Enhancing the detection capabilities of nano-avalanches via data-driven classification of acoustic emission signals

Emil Bronstein [1,*] Jonathan Zimmerman [2] Eugen Rabkin [2] Eilon Faran [1] Ronen Talmon [3] and Doron Shilo [1]

[1] *Faculty of Mechanical Engineering, Technion–Israel Institute of Technology, 3200003 Haifa, Israel*

[2] *Department of Materials Science and Engineering, Technion–Israel Institute of Technology, 3200003 Haifa, Israel*

[3] *Viterbi Faculty of Electrical and Computer Engineering, Technion–Israel Institute of Technology, 3200003 Haifa, Israel*

(Received 8 June 2023; revised 23 August 2023; accepted 5 September 2023; published 5 October 2023)

Acoustic emission (AE) is a powerful experimental method for studying discrete and impulsive events termed avalanches that occur in a wide variety of materials and physical phenomena. A particular challenge is the detection of small-scale avalanches, whose associated acoustic signals are at the noise level of the experimental setup. The conventional detection approach is based on setting a threshold significantly larger than this level, ignoring "false" events with low AE amplitudes that originate from noise. At the same time, this approach overlooks small-scale events that might be true and impedes the investigation of avalanches occurring at the nanoscale, constituting the natural response of many nanoparticles and nanostructured materials. In this work, we develop a data-driven method that allows the detection of small-scale AE events, which is based on two propositions. The first includes a modification of the experimental conditions by setting a lower threshold compared to the conventional threshold, such that an abundance of small-scale events with low amplitudes are considered. Second, instead of analyzing several conventional scalar features (e.g., amplitude, duration, energy), we consider the entire waveform of each AE event and obtain an informative representation using dynamic mode decomposition. We apply the developed method to AE signals measured during the compression of platinum nanoparticles and demonstrate a significant enhancement of the detection range toward small-scale events that are below the conventional threshold.

DOI: 10.1103/PhysRevE.108.045001

I. INTRODUCTION

Avalanches are discrete and impulsive events that occur during slow changes applied to materials, for example, via electric or magnetic fields, mechanical loads, or temperature changes [1,2]. A variety of materials display a multitude of avalanches as part of a wide range of phenomena. This includes crack propagation [3], plasticity in metals [4,5], phase transformations [6–8], domain switching in ferromagnetic [9] and ferroelectric materials [7,10,11], twinning reorientation in metals and ferroelastic minerals [12,13], and the collapse of granular materials (e.g., rocks) [14]. The same processes occur across many scales. At larger scales, they can be found at the level of machine parts, mechanical structures, civil engineering structures, and the motion of tectonic plates [15,16]. At the smaller and, particularly, the nanoscale, avalanches have been reported in nanocrystals, nanoparticles [17–26], and ferroelectric relaxors [27,28].

Avalanches can be experimentally studied by various methods [11,29–37]. Still, the most prevalent method for experimental avalanche investigation is the acoustic emission (AE) method. The main advantage of the AE method is its ability to detect events that occur at small scales of length and time [38,39]. This capability is especially important because meaningful investigations of avalanche phenomena often rely on statistical analyses of data distributed over a wide range. Indeed, for many studied materials and phenomena, avalanche scalar features measured by AE (e.g., amplitude, energy, and duration) are distributed over several decades [2,9,38,40–48]. At the same time, there are phenomena such as in the above-mentioned nanostructures that inherently occur only at small scales that are comparable to or even smaller than the detection limit of the AE method. In such cases, the capability of the AE method to detect small-scale events is insufficient, and, hence, it is difficult to analyze the distributions of measured avalanche features [27].

The conventional way to determine the detection limit of the AE method is based on the evaluation of the level of *false* (artificial) events, which are unrelated to the studied process and originate from the surroundings of the sample or from noise in the experimental system. The common procedure to evaluate this noise level is based on measuring AE before the experiment, i.e., when no external stimuli are applied to the studied sample and no avalanches are expected to occur. The highest peak amplitude of the measured AE signal before the application of the stimulus is considered the detection limit, and usually, the user sets a threshold significantly larger than this value to ensure high certainty in the measured signals. Any event with a peak AE amplitude larger than the threshold is then considered as a *true* event that is associated with the sample response. Contrarily, all the events with peak amplitudes smaller than the threshold are *suspected of being false* and are, therefore, ignored. We note that many of the events

*emilbr@campus.technion.ac.il

with peak amplitudes smaller than the threshold *might be true* events. Consequently, the differentiation between the two event groups has attracted interest [49,50].

The distinction between true and false AE events is a challenging task due to four main reasons. First, the abrupt and nonlinear nature of avalanches produces a variety of AE events that are dissimilar to each other. Second, each event is broadband, i.e., comprising a vast frequency content. Third, the noise may originate from several sources, e.g., mechanical setup, electrical devices, and electromagnetic waves, resulting in broadband, colored noise. Fourth, the frequency contents of both the avalanches and the noise are setup-specific and vary between different samples and mechanical apparatuses. Due to these reasons, currently, there is no systematic method that does not rely on *a priori* assumptions concerning the physical origin of the avalanches or the noise and distinguishes between true and false events for signals with peak amplitudes smaller than the threshold.

In this work, we develop a data-driven classification method of true and false AE events, which enables the detection of true small-scale events that are below the conventional detection limit. Our method relies on two propositions. The first includes a modification of the experimental conditions by setting a *lower threshold* compared to the conventional threshold, such that small-scale events, whether true or false, are acquired. The second proposition includes obtaining an informative representation of each AE signal based on its *entire waveform*, using a recent extension of dynamic mode decomposition (DMD) [51,52], which is shown to be less sensitive to measurement noise [53] and allows for a systematic distinction between true and false events. Our three-stage method comprises of the application of the DMD-based extension to the measured AE signals, followed by the construction of their representations, ending with the classification of the events. We apply the method to AE signals measured during a simultaneous compression of hundreds of thousands of Pt nanoparticles, and demonstrate a significant enhancement of the detection range toward small-scale events that are below the conventional threshold.

II. DYNAMIC MODE DECOMPOSITION AND DELAY-COORDINATES EMBEDDING

Dynamic mode decomposition (DMD) is a data-driven technique for the analysis of dynamical systems [51,54–59]

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{f}(\boldsymbol{x}(t)) \tag{1}$$

from their measurements. DMD linearly approximates the dynamics $\boldsymbol{f}$ of a state vector $\boldsymbol{x} \in \mathbb{R}^q$ at time t as

$$\dot{\boldsymbol{x}} = \boldsymbol{\mathcal{A}}\boldsymbol{x}, \tag{2}$$

whose solution is given by [60,61]

$$\boldsymbol{x}(t) = \sum_{i=1}^{q} \exp(\mu_i t)\boldsymbol{\psi}_i \nu_i, \tag{3}$$

where μ_i and $\boldsymbol{\psi}_i$ are the eigenvalues and eigenvectors of $\boldsymbol{\mathcal{A}}$, and $\nu_i = \langle \boldsymbol{\psi}_i, \boldsymbol{x}(0)\rangle$ using the standard inner product in $\mathbb{R}^q$. In case of a damped oscillating system having oscillation frequencies ω_i and damping ratios ζ_i, the state vector can be represented as

$$\boldsymbol{x}(t) = \sum_{i=1}^{q} \exp(-\zeta_i\omega_i t \pm j\omega_i t)\boldsymbol{\psi}_i \nu_i, \quad j = \sqrt{-1}. \tag{4}$$

By sampling the state vector $\boldsymbol{x}(t)$ in time at a constant sampling rate of $f_s = 1/\Delta t$, the discrete measurements $\boldsymbol{x}_k = \boldsymbol{x}(k\Delta t)$ are acquired, and the following discrete-time dynamics can be analyzed:

$$\boldsymbol{x}_{k+1} = \boldsymbol{A}\boldsymbol{x}_k, \quad \boldsymbol{A} \in \mathbb{R}^{q\times q}. \tag{5}$$

The goal of DMD is to optimally approximate the operator $\boldsymbol{A}$ that captures the dynamics of the system. This goal is achieved by arranging the measurements $\boldsymbol{x}_k$, $k = 0, \ldots, m$, into the measurement and time-shifted measurement matrices, $\boldsymbol{X}$ and $\boldsymbol{X}'$, respectively, as

$$\boldsymbol{X} = \begin{bmatrix} | & | & & | \\ \boldsymbol{x}_0 & \boldsymbol{x}_1 & \cdots & \boldsymbol{x}_{m-1} \\ | & | & & | \end{bmatrix} \in \mathbb{R}^{q\times m},$$
$$\boldsymbol{X}' = \begin{bmatrix} | & | & & | \\ \boldsymbol{x}_1 & \boldsymbol{x}_2 & \cdots & \boldsymbol{x}_m \\ | & | & & | \end{bmatrix} \in \mathbb{R}^{q\times m}, \tag{6}$$

and solving $\boldsymbol{X}' = \boldsymbol{A}\boldsymbol{X}$. Then, using the $\tilde{r} = \mathrm{rank}(\boldsymbol{X})$ leading spectral components of $\boldsymbol{A}$, i.e., its eigenvalues $\tilde{\lambda}_i$ and eigenvectors $\tilde{\boldsymbol{\phi}}_i$, the measurements $\boldsymbol{x}_k$ can be represented and reconstructed in a data-driven manner as

$$\boldsymbol{x}_k = \sum_{i=1}^{\tilde{r}} \tilde{\lambda}_i^k \tilde{\boldsymbol{\phi}}_i \tilde{b}_i, \tag{7}$$

where $\tilde{b}_i$ is the ith entry of the DMD amplitudes $\tilde{\boldsymbol{b}}$ satisfying $\boldsymbol{x}_0 = \tilde{\boldsymbol{\Phi}}\tilde{\boldsymbol{b}}$, and $\tilde{\boldsymbol{\Phi}}$ is matrix whose columns are $\tilde{\boldsymbol{\phi}}_i$.

