

Nanomechanical mass measurements through feature-based time series clustering

Cite as: Rev. Sci. Instrum. 95, 025001 (2024); doi: 10.1063/5.0176303

Submitted: 12 September 2023 • Accepted: 9 December 2023 •

Published Online: 6 February 2024



View Online



Export Citation



CrossMark

Adam P. Neumann,¹  Alfredo Gomez,¹  Alexander R. Nunn,¹ John E. Sader,^{2,a)} 
and Michael L. Roukes^{1,a)} 

AFFILIATIONS

¹ Department of Physics, California Institute of Technology, Pasadena, California 91125, USA

² Graduate Aerospace Laboratories and Department of Applied Physics, California Institute of Technology, Pasadena, California 91125, USA

^{a)} Authors to whom correspondence should be addressed: jsader@caltech.edu and roukes@caltech.edu

ABSTRACT

Recent years have seen explosive growth in miniaturized sensors that can continuously monitor a wide variety of processes, with applications in healthcare, manufacturing, and environmental sensing. The time series generated by these sensors often involves abrupt jumps in the detected signal. One such application uses nanoelectromechanical systems (NEMS) for mass spectrometry, where analyte adsorption produces a quick but finite-time jump in the resonance frequencies of the sensor eigenmodes. This finite-time response can lead to ambiguity in the detection of adsorption events, particularly in high event-rate mass adsorption. Here, we develop a computational algorithm that robustly eliminates this often-encountered ambiguity. A moving-window statistical test together with a feature-based clustering algorithm is proposed to automate the identification of single-event jumps. We validate the method using numerical simulations and demonstrate its application in practice using time-series data that are experimentally generated by molecules adsorbing onto NEMS sensors at a high event rate. This computational algorithm enables new applications, including high-throughput, single-molecule proteomics.

Published under an exclusive license by AIP Publishing. <https://doi.org/10.1063/5.0176303>

I. INTRODUCTION

In nanoelectromechanical systems (NEMS)-based mass spectrometry (MS), individual molecules adsorbing to a mechanical resonator induce a change in the resonance frequencies of the device. The fractional eigenfrequency shifts, Δf_n , of mode $n = 1, 2, \dots$, are related to the adsorbed mass, m , and its position, x , by¹⁻⁵

$$\Delta f_n = -\frac{m}{2M_{\text{device}}} \Phi_n^2(x), \quad (1)$$

where M_{device} is the device mass and $\Phi_n(x)$ is the corresponding eigenmode. The masses of separate analytes can then be measured experimentally by detecting individual adsorption events and quantifying the associated shifts in the resonant frequencies of the device.

The accurate measurement of frequency shifts due to individual adsorption events is complicated by several processes. In practice, the resonance frequency of each eigenmode of the device is tracked using a phase-locked loop (PLL). This is a control loop that

maintains stability with respect to high-frequency noise fluctuations while simultaneously achieving a controlled response to large jumps in frequency. A PLL optimized for these criteria creates finite-time transients when an adsorption event occurs. When multiple analytes adsorb within a narrow time interval, succeeding frequency shifts interrupt the finite-time transients of prior events and create ambiguity in the frequency shift due to the adsorption of a single analyte. Further complicating the problem, the multimodal time-series can have a low signal to noise ratio, drift, and noise processes correlated over long time intervals, which invalidates the assumptions of many conventional statistical approaches.

Prior studies for detecting adsorption events in NEMS-MS have fit the inter-sample difference of the multimode time-series to a multivariate Gaussian distribution. Adsorption events are then identified as outliers that deviate from the mean by a user-defined threshold.^{3,6} The disadvantage of this approach is that it fails to differentiate between frequency shifts due to single or multiple adsorption events.⁷ As we shall demonstrate, this can significantly distort the mass distribution predicted by NEMS-MS.

