Geosci Remote Sens Symp (IGARSS)*, 2020.
- [33] F. Provost, C. Hibert, and J.-P. Malet, "Automatic classification of endogenous landslide seismicity using the random forest supervised classifier," *Geophys. Res. Lett.*, vol. 44, no. 1, pp. 113–120, 2017.
- [34] M. Curilem, A. Cuevas, R. Soto, F. Huenupan, C. Martin, M. S. Khan, F. Gil, C. Cardona, and L. Franco, "Classification of volcanic seismic events: An expert knowledge analysis for feature selection," in *8th Proc. Int. Conf. Pattern Recognit.* IET, 2017, pp. 2–6.
- [35] A. Maggi, V. Ferrazzini, C. Hibert, F. Beaucaud, P. Boissier, and A. Amemoutou, "Implementation of a multistation approach for automated event classification at piton de la fournaise volcano," *Seismol. Res. Lett.*, vol. 88, no. 3, pp. 878–891, 2017.
- [36] S. M. Mousavi, S. P. Horton, C. A. Langston, and B. Samei, "Seismic features and automatic discrimination of deep and shallow induced-microearthquakes using neural network and logistic regression," *Geophys. J. Int.*, vol. 207, no. 1, pp. 29–46, 2016.
- [37] S. Zaugg, M. Van Der Schaar, L. Houégnyan, C. Gervaise, and M. André, "Real-time acoustic classification of sperm whale clicks and shipping impulses from deep-sea observatories," *Appl Acoust*, vol. 71, no. 11, pp. 1011–1019, 2010.
- [38] C. Levy, D. Jongmans, and L. Baillet, "Analysis of seismic signals recorded on a prone-to-fall rock column (vercors massif, french alps)," *Geophys. J. Int.*, vol. 186, no. 1, pp. 296–310, 2011.
- [39] J. Kortström, M. Uski, and T. Tiira, "Automatic classification of seismic events within a regional seismograph network," *Comput Geosci*, vol. 87, pp. 22–30, 2016.
- [40] X. Shang, X. Li, A. Morales-Esteban, and G. Chen, "Improving microseismic event and quarry blast classification using artificial neural networks based on principal component analysis," *Soil Dyn. Earthq. Eng.*, vol. 99, pp. 142–149, 2017.
- [41] G. Zhao, M. Ju, L. Dong, X. Li, G. Chen, and C. Zhang, "Classification of mine blasts and microseismic events using starting-up features in seismograms," *Trans Nonferrous Met Soc China*, vol. 25, no. 10, pp. 3410–3420, 2015.
- [42] T. Joachims, "Making large-scale svm learning practical," *JFE Tech. Rep.*, Tech. Rep., 1998.
- [43] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, pp. 5–32, 2001.
- [44] P. E. E. Lara, C. A. R. Fernandes, A. Inza, J. I. Mars, J.-P. Métaixian, M. Dalla Mura, and M. Malfante, "Automatic multichannel volcano-seismic classification using machine learning and emd," *IEEE J Sel Top Appl Earth Obs Remote Sens*, vol. 13, pp. 1322–1331, 2020.
- [45] P. Venegas, N. Pérez, D. S. Benítez, R. Lara-Cueva, and M. Ruiz, "Building machine learning models for long-period and volcano-tectonic event classification," in *IEEE Chil. Conf. Electr., Electron. Eng., Inf. Commun. Technol., CHILECON*. IEEE, 2019, pp. 1–6.
- [46] D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," *arXiv preprint arXiv:1211.0053*, 2012.
- [47] H. Cai, H. Ren, Q. Wu, L. Peng, and C. Wen, "Semi-supervised fast algorithm for seismic waveform classification," in *Leading Edge. Society of Exploration Geophysicists and Chinese Petroleum Society*, 2018, pp. 1175–1179.
- [48] A. Sandryhaila and J. M. Moura, "Classification via regularization on graphs," in *IEEE Glob. Conf. Signal Inf. Process., GlobalSIP*. IEEE, 2013, pp. 495–498.
- [49] K. He, L. Stankovic, J. Liao, and V. Stankovic, "Non-intrusive load disaggregation using graph signal processing," *IEEE Trans Smart Grid*, vol. 9, no. 3, pp. 1739–1747, 2018.
- [50] P. Milanfar, "A tour of modern image filtering: New insights and methods, both practical and theoretical," *IEEE Signal Process Mag*, vol. 30, no. 1, pp. 106–128, 2012.
- [51] S. Taheri and J. Behboodian, "Neyman-pearson lemma for fuzzy hypotheses testing," *Metrika*, vol. 49, no. 1, pp. 3–17, 1999.
- [52] A. Trnkoczy, "Understanding and parameter setting of sta/lta trigger algorithm," in *New Manual of Seismological Observatory Practice (NMSOP)*. Deutsches GeoForschungsZentrum GFZ, 2009, pp. 1–20.
- [53] G. Roffo, "Feature selection library (matlab toolbox)," *arXiv preprint arXiv:1607.01327*, 2016.
- [54] P. A. Estévez, M. Tesmer, C. A. Perez, and J. M. Zurada, "Normalized mutual information feature selection," *IEEE Trans Neural Netw Learn Syst*, vol. 20, no. 2, pp. 189–201, 2009.
- [55] C. Yang, G. Cheung, and V. Stankovic, "Alternating binary classifier and graph learning from partial labels," in *Asia-Pac. Signal Inf. Process. Assoc. Annu. Summit Conf., APSIPA ASC - Proc.* IEEE, 2018, pp. 1137–1140.
- [56] S. M. Mousavi and C. A. Langston, "Automatic noise-removal/signal-removal based on general cross-validation thresholding in synchrosqueezed domain and its application on earthquake data," *Geophysics*, vol. 82, no. 4, pp. V211–V227, 2017.
- [57] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "Smote: synthetic minority over-sampling technique," *J Artif Intell Res*, vol. 16, pp. 321–357, 2002.
- [58] A. Fernández, S. Garcia, F. Herrera, and N. V. Chawla, "Smote for learning from imbalanced data: progress and challenges, marking the 15-year anniversary," *J Artif Intell Res*, vol. 61, pp. 863–905, 2018.