Using the DMD eigenvalues $\tilde{\lambda}_i$, the characteristics of the dynamical system, i.e., its oscillation frequencies ω_i and damping ratios ζ_i, can be obtained via [62]

$$\omega_i = \frac{|\ln(\tilde{\lambda}_i)|}{\Delta t}, \quad \zeta_i = -\frac{\mathrm{Re}(\ln(\tilde{\lambda}_i))}{|\ln(\tilde{\lambda}_i)|}. \tag{8}$$

In other words, Eq. (8) provides relations between the DMD eigenvalues and the characteristics of the system's behavior that have physical meanings.

Since the introduction of DMD by Schmid and Sesterhenn [51,52], it has gained significant interest in a variety of fields [63–65] and, particularly, in fluid dynamics [54,55,57,59,66,67]. Also, numerous variants of DMD have been presented ever since, such as the exact DMD [55], extended DMD [56], higher order DMD [68], and sparsity-promoting DMD [69]. Moreover, many utilities of DMD have been shown, such as its ability to predict the system's state at future, nonmeasured time instances [60,64,70,71].

One of the main limitations of DMD is that its straightforward application to one-dimensional signals (i.e., $q = 1$), such as AE signals, results in a degenerate (scalar) representation of the dynamics $\boldsymbol{A}$ in (5). The combination of DMD and the delay-coordinates embedding [72], which is a method for augmenting past measurements to the present measurement, has been shown to circumvent this limitation [55,60,68,70]. In

this combination, termed delay-coordinates DMD, the measurement matrices in (6) are *augmented* by an augmentation number $s < m$, $s \in \mathbb{N}$, such that

$$\hat{\boldsymbol{X}} = \begin{bmatrix} x_0 & x_1 & \cdots & x_{m-s-1} \\ x_1 & x_2 & \cdots & x_{m-s} \\ \vdots & \vdots & \vdots & \vdots \\ x_s & x_{s+1} & \cdots & x_{m-1} \end{bmatrix} \in \mathbb{R}^{[s+1]\times[m-s]},$$

$$\hat{\boldsymbol{X}}' = \begin{bmatrix} x_1 & x_2 & \cdots & x_{m-s} \\ x_2 & x_3 & \cdots & x_{m-s+1} \\ \vdots & \vdots & \vdots & \vdots \\ x_{s+1} & x_{s+2} & \cdots & x_m \end{bmatrix} \in \mathbb{R}^{[s+1]\times[m-s]}, \quad (9)$$

where $\hat{\boldsymbol{X}}$ and $\hat{\boldsymbol{X}}'$ are the augmented observation matrices, and x_k, $k = 0, \ldots, m$, are the measurements of a one-dimensional signal. Now $\hat{\boldsymbol{A}} \in \mathbb{R}^{[s+1]\times[s+1]}$, which can be found from $\hat{\boldsymbol{X}}' = \hat{\boldsymbol{A}}\hat{\boldsymbol{X}}$, is no longer degenerate, and its spectral components can be computed and used in (7) and (8).

Recently, by studying delay-coordinates DMD, we have revealed a unique spatiotemporal coupling between the DMD eigenvalues and modes, which we term the *spatiotemporal coupling in delay coordinates DMD* (scdcDMD) [53]. Based on this coupling, we proposed an algorithm for obtaining the r DMD components that truly describe the dynamical system, from which compact representations of signals similar to (7) are constructed as

$$\boldsymbol{x}_k = \sum_{i=1}^{r} \lambda_i^k \boldsymbol{\phi}_i b_i, \quad (10)$$

where λ_i, $\boldsymbol{\phi}_i$, and b_i are the DMD eigenvalues, modes, and amplitudes, respectively, required for obtaining this compact representation. We have shown that using the scdcDMD-based algorithm [53], successful representations can be achieved even when the measurements $\boldsymbol{x}_k$ are corrupted with white standard Gaussian noise. As any measured signal is corrupted with some noise arising from the experimental setup, including the AE signals analyzed in this paper, the scdcDMD-based algorithm is utilized to construct the representations of the signals as part of our proposed method.

III. EXPERIMENTAL SETUP AND STUDIED MATERIAL

The experimental setup, illustrated in Fig. 1, is designed to measure AE during plastic deformation in platinum (Pt) nanoparticles induced by their slow compression in an INSTRON 5966 testing machine. The nanoparticles are fabricated using a solid-state dewetting method, which has been described in detail in several previous works [22,73,74]. In short, a 30 nm Pt film is deposited on a 2.5 × 2.5 mm^2 basal plane-oriented sapphire substrate and annealed for 24 hours in air at the temperature of 1000 °C. The film breaks into an array of about 800 000 dislocation-free nanoparticles with an average diameter of 400 nm and spacing of approximately 1 μm.

In this fabrication method, previous studies based on microcompression of single nanoparticles have shown that the deformation mode of the nanoparticles begins with a purely elastic strain up to ≈10%, followed by massive avalanche-like

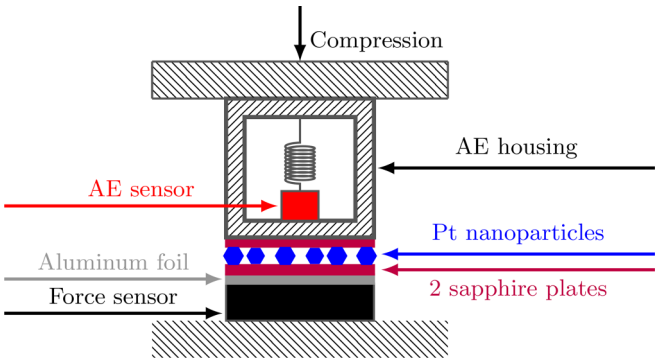

FIG. 1. Illustration of the experimental setup. Platinum (Pt) nanoparticles, which are deposited on the bottom sapphire plate (substrate), are placed on a thin aluminum foil layer, and both are stacked on the INSTRON force sensor. The AE sensor and the mechanical spring that presses on it are placed in the AE housing. The top surface of the housing is connected to the moving bridge of the INSTRON machine, and its bottom surface compresses the nanoparticles through the top sapphire plate. The aluminum foil accommodates possible misalignment during the compression.

dislocation nucleation and particle collapse [21,22,24,26,75–77]. Due to the small volume of the tested particles, the abrupt and stochastic nature of dislocation nucleation events, and the subsequent fast dislocation glide, the testing instrument (the nanoindenter's rigidity) and loading regime (load or displacement control) very likely play an unknown role in the initiation of dislocation avalanches in these tests.

The nanoparticles and substrate are placed on an INSTRON 10 kN force sensor. A flat sapphire plate (similar to the substrate) is placed above the nanoparticles and a thin layer of aluminum foil is placed between the samples and the force sensor to maintain good contact and accommodate possible misalignment during the experiment. A broadband AE sensor (Fujicera 1045S), having a bandwidth of 100–1500 kHz, is placed in an aluminum housing in accordance with common practice [12,43], where a mechanical spring with small stiffness pushes against the sensor. This assembly ensures that the sensor is subjected to a constant load and protects the sensor's surface from scratches. The housing is rigidly connected to the moving bridge of the INSTRON machine.

During an experiment, the nanoparticles are compressed at a constant rate of 0.05 mm/min, inducing plastic collapse events in them. Each such event (avalanche) produces acoustic waves that are detected by the AE sensor, amplified by a dedicated amplifier with a gain of 40 dB (Vallen AEP5 preamplifier), and sampled at 10 MSa/s by a Vallen AMSY-6 acquisition system. The force sensor measures the overall force applied to the samples and is simultaneously acquired (by the same acquisition system) at 1 kSa/s. The mass particle compression by a probe (AE housing), whose size and weight are larger by orders of magnitude than those of individual metal particles, is expected to circumvent uncertainties associated with microcompression testing instruments and answer the question of whether the dislocation avalanches can also occur when the nanoparticles are compressed by an "ideal" instrument that is unaffected by the individual avalanche events.