A robust detection methodology is required that can accurately measure the change in resonance frequencies and distinguish between single and multiple mass adsorption events occurring in proximity. This is crucial since accurate frequency shift measurements can only be taken from single-event jumps. Indeed, such a scheme is a necessity for scaling NEMS-MS to high-throughput applications where so called, *multi-event jumps* become frequent. It is also highly desirable that this detection method performs in near real time,⁸ which would enable the analysis and selective response to individual molecular adsorption events. This would allow NEMS-MS to act as a pre-selector for further downstream measurements.

Here, we present a modular, noise-driven approach for detecting single adsorption events within noisy multimodal NEMS-MS time-series data. The approach consists of three stages: jump detection, classification, and mean frequency shift measurement. In the detection phase, the algorithm identifies abrupt shifts in the time series. In the subsequent classification stage, time-domain features of the events are extracted, and events are categorized as either single-event jumps, caused by a single mass absorbent, or multi-event jumps, caused by multiple mass absorbents with overlapping transients. The final measurement phase retains only single-event jumps and measures the change in mean resonance frequency before and after the event. This multi-step process enables the robust measurement of single-event frequency shifts and excludes ambiguous multi-event jumps.

The study is organized as follows: In Sec. II, the methodology and theory of our method are presented, and in Sec. III, adsorption events are numerically simulated on a NEMS device to produce a multimodal time-series with known adsorption times. This is used to evaluate the performance of the method as the level of noise increases. In Sec. IV, we demonstrate the application of the method to experimental data obtained from a hybrid Orbitrap-NEMS system, reported elsewhere.^{7,9} The frequency-shift response due to single macromolecule absorption events is extracted from noisy multimodal time-series data, and the mono-disperse mass distribution is recovered. A discussion of results is given in Sec. V, with the current Python implementation of the method provided in the Appendix.

II. METHODOLOGY

In this section, we outline the proposed methodology for processing NEMS-MS time-series data. Consisting of jump identification, single-event classification, and output phases, the approach is modular, allowing practitioners to interchange methods at each step as needed. A diagram of the approach is shown in Fig. 1.

To detect shifts in noisy time-series data during the detection phase, we utilize methods derived from change point analysis.^{10–12} These methods can be broadly categorized as either exact or approximate, with the former using the entire time-series (offline) and the latter providing an estimate based on the observed history (online). To analyze the time-series data of our problem, which is embedded in non-stationary and long-time correlated noise processes, we employ a window-based approximate change point method. These methods are robust regardless of the number or position of the frequency shifts, making them well-suited for NEMS-MS analysis in

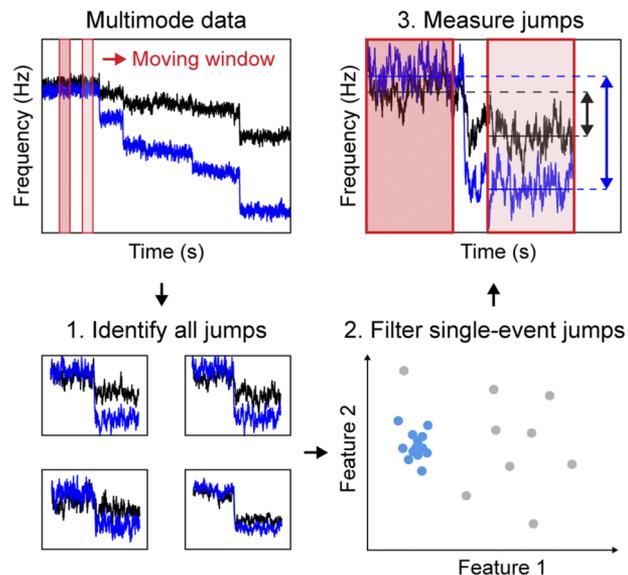


FIG. 1. Event measurement scheme for multimode time series data. A sliding window approach with a built-in gap for the response time of the sensor is used to calculate a two-sample statistic. Jumps are detected when the statistic exceeds a statistical significance threshold. Features of the time series surrounding each jump are then extracted, and jumps are filtered to single-event jumps using a clustering algorithm based on similarity to the central tendency of the extracted features. Finally, the jump magnitudes of the remaining jumps are measured and reported.

real time. The statistical test used to identify shifts between rolling sample windows is described in Sec. II A.