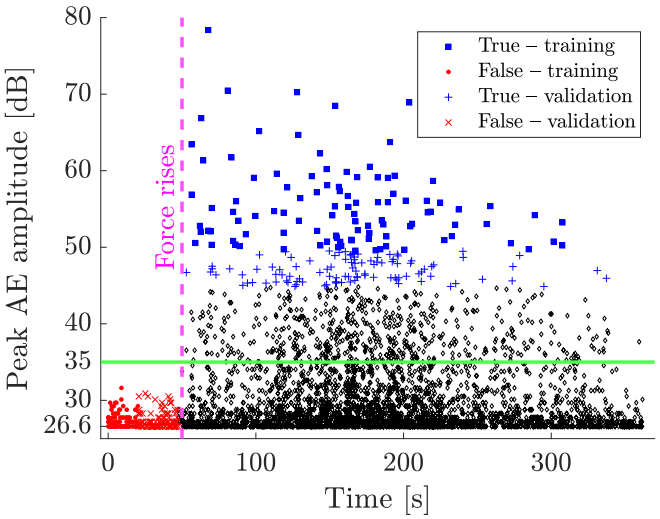

FIG. 2. The AE events detected in the experiment using a threshold of 26.6 dB. Out of them, four subsets of data points are marked as follows. The first 200 detected events are false and are equally divided between the subsets, such that $S_{\text{false}}^{\text{training}}$ contains events 1–100, marked by red dots, and $S_{\text{false}}^{\text{validation}}$ contains events 101–200, marked by red crosses. The subsets $S_{\text{true}}^{\text{training}}$ and $S_{\text{true}}^{\text{validation}}$ contain the 102 and 103 true events with the largest and next largest peak amplitudes, respectively. The former and latter are denoted by blue squares and plus signs, respectively. The rest of the data points are marked by black diamonds. After 50 s from the onset of the experiment, marked by the vertical dashed magenta line, the force significantly rises. The green horizontal line denotes a threshold of 35 dB.

Figure 2 shows 3419 hits (AE events) detected in the experiment using a threshold of 26.6 dB. This value was set to be relatively low to increase the sensitivity of the measurement system to small-scale events at the expense of possible detection of false events. Yet, out of the total amount of events, some can be classified as true with high certainty or as definitely false. For instance, the first 200 events occurring in the first 47 s of the experiment, denoted by red dots and crosses, are definitely false events because they were detected before any significant load was applied to the samples. Additionally, the 205 events with the largest peak amplitudes, denoted by blue squares and plus signs, can be considered true events because their peak amplitudes are significantly larger than 35 dB (green horizontal line), which is the threshold value that is larger than the peak amplitudes of all of the first 200 false events.

The subsets of the 200 false and 205 true events are denoted by S_{false} and S_{true}, respectively. These two subsets are utilized for the classification and validation of the AE events in Secs. IV and V. To this end, S_{false} and S_{true} are split into training and validation data sets as follows. S_{false} is equally split with respect to the elapsed experiment time; that is, $S_{\text{false}}^{\text{training}}$ and $S_{\text{false}}^{\text{validation}}$ contain events 1–100 (training, red dots) and events 101–200 (validation, red crosses), respectively. The events in S_{true} are ordered in descending order according to their peak amplitudes, and this subset is split such that $S_{\text{true}}^{\text{training}}$ contains the first 102 training events (blue squares) and $S_{\text{true}}^{\text{validation}}$ contains the remaining 103 validation events (blue plus signs).

Figure 3 shows the application of the scdcDMD-based algorithm to four representative measured signals (solid lines), demonstrating their successful reconstructions, as shown by the dotted black lines. Consequently, the DMD components used in these reconstructions, and, specifically, the eigenvalues from which the characteristics of the signals are extracted via (8), can be considered to accurately describe the signals. To demonstrate the complexity of the problem, Fig. 3 presents a true event from $S_{\text{true}}^{\text{training}}$ that corresponds to 55.9 dB [Fig. 3(a)], the first false event in $S_{\text{false}}^{\text{training}}$ [Fig. 3(b)], and two events with unknown labels, one below the threshold and one above it, corresponding to peak amplitudes of 27.2 and 36.9 dB [Figs. 3(c) and 3(d), respectively]. It can be seen that the waveform of the smaller event in Fig. 3(c) resembles that of the true event in Fig. 3(a), despite having a peak amplitude below the threshold. At the same time, the waveform of the larger event in Fig. 3(d) resembles that of the false event in Fig. 3(b), despite having a peak amplitude above the threshold.

IV. METHOD FOR DATA-DRIVEN CLASSIFICATION OF ACOUSTIC EMISSION SIGNALS

Currently, a classification method of AE events is absent, where the threshold-based method circumvents the classification task, and merely considers *all* events whose peak amplitudes are above the threshold as true, as the top red route in Fig. 4 shows. To address this issue, we develop the data-driven method illustrated by the bottom blue route in Fig. 4. Our method relies on the scdcDMD-based algorithm to obtain physically informative variables that characterize each AE signal, which are then transferred to high-dimensional representations of the signals, followed by their dimensionality reduction. The method ends with a classification model, which is built based on the training subset and excellently classifies events from the (different) validation subset.

Our three-stage classification method begins with the application of the scdcDMD-based algorithm to each of the measured AE signals, followed by the extraction of $\{\omega_i, \zeta_i, b_i\}_{i=1}^{r}$ using (8) and (10). That is, the oscillation frequencies ω_i, damping ratios ζ_i, and DMD amplitudes b_i, respectively, that correspond to each AE signal are obtained. Then ω_i and ζ_i are converted to the physically informative variables $f_i = \omega_i/(2\pi)$ with units of Hz and $\tau_i = (\omega_i\zeta_i)^{-1}$, which have units of seconds and represent the decay or growth times of a signal. Eventually, we create the triples $\{f_i, \tau_i, b_i\}_{i=1}^{r}$ that characterize each AE signal.

Note that each representation $\{f_i, \tau_i, b_i\}_{i=1}^{r}$ of an AE signal has a different amount of DMD components, denoted by r, and hence, a different number of triplets corresponding to each signal. Considering all 3419 events detected in our experiment (see Fig. 2), r varies from 117 to 1163, with a median value of 670. Namely, each AE signal is represented by a multitude of oscillation frequencies and decay or growth times.

The AE signals cannot be represented well by only several dominant DMD components that correspond to the largest DMD amplitudes. For example, the five dominant DMD components of the signals in the false and true training subsets constitute on average 15% and 24% of the total sum of all DMD amplitudes of these signals, respectively. The same

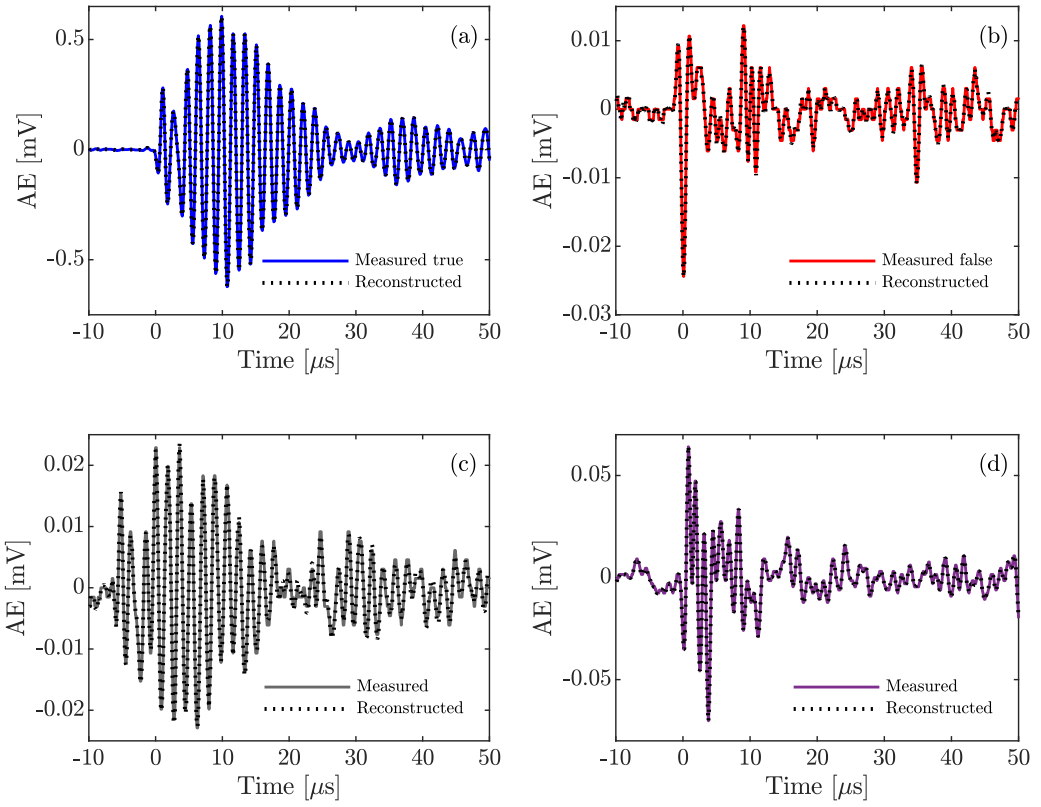

FIG. 3. The first 50 μs of four measured AE signals (solid lines) and their reconstructions (dotted black lines) using our scdcDMD-based algorithm [53]. (a) A true event. (b) A false event. Events with unknown labels having peak amplitudes of 27.2 dB (c) and 36.9 dB (d).