In the classification step, we employ techniques from *time-series clustering* to group detected events based on similarity.^{13,14} Time-series clustering methods can be broadly divided into raw data-based, model-based, or feature-based methods. Raw data-based methods include every sample of an event in the classification. This approach has the advantage of making minimal assumptions about the characteristics of events but performs poorly on datasets with high noise and a small sample size. In contrast, model-based approaches make strong assumptions about the characteristics of events and fit each event to a model. However, since the characteristics of events in NEMS-MS cannot be determined precisely, we consider feature-based time-series clustering. By extracting several key features related to the shape of the event peak in the moving window statistic, such as the full-width at half maximum (FWHM), this approach has the advantages of being robust to noise, significantly reducing the dimensionality, and thereby complexity, of the classification problem, and including minimal assumptions about event behavior. We describe feature-based time-series clustering in Sec. II B. By performing this clustering, we demonstrate that multi-event jumps can be identified automatically by their deviations away from the central single-event jump tendency of the dataset in feature space.

We conclude by describing a consistent approach for measuring the mean frequency shift due to a single-event jump in Sec. II C.

A. Step 1—Identify all frequency jumps

Consider two samples, X and Y , representing the multimode frequency time series before and after each event. Samples X and Y are taken using a simple moving window approach with a time gap, t_{jump} , between the samples to exclude the jump transient. For NEMS devices, t_{jump} can be initially estimated from the PLL time chosen for the experiment, but as will be shown later, it can be fine-tuned after a relatively insensitive initial guess. Previous applications of jump detection in NEMS-MS utilized the student's t -test to automatically find frequency jumps.⁷ The embodiment in this article applies Hotelling's two-sample t^2 statistic, a multivariate generalization of the t statistic,¹⁵ which is given as

$$t^2 = \frac{n_x n_y}{n_x + n_y} (\bar{X} - \bar{Y})^T \Sigma^{-1} (\bar{X} - \bar{Y}), \quad (2)$$

where n_x and n_y represent the number of degrees of freedom in X and Y , \bar{X} and \bar{Y} represent the vector of the column mean for X and Y , and Σ is the pooled covariance matrix.¹⁶ In the case where $n_x = n_y$, Σ is the mean of Σ_X and Σ_Y (the covariance matrices for X and Y , respectively). Equation (2) can be related to an F -distribution, which has more widespread support in standard software packages, with a simple multiplicative factor required for conversion,¹⁷

$$\frac{n_x + n_y - N - 1}{N(n_x + n_y - 2)} t^2 \sim F(N, n_x + n_y - N - 1), \quad (3)$$

where N represents the number of dimensions, in this case, modes.

If each data point is drawn independently from two independent multivariate normal distributions, n_x and n_y also represent the number of data points in X and Y . With NEMS experimental data, however, there is a high degree of autocorrelation in the time series due to the non-white nature of the noise as well as the dynamical nature of the PLL, which incorporates an integrator, low-pass filter, and feedback. For this reason, the value of the F -statistic calculated in Eq. (3) with n_x and n_y set to the number of data points in X and Y is not comparable to that obtained from standard models and cannot directly be converted to a confidence interval or p -value.

To relate the F -statistic calculated in Eq. (3) to a p -value, we employed a non-parametric bootstrapping method known as