calculations for the 10 dominant DMD components provide the values 23% and 37%, respectively. These results indicate that representations of AE signals based on a few dominant components are deficient, and a different, high-dimensional representation that accounts for all (or, at least, most) DMD components is required. We propose such a representation in the next stage of our method.

In stage 2, we propose a representation for the AE signals. Specifically, we begin by representing all the signals via histograms by accounting for all of their DMD components and using the same set of parameters. For this purpose, a binning process is performed on the triplets $\{f_i, \tau_i, b_i\}_{i=1}^{r}$.

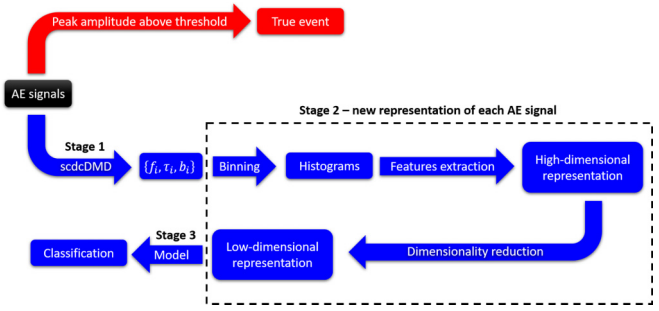

FIG. 4. Currently, a method that classifies AE events as *true* and *false* does not exist, where the common practice is to consider *every* detected event whose peak amplitude exceeds a predetermined threshold as *true* (top red route). The bottom blue route schematically shows our three-stage data-driven classification method.

Concretely, the f_i and τ_i of each event n are divided into the same bins of equal width. As the frequency bandpass of the AE sensor is 0.1–1.5 MHz, we divide this range into N_{bins} of width $1.4\,\text{MHz}/N_{\text{bins}}$, and discard f_i that are not in this range and their corresponding b_i, where N_{bins} is the number of bins. Since the majority of the τ_i in our data are found in the range $[-0.2, 0.2]$ ms, we divide them into N_{bins} of equal widths given by $0.4\,\text{ms}/N_{\text{bins}}$. The rest of the τ_i are assigned to two additional bins related to $\tau_i < -0.2$ ms and $\tau_i > 0.2$ ms. Additionally, we define the normalized DMD amplitudes as $\hat{b}_i = |b_i|/\sum_k |b_k|$.

Then the counts $c_n^{(j)}$ are calculated as the sums of $\hat{b}_i$ that correspond to each bin, where $1 \leqslant j \leqslant N_{\text{bins}}$. Namely, the calculations of the counts that correspond to the oscillation frequencies (denoted by the subscript f) and the decay or growth times (denoted by the subscript τ) of event n are

$$c_{n,f}^{(j)} = \sum_{l=1}^{N_f} \hat{b}_l, \quad c_{n,\tau}^{(j)} = \sum_{l=1}^{N_\tau} \hat{b}_l, \tag{11}$$

where N_f and N_τ are the amounts of f_l and τ_l in each bin. The result of the binning process can be visualized by histograms whose horizontal axes are bins of f_i or τ_i, and the vertical axes are the counts $c_{n,f}^{(j)}$ or $c_{n,\tau}^{(j)}$ in (11), respectively. Figure 5 shows the histograms of f_i and τ_i for the choice of $N_{\text{bins}} = 40$ of the false events $n = 1$ [Figs. 5(a) and 5(b)] and $n = 101$ [Figs. 5(c) and 5(d)], where the waveform of the former is depicted in Fig. 3(b).

We note that although false signals originate from noise in the experimental setup, their waveforms have some structure

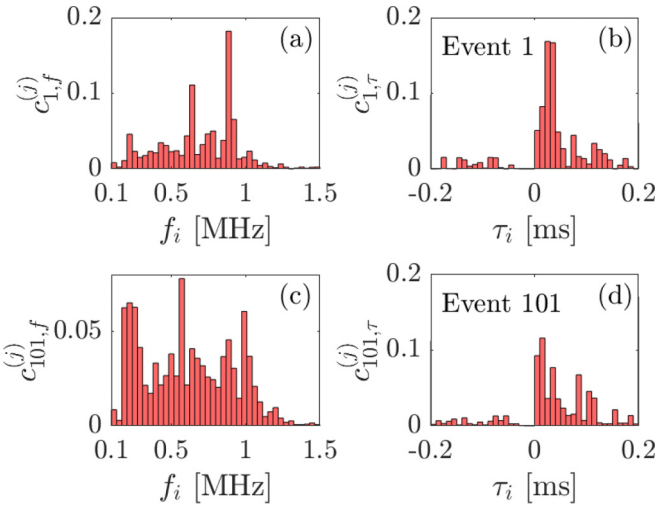

FIG. 5. Histograms of the frequencies f_i and the decay or growth times τ_i of the two false events 1 (a, b) and 101 (c, d). The histograms are obtained by applying the binning process described in stage 2 of the developed method to these events with $N_{\text{bins}} = 40$.

(i.e., some dominant frequency and decay or growth time ranges), and their characteristics differ from those of white noise. The demonstration in Fig. 5 conveys the complexity of identifying false AE signals that have their own characteristic patterns, which differ from one false event to another false event.

Eventually, a features vector $\boldsymbol{c}_n$ of dimension $(2N_{\text{bins}} + 2)$ that represents the nth AE signal is defined as

$$\boldsymbol{c}_n = \left[c_{n,\tau}^{(1)} \quad \cdots \quad c_{n,\tau}^{(N_{\text{bins}}+2)} \quad c_{n,f}^{(1)} \quad \cdots \quad c_{n,f}^{(N_{\text{bins}})}\right]. \tag{12}$$

Using (12), a features matrix $\boldsymbol{M}$ of the events is built, such that its rows are the features vectors $\boldsymbol{c}_n$ that correspond to each AE event n, $1 \leqslant n \leqslant N_{\text{events}}$, as

$$\begin{aligned}\boldsymbol{M} &= \begin{bmatrix} \boldsymbol{c}_1 \\ \vdots \\ \boldsymbol{c}_{N_{\text{events}}} \end{bmatrix} \\ &= \begin{bmatrix} c_{1,\tau}^{(1)} & \cdots & c_{1,\tau}^{(N_{\text{bins}}+2)} & c_{1,f}^{(1)} & \cdots & c_{1,f}^{(N_{\text{bins}})} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ c_{N_{\text{events}},\tau}^{(1)} & \cdots & c_{N_{\text{events}},\tau}^{(N_{\text{bins}}+2)} & c_{N_{\text{events}},f}^{(1)} & \cdots & c_{N_{\text{events}},f}^{(N_{\text{bins}})} \end{bmatrix} \\ &\times \in \mathbb{R}^{N_{\text{events}} \times [2N_{\text{bins}}+2]}. \end{aligned} \tag{13}$$

Namely, the matrix $\boldsymbol{M}$ constitutes a high-dimensional representation of the events.

Next, the principal components analysis (PCA) [78,79] is applied to $\boldsymbol{M}$, resulting in the dimensionality reduction of its columns from a space of $(2N_{\text{bins}} + 2)$ dimensions to a two-dimensional (2D) plane. PCA linearly transforms a data set from its original space into a coordinate system that captures most of the information in the data with fewer dimensions. More concretely, the two principal directions of the coordinate system obtained by PCA describe the largest variation in the data in two dimensions [80].