moving block bootstrapping^{18–20} on an event-free (i.e., pure noise) dataset. This is always possible to obtain with NEMS devices. If such a dataset is not available, it might be possible to approximate one by running through the jump detection framework with an initial pass, excluding data near all suspected jumps, and appending the remaining time series. Briefly, X and Y are sampled randomly with replacement, with the resulting F -statistics sorted and plotted alongside the corresponding fraction of the collected samples. The pure noise dataset acquired with the same device used to collect the experimental data is provided along with the bootstrapping results in Fig. 2. In addition, shown in Fig. 2 are contour plots in relative frequency space corresponding to 1-, 2-, and 3-sigma significance. Further details about the experimental data and device are provided in Sec. IV. These contour plots offer an alternative representation of the minimum frequency fluctuation detectable by the device, which is more typically calculated using a single-mode Allan deviation; this is a measure of frequency stability involving the standard deviation of the differences of time-averaged fractional frequencies.²¹ Importantly, the gap between samples required for the jump is automatically incorporated into this bootstrapping calculation, in contrast with the Allan deviation, which requires careful modification of the standard formula. The window length (the size of X and Y) can be fixed by choosing the value that minimizes the volume of these frequency fluctuation contours in N -dimensional space.

Once bootstrapping has been performed, jumps can be detected when the F -statistic crosses above a threshold related to the p -value. We chose $p = 0.003$, which is approximately equivalent to a 3-sigma deviation. It should be emphasized that both the jump threshold and window length are directly informed by our application and data, which is not always the case and is typically cited as a weakness of moving-window based change point detection.

Following event detection, an additional processing step called *filtering* is performed to eliminate some erroneous events. These filtering steps are manually chosen to select events with clear outlier behavior and include (1) eliminating jumps with too little detected time spent above the threshold—these are jumps that are presumed to be noise fluctuations, (2) removing identified events that are near one another, and (3) jumps with a positive frequency shift. The latter

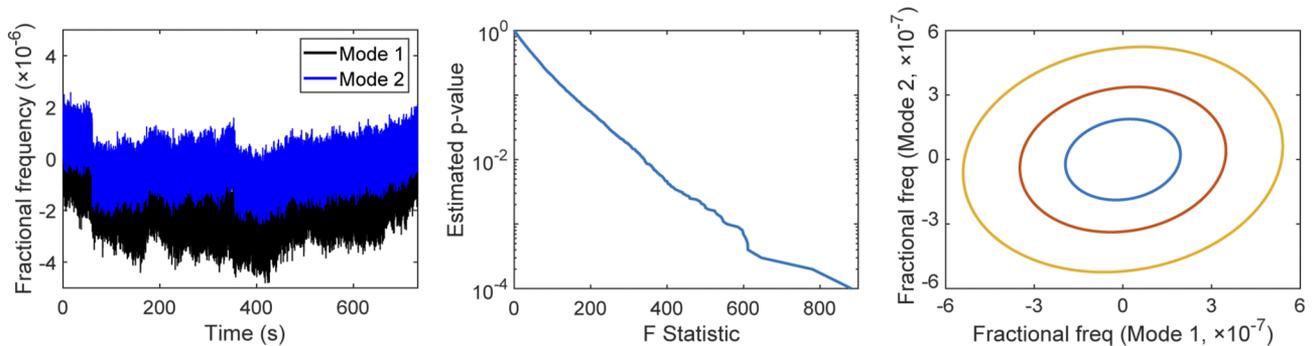


FIG. 2. Noise characterization and significance bootstrapping. Fractional frequency noise is shown for experimental data acquired with no events. With these data, moving block bootstrapping is used to empirically determine the significance of any given window sample based on its F -statistic, which is used to select a detection threshold.

is specific to our application of NEMS-based mass sensing, associated with the constraint that we are only interested in detecting physical events with positive added mass.

B. Step 2—Reduce to single-event jumps

With the observation that the F-statistic vs time exhibits different dynamics based on the presence (or not) of multiple jumps in proximity, we seek to develop statistical features that reduce the dimensionality in a manner that is robust to noise and not sensitive to the knowledge of the exact jump time. In this way, we aim to develop inputs for a feature-based clustering algorithm that could automate the detection of single-event jumps. The general approach considered here involves the expansion of the F-statistic vs time, $F(t)$, near each detected jump in terms of statistical moments, treating it as a probability distribution. The first five moments, normalized to give magnitude and width independence, are