The PCA of $\boldsymbol{M}$ begins with its singular value decomposition (SVD),

$$\boldsymbol{M} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^*, \tag{14}$$

where $\boldsymbol{U} \in \mathbb{C}^{N_{\text{events}} \times r_M}$, $\boldsymbol{\Sigma} \in \mathbb{C}^{r_M \times r_M}$, $\boldsymbol{V} \in \mathbb{C}^{[2N_{\text{bins}}+2] \times r_M}$, $r_M = \text{rank}(\boldsymbol{M})$, and $(\cdot)^*$ is the complex conjugate transpose of its argument. The columns of $\boldsymbol{V}$ are the *principal directions* for $\boldsymbol{M}$. Hence, the projection $\boldsymbol{MV}$ expresses the events in $\boldsymbol{M}$ in terms of the principal directions in $\boldsymbol{V}$. To obtain the 2D expression of the events, we write $\boldsymbol{V} = [\boldsymbol{V}_{1-2}, \boldsymbol{V}_{3-r_M}]$, where $\boldsymbol{V}_{1-2}$ and $\boldsymbol{V}_{3-r_M}$ contain the first two and the remaining $(r_M - 2)$ columns of $\boldsymbol{V}$, respectively, and define

$$\boldsymbol{v} = \boldsymbol{M}\boldsymbol{V}_{1-2} \in \mathbb{R}^{N_{\text{events}} \times 2}. \tag{15}$$

At the end of stage 2, each event can be alternatively represented (and conveniently visualized) in two dimensions via (15).

We note that the amount of dimensionality reduction techniques is vast, e.g., linear discriminant analysis (LDA) [81], quadratic discriminant analysis (QDA) [82], diffusion maps [83], and alternating diffusion [84]. We choose to use PCA because, on the one hand, it is very simple and easy to implement as Eq. (14) and (15) suggest, and on the other hand, it provides excellent results on the events studied in this paper, as presented in Sec. V.

By employing Eq. (13)–(15) on $S_{\text{false}}^{\text{training}}$ and $S_{\text{true}}^{\text{training}}$, a 202-by-2 vector $\boldsymbol{v}^{\text{training}}$ is obtained as

$$\boldsymbol{v}^{\text{training}} = \begin{bmatrix} \boldsymbol{v}_{\text{false}}^{\text{training}} \\ \\ \boldsymbol{v}_{\text{true}}^{\text{training}} \end{bmatrix}, \tag{16}$$

where $\boldsymbol{v}_{\text{false}}^{\text{training}}$ and $\boldsymbol{v}_{\text{true}}^{\text{training}}$ contain the 100 and 102 principal components of the false and true training events, respectively. Figure 6(a) depicts $\boldsymbol{v}_{\text{false}}^{\text{training}}$ in red dots and $\boldsymbol{v}_{\text{true}}^{\text{training}}$ in blue squares, where a nearly complete separation between the true and false training events can be observed.

In the last stage of the developed method, a classification model is built based on the principal components in $\boldsymbol{v}^{\text{training}}$, which determines whether a newly introduced event (i.e., not included in the training subsets) is true or false. Once the low-dimensional representation of the events as in Fig. 6(a) is at hand, different classification methods can be employed, e.g., logistic regression [85], k-nearest neighbors [86,87], etc.

As the principal components in $\boldsymbol{v}_{\text{false}}^{\text{training}}$ are concentrated at the top half-plane in Fig. 6(a) and almost do not mix with the principal components in $\boldsymbol{v}_{\text{true}}^{\text{training}}$ that are concentrated at the bottom half-plane, we use support vector machine (SVM) [88] as the binary classification model. Since the training events in Fig. 6(a) are not linearly separable, we train an off-the-shelf soft-margin classifier as a part of our method by using the fitcsvm MATLAB function with a regularization term (box constraint) of 500 [89], which produces the hyperplane given by the magenta line. This hyperplane divides the plane spanned by principal components 1 and 2 into top and bottom halves corresponding to regions of false and true events, respectively. Evidently, the resulting hyperplane separates the training subsets well, as 100% of the true training events (blue

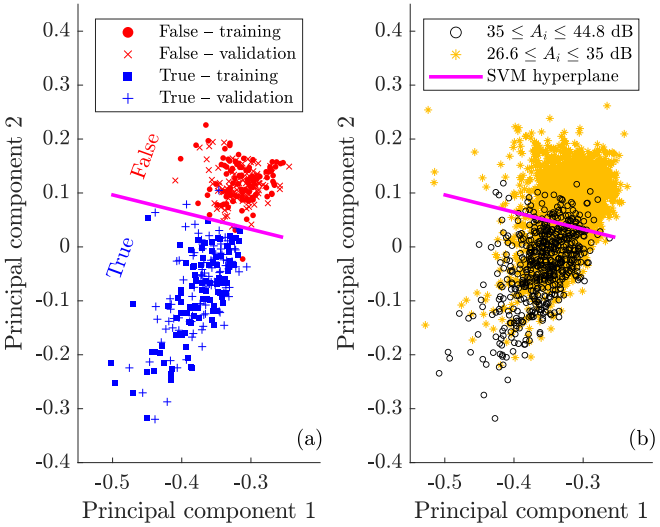

FIG. 6. (a) The principal components of $S_{\text{false}}^{\text{training}}$ (red dots), $S_{\text{false}}^{\text{validation}}$ (red crosses), $S_{\text{true}}^{\text{training}}$ (blue squares), and $S_{\text{true}}^{\text{validation}}$ (blue plus signs). The magenta line represents the support vector machine (SVM) hyperplane that serves as the classification model in our method, dividing the data points into two groups. A nearly complete separation between the true and false training events, as well as a successful classification of the true and false validation events, are achieved using our method. (b) The principal components of the events whose peak AE amplitudes, A_i, correspond to $35 \leqslant A_i \leqslant 44.8$ dB (black circles) and $26.6 \leqslant A_i \leqslant 35$ dB (orange asterisks), where 35 dB is the conventional threshold. The events in (a) do not appear in (b) and vice versa, and their union constitutes the entire data set of events.

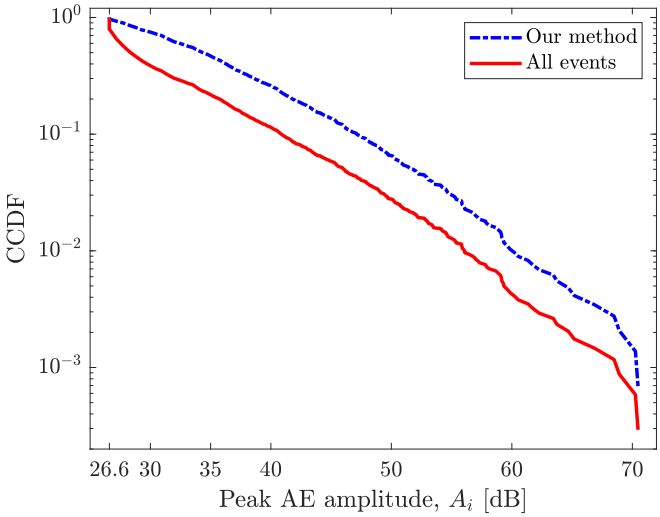

FIG. 7. The complementary cumulative distribution function (CCDF) of the peak AE amplitudes, A_i, of the events classified as true by our method (dashed blue curve) and of all the events detected using the lower threshold of 26.6 dB (solid red curve), without any screening process applied to them.

squares) and 98% of the false training events (red dots) are found on either side of the hyperplane.

V. RESULTS

To evaluate the performance of the developed method in Sec. IV, it is employed on the validation data set comprising the subsets $S_{\text{true}}^{\text{validation}}$ and $S_{\text{false}}^{\text{validation}}$ that contain 103 true and 100 false events, which are denoted by the blue plus signs and by the red crosses in Fig. 6(a), respectively. That is, the label (true or false) of each event in these subsets is known. It is observed that the developed method correctly identifies 98% of the true validation events and 99% of the false validation events. These results mean that the model, which is built based on the training data set, successfully classifies events from the validation data set.

Figure 6(b) shows two additional subsets of events that are not in the training or validation data sets, and whose peak amplitudes, A_i, are in the range $35 \leqslant A_i \leqslant 44.8$ dB (black circles, above the threshold) and $26.6 \leqslant A_i \leqslant 35$ dB (orange asterisks, below the threshold). Almost all of the events in the former subset, denoted by S_{above}, and many of the events in the latter subset are labeled true by our method.

Now we compare the results of our classification method with the threshold-based method. First, we consider the 764 events above the threshold of 35 dB [black circles in Fig. 6(b)], most of which are expected to be true. Indeed, our method classifies 91% (695 out of 764) of them as true. Further, 754 small-scale events with peak amplitudes below 35 dB [orange asterisks in Fig. 6(b)], which are not captured by the threshold-based method, are detected and classified as true by our method. In total, our method classifies 1449 events as true, thus doubling the population of true events compared to the threshold-based method. This expansion provides more information about nano-avalanches.