$$\begin{aligned} F_0 &= \int_{t_0}^{t_1} F(t) dt, & F_1 &= \int_{t_0}^{t_1} tF(t) dt / F_0, \\ F_2 &= \int_{t_0}^{t_1} (t - F_1)^2 F(t) dt / F_0, \\ F_3 &= \int_{t_0}^{t_1} (t - F_1)^3 F(t) dt / (F_0 F_2^{3/2}), \\ F_4 &= \int_{t_0}^{t_1} (t - F_1)^4 F(t) dt / (F_0 F_2^2), \end{aligned} \quad (4)$$

where t_0 and t_1 are the start and end of the detected jump (as determined by the times it crosses above and below the detection threshold), respectively, and the integrals are calculated numerically. The first two features represent the time average of $F(t)$ relative to the detected jump time and the standard deviation. The next two features represent the skewness and kurtosis, i.e., the third and fourth central moments normalized by the standard deviation (to create new independent features), respectively. Finally, FWHM represents the fifth feature. FWHM, as introduced in Sec. II, is observed to be highly sensitive to the sharper peaks with nearly instantaneous jumps—those with an event beginning before a prior event ends. In the proposed method, we used the five quantities $F_1 - t_{jump}$, $\sqrt{F_2}$, F_3 , F_4 , and FWHM.

We now investigate the use of clustering algorithms on the extracted statistical features to isolate single-event jumps in a process. This takes advantage of the observed central tendency of the single-event jump in a manner that minimizes human bias. To target this spatial relationship, we use DBSCAN (Density Based Spatial Clustering of Applications with Noise), which employs a density-based selection of clusters of arbitrary shapes with minimal domain knowledge.¹⁴ Moreover, DBSCAN is an unsupervised algorithm, meaning that exact knowledge about the behavior of single- and double-jump events is not a necessity, which is important when analyzing experimental data. DBSCAN identifies both clusters and noise (points not belonging to any cluster) for data using two parameters: N_{min} and ϵ , which denote the minimum number of points required to form a cluster and neighborhood radius about each point, respectively. The parameter N_{min} is typically set to twice the number of dimensions—ten for five-dimensional data—while the

neighborhood radius, ϵ , is the freely adjustable parameter. To give each feature equal weight when using Euclidean distance as the distance metric, features are normalized such that 50% of the data would lie within the range -1 to $+1$.

The original formulation of DBSCAN is accompanied by a heuristic for choosing a value for the neighborhood radius, ϵ , that separates the noise from the clusters.¹⁴ This involves estimating the percent of noise and selecting a threshold near this estimate based on the visual behavior of the data sorted according to the distance to the k -th nearest point (where k is equal to N_{min}). In our application, there is no clear opportunity for making a specific choice of ϵ based on this heuristic; this might be because the noise can be arbitrarily close to the data with no clear transition, and the low overall SNR for some datasets. For this reason, we find that estimating the percent of noise and using this choice to determine ϵ is the simplest and most consistent method.

For our dataset, “noise” consists of (1) fluctuations misclassified as events (false positives) and (2) multiple-jump events. The first type of noise depends on the choice of the F-statistic threshold used to detect events and can be estimated by finding the number of events detected on an event-free (pure noise) dataset. For NEMS experiments, it is always possible to acquire these data separately. Simulations in Sec. III A show this is a relatively small contribution. The second type of noise is estimated for our application by assuming the experimental events occur at a constant rate and follow a Poisson process. The number of single-event jumps is estimated by finding the event rate from the total number of detected events, then calculating the probability that no events will occur within a jump window Δt_{jump} and two measurement windows t_{meas} (to account for the measurement before and after a given jump). This follows from the Poisson process because events occur independently of each other. Thus, estimating the probability of events within a time period, $2t_{meas} + \Delta t_{jump}$, is equivalent to estimating the probability of just one event having occurred in that time period, conditional upon a jump having already been detected with jump time, t_{jump} .

With observations on the simulated datasets, we adapt DBSCAN to only consider the largest discovered cluster based on the above noise estimations. Simulations show there is a trade-off between accepting multiple-jump events and removing low-magnitude events.