Figure 7 shows the complementary cumulative distribution function (CCDF) of A_i of the events labeled true by our method (dashed blue curve) and of all the events detected using the lower threshold of 26.6 dB (solid red curve), without any screening process applied to them. As expected, both curves display a similar (almost identical) behavior for $A_i \geqslant 35$ dB, since our method classifies nearly all the events in this range as true. Contrarily, the two curves show entirely different trends for $26.6 \leqslant A_i \leqslant 35$ dB, where the upward deflection of the solid red curve implies an excess of false events in this range.

To verify that the peak AE amplitudes of the events, A_i, do not have a strong influence on our classification results, we compute the distances, d_i, of each point in Fig. 6 from the hyperplane (the magenta line). Figure 8 shows A_i vs d_i, where the markings are as in Fig. 6. Naturally, because most of the false events are below the threshold, there is some correlation between d_i and A_i. Yet, for the events with peak amplitudes above the threshold, this correlation appears to be weak. To quantify this observation, we calculate the Pearson correlation coefficient r_{dA} between d_i and A_i. The calculations of r_{dA} for $S_{\text{true}}^{\text{training}}$, $\{S_{\text{true}}^{\text{training}} \cup S_{\text{true}}^{\text{validation}}\}$, and $\{S_{\text{true}}^{\text{training}} \cup S_{\text{true}}^{\text{validation}} \cup S_{\text{above}}\}$ yield the values 0.05, −0.02, and −0.14, respectively. Additionally, the waveforms shown in Figs. 3(c) and 3(d) correspond to the events marked in Fig. 8 by the gray triangle and by the purple star, respectively. Due to the resemblance of the waveform of the true event in Fig. 3(a) to the waveform in Fig. 3(c) of the event marked by the gray triangle, it is plausible that the latter is true despite its small peak

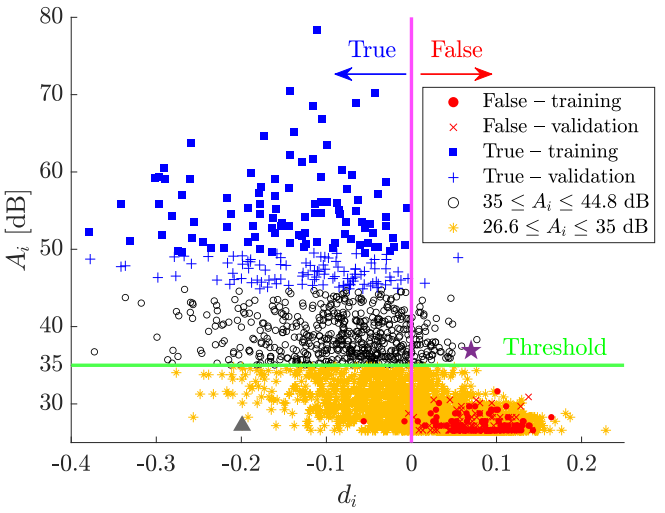

FIG. 8. The peak amplitudes of the events, A_i, vs the distances, d_i, from the SVM hyperplane to the low-dimensional representations of the events in Fig. 6. Events to the left and to the right of the SVM hyperplane (vertical magenta line) are classified as true and false by our method, respectively. Events above and below the conventional threshold (horizontal green line) are considered and discarded by the threshold-based method, respectively. The waveforms of the events marked by the gray triangle and by the purple star appear in Figs. 3(c) and 3(d), respectively.

amplitude and in accordance with the classification of our method. Similarly, due to the resemblance of the waveform of the false event in Fig. 3(b) to the waveform in Fig. 3(d) of the event marked by the purple star, it is plausible that the latter is false despite its large peak amplitude, also in accordance with the classification of our method.

VI. CONCLUSIONS

AE is a highly sensitive technique that is capable of detecting small-scale avalanche events. Commonly, the AE detection limit is determined based solely on the signal amplitude by defining a threshold that is significantly larger than the noise level. This method ignores numerous true avalanche events with amplitudes below the threshold and limits the detection capabilities of the AE method.

In this work, we introduced a method that enhances the detection capability of the AE method by analyzing the *entire waveforms* of the AE signals. Using the scdcDMD-based algorithm, we found the spectral components of all captured signals, both the true ones that are generated within the sample by the studied process, and the false ones that are generated by the surroundings of the sample or by the experimental system. Our method has been shown to successfully extract physically informative variables that characterize both true and false AE signals, based on which high-dimensional representations of the events were built. Our results demonstrate that false events are different from white noise and have some spectral characteristics. Therefore, their identification is generally challenging and, particularly, is extremely difficult (and, perhaps, impossible) to achieve based on physical intuition or prior experience.

We developed a data-driven classification method that distinguishes between true and false AE events, thus expanding the detection range towards nanoscale events that are below the detection capability of the conventional threshold-based approach. Application of PCA to the high-dimensional representations of all detected events resulted in their dimensionality reduction to a 2D plane, whose axes capture most of the information in the data. At the low-dimensional space, we used two data sets with known labels (true or false) to propose an SVM-based classification model, which has been shown to excellently classify events from a validation data set it has not been trained on. The classification method does not rely on any *a priori* assumption concerning the physical origin of the avalanches or the noise and is, therefore, material- and process-agnostic, and can be utilized to study a variety of phenomena, especially processes that occur at the nanoscale.

We demonstrated the capabilities of the method in a study of avalanches that occur during the plastic collapse of Pt nanoparticles. Application of our method to all detected AE events increased the number of events classified as true by almost twofold compared to the number of events considered by the threshold-based method. From a materials science aspect, the avalanche events identified in the present study during mass compression of Pt nanoparticles indicate that dislocation avalanches represent an intrinsic plasticity mode of pristine metal nanoparticles that is unaffected by the interference from the testing instrument.

This demonstration shows that the developed method can be applied in studies of nanoprocesses, in which avalanches inherently occur only at small scales that are comparable to or smaller than the detection limit of the AE method. Thus, it paves the way for studying avalanches in materials such as nanocrystals and nanoparticles, ferroelectric relaxors, strain glasses, and other nanostructured materials. Furthermore, the scdcDMD-based representations of entire AE signals lay the foundation for other classification tasks, such as the distinction between different physical processes that give rise to the AE signals.

ACKNOWLEDGMENTS

E.B., E.F., and D.S. wish to thank the Israel Science Foundation (Grant No. 1600/22) for partial support of this study. J.Z. and E.R. wish to thank the Israel Science Foundation (Grant No. 617/19) for partial support of this study. E.B. is grateful to the Azrieli Foundation for the award of an Azrieli Fellowship. R.T. acknowledges the support of the Schmidt Career Advancement Chair. Helpful discussions with Prof. Noy Cohen are heartily appreciated.

[1] J. P. Sethna, K. A. Dahmen, and C. R. Myers, Crackling noise, Nature (London) **410**, 242 (2001).

[2] E. K. Salje and K. A. Dahmen, Crackling noise in disordered materials, Annu. Rev. Condens. Matter Phys. **5**, 233 (2014).

[3] T. Roberts and M. Talebzadeh, Acoustic emission monitoring of fatigue crack propagation, J. Constr. Steel Res. **59**, 695 (2003).

[4] M. Lebyodkin, N. Kobelev, Y. Bougherira, D. Entemeyer, C. Fressengeas, V. Gornakov, T. Lebedkina, and I. Shashkov, On the similarity of plastic flow processes during smooth and jerky flow: Statistical analysis, Acta Mater. **60**, 3729 (2012).

[5] M. Lebyodkin, N. Kobelev, Y. Bougherira, D. Entemeyer, C. Fressengeas, T. Lebedkina, and I. Shashkov, On the similarity of plastic flow processes during smooth and jerky flow in dilute alloys, Acta Mater. **60**, 844 (2012).

[6] L. Straka, V. Novák, M. Landa, and O. Heczko, Acoustic emission of Ni–Mn–Ga magnetic shape memory alloy in different straining modes, Mater. Sci. Eng., A **374**, 263 (2004).

[7] M. Roth, E. Mojaev, E. Dul'kin, P. Gemeiner, and B. Dkhil, Phase transition at a nanometer scale detected by acoustic emission within the cubic phase $Pb(Zn_{1/3}Nb_{2/3})O_{3-x}PbTiO_3$ relaxor ferroelectrics, Phys. Rev. Lett. **98**, 265701 (2007).

[8] G. F. Nataf and E. K. Salje, Avalanches in ferroelectric, ferroelastic and coelastic materials: Phase transition, domain switching and propagation, Ferroelectrics **569**, 82 (2020).

[9] D. L. Beke, L. Daróczi, L. Z. Tóth, M. K. Bolgár, N. M. Samy, and A. Hudák, Acoustic emissions during structural changes in shape memory alloys, Metals **9**, 58 (2019).