C. Step 3—Output frequency shifts of single-event jumps

The optimal measurement of the jump magnitude, $\bar{Y} - \bar{X}$, requires positioning the measurement windows as close as possible (to reduce the impact of noise and drift) while ensuring that the jump transient is excluded. For datasets with no time-dependent response, i.e., those with instantaneous jumps, the jump time is simply calculated as the peak in the moving-window statistic¹⁰ described in Sec. II A. For this application with a finite-time response, a custom peak-finding routine is developed to establish a precise and consistent jump time, as follows: First, we calculate the peak of the F-statistic, F_{peak} , near each jump, then locate the time in which the F-statistic reaches half that value, working backward in time from the end of the jump time (the time at which the F-statistic crosses below the detection threshold). This algorithm consistently places this jump time, t_{jump} , slightly to the right (i.e., ahead in time) of

the final peak, even for low-magnitude jumps with a flat peak and situations with multiple jumps back-to-back.

Next, the jump window is chosen with a specific jump window, Δt_{jump} , and time offset, t_{offset} , relative to the jump time, t_{jump} , for the jump window to be precisely adjusted to start and end relative to the jump time in a way that fully encompasses the jump dynamics. While this can be done using an initial guess based on the known PLL time, it can be further fine-tuned by adjusting the Δt_{jump} and t_{offset} of the measurement. For validating choices of Δt_{jump} and t_{offset} , an additional validation method is developed through the notion of a jump signature. The time series surrounding each event is collected, aligned at the jump time, and normalized such that the jump decreases from 1 to 0, then the median of such data is obtained and plotted along with the jump windows. The variables Δt_{jump} and t_{offset} are then adjusted, and the jump signature is recalculated.

III. NUMERICAL SIMULATIONS OF THE PROTOCOL

To validate our detection and classification methodology, we numerically simulate mass absorption events on a NEMS device to produce noisy multimodal time-series data with known event times. Mirroring the experimental setup in Sec. IV, COMSOL simulations are performed to calculate the frequency shifts induced by adsorbing analytes. The analytes are modeled as half-sphere primitives and are positioned randomly over the surface of the device, with the mass of the analyte and device chosen to reflect the experiment. To calculate the precise frequency shift, a stationary eigenfrequency analysis is conducted with the half-sphere primitive first set at zero density and then at the final density using the same mesh. Multimodal time-series are then created by generating mass absorption events at random times following a Poisson distribution with an event rate of 2 events/s. An exponential transient decay (time constant of 10 ms) is introduced to simulate the transient behavior of the phase lock loop (PLL). This noise free time-series with known event times is then embedded in experimental noise collected from the device described in Sec. IV when no molecules are sent to the device.

Different noise levels compared with the experiment are generated by altering the relative magnitude of the particle mass compared with the molecule used in the experiment. Simulations are conducted at $1\times$, $10\times$, and $100\times$ scaling of the original signal-to-noise ratio (SNR). Larger scaling datasets are used for clarity with the initial algorithm development and validation, and the $1\times$ SNR dataset is used to assess the performance of the experimental dataset. We now implement the three steps for single event jump detection detailed in Sec. II.

A. Step 1—Identify all frequency jumps

The simulated time-series provides insight into how the proximity of events affects the behavior of the F-statistic over time. As shown in Fig. 3, events that are well-separated in time produce distinct peaks in the F-statistic above the detection threshold. Conversely, when events occur nearly simultaneously, with one interrupting the transient phase of the other, the peaks of the F-statistic merge into a single peak. Confoundingly, this merged peak often appears narrower than the peaks resulting from individual events when measured by the full-width at half-maximum (FWHM). Another complication arises because the shape of the F-statistic can vary depending on the magnitude of the frequency shift. Small frequency shifts lead to broad, flat peaks in the F-statistic over time, while larger frequency shifts produce sharper, skewed peaks. As we later show, variation in the shape of the F-statistic peak can be used to distinguish single and multiple adsorption events and noise fluctuations.