[10] G. Nataf, M. Guennou, J. Gregg, D. Meier, J. Hlinka, E. Salje, and J. Kreisel, Domain-wall engineering and topological defects in ferroelectric and ferroelastic materials, Nat. Rev. Phys. **2**, 634 (2020).

[11] B. Casals, G. F. Nataf, D. Pesquera, and E. K. Salje, Avalanches from charged domain wall motion in $BaTiO_3$ during ferroelectric switching, APL Mater. **8**, 011105 (2020).

[12] G. F. Nataf, P. O. Castillo-Villa, P. Sellappan, W. M. Kriven, E. Vives, A. Planes, and E. K. Salje, Predicting failure: Acoustic emission of berlinite under compression, J. Phys.: Condens. Matter **26**, 275401 (2014).

[13] E. K. Salje, A. Saxena, and A. Planes, *Avalanches in Functional Materials and Geophysics* (Springer, Berlin, 2017).

[14] H. M. Jaeger, C.-H. Liu, and S. R. Nagel, Relaxation at the angle of repose, Phys. Rev. Lett. **62**, 40 (1989).

[15] E. L. Majer and T. W. Doe, Studying hydrofractures by high frequency seismic monitoring, Int. J. Rock Mech. Mining Sci. & Geomechanics Abs. **23**, 185 (1986).

[16] B. Rouet-Leduc, C. Hulbert, N. Lubbers, K. Barros, C. J. Humphreys, and P. A. Johnson, Machine learning predicts laboratory earthquakes, Geophys. Res. Lett. **44**, 9276 (2017).

[17] X. Tong, G. Wang, J. Yi, J. Ren, S. Pauly, Y. Gao, Q. Zhai, N. Mattern, K. Dahmen, P. Liaw *et al.*, Shear avalanches in plastic deformation of a metallic glass composite, Int. J. Plast. **77**, 141 (2016).

[18] N. Friedman, A. T. Jennings, G. Tsekenis, J.-Y. Kim, M. Tao, J. T. Uhl, J. R. Greer, and K. A. Dahmen, Statistics of dislocation slip avalanches in nanosized single crystals show tuned critical behavior predicted by a simple mean field model, Phys. Rev. Lett. **109**, 095507 (2012).

[19] S. Brinckmann, J.-Y. Kim, and J. R. Greer, Fundamental differences in mechanical behavior between two types of crystals at the nanoscale, Phys. Rev. Lett. **100**, 155502 (2008).

[20] F. F. Csikor, C. Motz, D. Weygand, M. Zaiser, and S. Zapperi, Dislocation avalanches, strain bursts, and the problem of plastic forming at the micrometer scale, Science **318**, 251 (2007).

[21] D. Mordehai, S.-W. Lee, B. Backes, D. J. Srolovitz, W. D. Nix, and E. Rabkin, Size effect in compression of single-crystal gold microparticles, Acta Mater. **59**, 5202 (2011).

[22] A. Sharma, J. Hickman, N. Gazit, E. Rabkin, and Y. Mishin, Nickel nanoparticles set a new record of strength, Nat. Commun. **9**, 4102 (2018).

[23] P. D. Ispánovity, D. Ugi, G. Péterffy, M. Knapek, S. Kalácska, D. Tüzes, Z. Dankházi, K. Máthis, F. Chmelík, and I. Groma, Dislocation avalanches are like earthquakes on the micron scale, Nat. Commun. **13**, 1975 (2022).

[24] W.-Z. Han, L. Huang, S. Ogata, H. Kimizuka, Z.-C. Yang, C. Weinberger, Q.-J. Li, B.-Y. Liu, X. Zhang, J. Li, E. Ma, and Z.-W. Shan, From "smaller is stronger" to "size-independent strength plateau": Towards measuring the ideal strength of iron, Adv. Mater. **27**, 3385 (2015).

[25] I. M. Padilla Espinosa, S. Azadehranjbar, R. Ding, A. J. Baker, T. D. Jacobs, and A. Martini, Platinum nanoparticle compression: Combining in situ tem and atomistic modeling, Appl. Phys. Lett. **120**, 013101 (2022).

[26] C. Griesbach, S.-J. Jeon, D. F. Rojas, M. Ponga, S. Yazdi, S. Pathak, N. Mara, E. L. Thomas, and R. Thevamaran, Origins of size effects in initially dislocation-free single-crystal silver micro-and nanocubes, Acta Mater. **214**, 117020 (2021).

[27] E. Salje, E. Dul'Kin, and M. Roth, Acoustic Emission During the Ferroelectric Transition $Pm\bar{3}m$ to P4mm in $BaTiO_3$ and the ferroelastic transition $R\bar{3}m$-C2/c in $Pb_3(PO_4)_2$, Appl. Phys. Lett. **106**, 152903 (2015).

[28] B. Casals, G. F. Nataf, and E. K. Salje, Avalanche criticality during ferroelectric/ferroelastic switching, Nat. Commun. **12**, 345 (2021).

[29] E. Bronstein, L. Z. Tóth, L. Daróczi, D. L. Beke, R. Talmon, and D. Shilo, Tracking twin boundary jerky motion at nanometer and microsecond scales, Adv. Funct. Mater. **31**, 2106573 (2021).

[30] N. Zreihan, E. Faran, E. Vives, A. Planes, and D. Shilo, Coexistence of a well-determined kinetic law and a scale-invariant power law during the same physical process, Phys. Rev. B **97**, 014103 (2018).

[31] N. Zreihan, E. Faran, E. Vives, A. Planes, and D. Shilo, Relations between stress drops and acoustic emission measured during mechanical loading, Phys. Rev. Mater. **3**, 043603 (2019).

[32] E. Faran, E. K. Salje, and D. Shilo, The exploration of the effect of microstructure on crackling noise systems, Appl. Phys. Lett. **107**, 071902 (2015).

[33] M. C. Gallardo, J. Manchado, F. J. Romero, J. delCerro, E. K. H. Salje, A. Planes, E. Vives, R. Romero, and M. Stipcich, Avalanche criticality in the martensitic transition of $Cu_{67.64}Zn_{16.71}Al_{15.65}$ shape-memory alloy: A calorimetric and acoustic emission study, Phys. Rev. B **81**, 174102 (2010).

[34] F. Romero, J. Manchado, J. Martin-Olalla, M. Gallardo, and E. Salje, Dynamic heat flux experiments in $Cu_{67.64}Zn_{16.71}Al_{15.65}$: Separating the time scales of fast and ultra-slow kinetic processes in martensitic transformations, Appl. Phys. Lett. **99**, 011906 (2011).

[35] L. Z. Tóth, S. Szabó, L. Daróczi, and D. L. Beke, Calorimetric and acoustic emission study of martensitic transformation in single-crystalline Ni_2MnGa alloys, Phys. Rev. B **90**, 224103 (2014).

[36] L. Daróczi, S. Gyöngyösi, L. Z. Tóth, and D. L. Beke, Effect of the martensite twin structure on the deformation induced

magnetic avalanches in Ni_2MnGa single crystalline samples, Scr. Mater. **114**, 161 (2016).
[37] F. J. Pérez-Reche, E. Vives, L. Mañosa, and A. Planes, Athermal character of structural phase transitions, Phys. Rev. Lett. **87**, 195701 (2001).
[38] D. Soto-Parra, X. Zhang, S. Cao, E. Vives, E. K. Salje, and A. Planes, Avalanches in compressed Ti-Ni shape-memory porous alloys: An acoustic emission study, Phys. Rev. E **91**, 060401(R) (2015).
[39] Y. Chen, B. Gou, X. Ding, J. Sun, and E. K. Salje, Real-time monitoring dislocations, martensitic transformations and detwinning in stainless steel: Statistical analysis and machine learning, J. Mater. Sci. Technol. **92**, 31 (2021).
[40] R. Niemann, J. Kopeček, O. Heczko, J. Romberg, L. Schultz, S. Fähler, E. Vives, L. Mañosa, and A. Planes, Localizing sources of acoustic emission during the martensitic transformation, Phys. Rev. B **89**, 214118 (2014).
[41] V. Navas-Portella, Á. Corral, and E. Vives, Avalanches and force drops in displacement-driven compression of porous glasses, Phys. Rev. E **94**, 033005 (2016).
[42] E. K. H. Salje, A. Planes, and E. Vives, Analysis of crackling noise using the maximum-likelihood method: Power-law mixing and exponential damping, Phys. Rev. E **96**, 042122 (2017).
[43] E. K. Salje, D. E. Soto-Parra, A. Planes, E. Vives, M. Reinecker, and W. Schranz, Failure mechanism in porous materials under compression: Crackling noise in mesoporous SiO_2, Philos. Mag. Lett. **91**, 554 (2011).
[44] E. K. H. Salje, D. Xue, X. Ding, K. A. Dahmen, and J. Scott, Ferroelectric switching and scale invariant avalanches in $BaTiO_3$, Phys. Rev. Mater. **3**, 014415 (2019).
[45] L. Z. Tóth, E. Bronstein, L. Daróczi, D. Shilo, and D. L. Beke, Scaling of average avalanche shapes for acoustic emission during jerky motion of single twin boundary in single-crystalline Ni_2MnGa, Materials **16**, 2089 (2023).
[46] L. Z. Tóth, L. Daróczi, S. Szabó, and D. L. Beke, Simultaneous investigation of thermal, acoustic, and magnetic emission during martensitic transformation in single-crystalline Ni_2MnGa, Phys. Rev. B **93**, 144108 (2016).
[47] G. F. Nataf, P. O. Castillo-Villa, J. Baró, X. Illa, E. Vives, A. Planes, and E. K. H. Salje, Avalanches in compressed porous SiO_2-based materials, Phys. Rev. E **90**, 022405 (2014).
[48] B. Casals, K. A. Dahmen, B. Gou, S. Rooke, and E. K. Salje, The duration-energy-size enigma for acoustic emission, Sci. Rep. **11**, 5590 (2021).
[49] E. Agletdinov, D. Merson, and A. Vinogradov, A new method of low amplitude signal detection and its application in acoustic emission, Appl. Sci. **10**, 73 (2019).
[50] E. Pomponi, A. Vinogradov, and A. Danyuk, Wavelet based approach to signal activity detection and phase picking: Application to acoustic emission, Signal Process. **115**, 110 (2015).
[51] P. J. Schmid, Dynamic mode decomposition of numerical and experimental data, J. Fluid Mech. **656**, 5 (2010).
[52] P. J. Schmid and J. L. Sesterhenn, Dynamic mode decomposition of experimental data, in *8th International Symposium on Particle Image Velocimetry, Melbourne, Victoria, Australia, Bulletin of the American Physical Society, 61st APS meeting, San Antonio* (American Physical Society, 2008), p. 208.
[53] E. Bronstein, A. Wiegner, D. Shilo, and R. Talmon, The spatiotemporal coupling in delay-coordinates dynamic mode decomposition, Chaos **32**, 123127 (2022).
[54] C. W. Rowley, I. Mezić, S. Bagheri, P. Schlatter, and D. S. Henningson, Spectral analysis of nonlinear flows, J. Fluid Mech. **641**, 115 (2009).
[55] J. H. Tu, C. W. Rowley, D. M. Luchtenburg, S. L. Brunton, and J. N. Kutz, On dynamic mode decomposition: Theory and applications, J. Comput. Dyn. **1**, 391 (2014).
[56] M. O. Williams, I. G. Kevrekidis, and C. W. Rowley, A data–driven approximation of the koopman operator: Extending dynamic mode decomposition, J. Nonlinear Sci. **25**, 1307 (2015).
[57] M. Budišić, R. Mohr, and I. Mezić, Applied Koopmanism, Chaos **22**, 047510 (2012).
[58] I. Mezić, Spectral properties of dynamical systems, model reduction and decompositions, Nonlinear Dyn. **41**, 309 (2005).
[59] I. Mezić, Analysis of fluid flows via spectral properties of the Koopman operator, Annu. Rev. Fluid Mech. **45**, 357 (2013).
[60] J. N. Kutz, S. L. Brunton, B. W. Brunton, and J. L. Proctor, *Dynamic Mode Decomposition: Data-Driven Modeling of Complex Systems* (SIAM, 2016).
[61] W. E. Boyce, R. C. DiPrima, and D. B. Meade, *Elementary Differential Equations and Boundary Value Problems* (John Wiley & Sons, 2021).
[62] P. J. Schmid, Application of the dynamic mode decomposition to experimental data, Exp. Fluids **50**, 1123 (2011).
[63] B. W. Brunton, L. A. Johnson, J. G. Ojemann, and J. N. Kutz, Extracting spatial–temporal coherent patterns in large-scale neural recordings using dynamic mode decomposition, J. Neurosci. Methods **258**, 1 (2016).
[64] S. L. Brunton, B. W. Brunton, J. L. Proctor, E. Kaiser, and J. N. Kutz, Chaos as an intermittently forced linear system, Nat. Commun. **8**, 19 (2017).
[65] S. L. Brunton and J. N. Kutz, Methods for data-driven multiscale model discovery for materials, J. Phys. Mater. **2**, 044002 (2019).
[66] D. Duke, J. Soria, and D. Honnery, An error analysis of the dynamic mode decomposition, Exp. Fluids **52**, 529 (2012).
[67] T. W. Muld, G. Efraimsson, and D. S. Henningson, Flow structures around a high-speed train extracted using proper orthogonal decomposition and dynamic mode decomposition, Comput. Fluids **57**, 87 (2012).
[68] S. Le Clainche and J. M. Vega, Higher order dynamic mode decomposition, SIAM J. Appl. Dyn. Syst. **16**, 882 (2017).
[69] M. R. Jovanović, P. J. Schmid, and J. W. Nichols, Sparsity-promoting dynamic mode decomposition, Phys. Fluids **26**, 024103 (2014).
[70] S. Pan and K. Duraisamy, On the structure of time-delay embedding in linear models of non-linear dynamical systems, Chaos **30**, 073135 (2020).
[71] C. T. Elmore and A. W. Dowling, Learning spatiotemporal dynamics in wholesale energy markets with dynamic mode decomposition, Energy **232**, 121013 (2021).
[72] F. Takens, Detecting strange attractors in turbulence, in *Dynamical Systems and Turbulence, Lecture Notes in Mathematics* (Springer, Berlin, 1981).
[73] J. Zimmerman, A. Bisht, Y. Mishin, and E. Rabkin, Size and shape effects on the strength of platinum nanoparticles, J. Mater. Sci. **56**, 18300 (2021).

[74] A. Bisht, R. K. Koju, Y. Qi, J. Hickman, Y. Mishin, and E. Rabkin, The impact of alloying on defect-free nanoparticles exhibiting softer but tougher behavior, Nat. Commun. **12**, 2515 (2021).

[75] P. Herre, J. Will, M. Dierner, D. Wang, T. Yokosawa, T. Zech, M. Wu, T. Przybilla, S. Romeis, T. Unruh *et al.*, Rapid fabrication and interface structure of highly faceted epitaxial Ni-Au solid solution nanoparticles on sapphire, Acta Mater. **220**, 117318 (2021).

[76] A. Sharma, R. Kositski, O. Kovalenko, D. Mordehai, and E. Rabkin, Giant shape-and size-dependent compressive strength of molybdenum nano-and microparticles, Acta Mater. **198**, 72 (2020).

[77] A. Sharma, N. Gazit, L. Klinger, and E. Rabkin, Pseudoelasticity of metal nanoparticles is caused by their ultrahigh strength, Adv. Funct. Mater. **30**, 1807554 (2020).

[78] K. Pearson, Liii. on lines and planes of closest fit to systems of points in space, London Edinburgh Dublin Philos. Mag. J. Sci. **2**, 559 (1901).

[79] H. Hotelling, Analysis of a complex of statistical variables into principal components. J. Educ. Psychol. **24**, 417 (1933).

[80] I. T. Jolliffe and J. Cadima, Principal component analysis: A review and recent developments, Philos. Trans. R. Soc. A **374**, 20150202 (2016).

[81] J. Cohen, P. Cohen, S. G. West, and L. S. Aiken, *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences* (Routledge, 2013).

[82] A. Tharwat, Linear vs. quadratic discriminant analysis classifier: A tutorial, Int. J. Appl. Pattern Recognit. **3**, 145 (2016).

[83] R. R. Coifman and S. Lafon, Diffusion maps, Appl. Comput. Harmon. Anal. **21**, 5 (2006).

[84] R. R. Lederman and R. Talmon, Learning the geometry of common latent variables using alternating-diffusion, Appl. Comput. Harmon. Anal. **44**, 509 (2018).

[85] D. W. Hosmer Jr, S. Lemeshow, and R. X. Sturdivant, *Applied Logistic Regression*, Wiley Series in Probability and Statistics, Vol. 398 (John Wiley & Sons, 2013).

[86] E. Fix, *Discriminatory Analysis: Nonparametric Discrimination, Consistency Properties*, vol. 1 (USAF School of Aviation Medicine, 1985).

[87] T. Cover and P. Hart, Nearest neighbor pattern classification, IEEE Trans. Inf. Theory **13**, 21 (1967).

[88] V. N. Vapnik and A. Y. Chervonenkis, On the uniform convergence of relative frequencies of events to their probabilities, in *Measures of Complexity*, edited by V. Vovk, H. Papadopoulos, and A. Gammerman (Springer, Cham, 2015).

[89] M. Awad and R. Khanna, Support vector machines for classification, in *Efficient Learning Machines* (Apress, Berkeley, CA, 2015).