To quantify the performance of the methodology in both the detection and classification phases, we treat our task as a binary classification problem. Commonly used in change point detection,⁸ this paradigm measures the performance of the detection phase according to how accurately detection windows are classified as containing an event or no event. A confusion matrix for the classification problem is shown in Table I.

The performance of a binary classification problem can be summarized using the metrics of precision and recall,

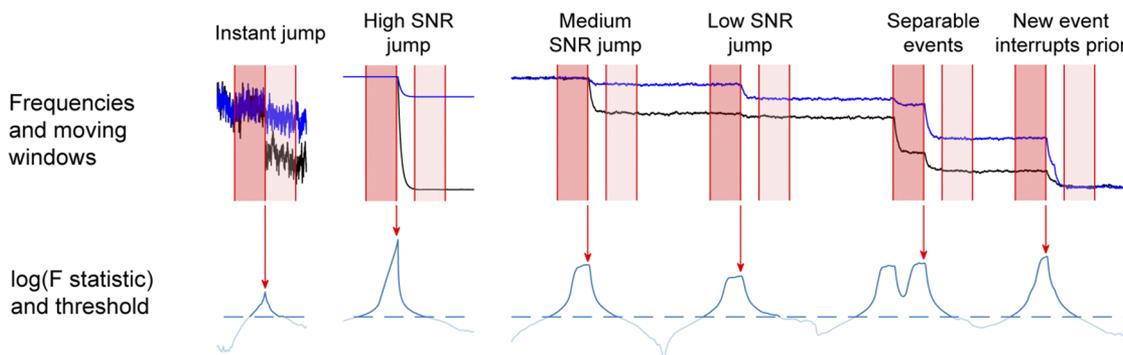


FIG. 3. Simulated events with a time-varying F-statistic. COMSOL was used to simulate frequency shifts from adsorbed particles obtained from the GroEL NEMS-MS measurements, and these frequency shifts were embedded in experimental noise. From left, a single-event instant jump at $1\times$ signal-to-noise (SNR), followed by jumps with an exponential decay of 10 ms: first, a single-event jump at $100\times$ SNR, followed by prototypical jump types at $10\times$ SNR: a single-event jump of medium and small magnitude, two adsorption events close together with isolated peaks in the F-statistic, then two events with the second event beginning before the first ends. Jump times are indicated with a red arrow, and the location of the measurement windows in relation to the jump time is shown with red rectangles; the procedure for determining both is discussed in the text.

TABLE I. Event Detection—Confusion Matrix. Categorization of event detection outcomes.

Event category	Event detected	No. event detected
$N_{events} \geq 1$	True positive (TP)	False negative (FN)
$N_{events} = 0$	False positive (FP)	True negative (TN)

TABLE II. Detection Evaluation. The F-score is calculated with $\beta = 0.5$ to reflect the preference for precision over recall.

Noise level	True positive	False positive	False negative	Precision	Recall	F-score
$1 \times \text{SNR}$	667	16	1333	0.98	0.33	0.70
$10 \times \text{SNR}$	727	15	1273	0.98	0.36	0.73
$100 \times \text{SNR}$	758	19	1242	0.98	0.38	0.74

$$precision = \frac{TP}{TP + FP}, \quad recall = \frac{TP}{TP + FN}. \quad (5)$$

Precision is the fraction of detected events that are true events, while *recall* is the fraction of events in the dataset detected. In our

case, correctly identifying events (high precision) is preferable to ensuring that all events are detected (high recall). A useful metric for evaluating the classification, given this preference, is

$$F\text{-score} = (1 + \beta^2) \frac{precision \cdot recall}{\beta^2 \cdot precision + recall}, \quad (6)$$

which attaches β -times as much importance to *recall* compared with *precision*. For the remainder of this study, we consider F-scores with $\beta = 1/2$ to favor precision over recall.

In **Table II**, we quantify the performance of the F-statistic threshold in the detection phase for increasing noise levels using these metrics. Across all noise levels, the precision of the detection method is very high, indicating that only a small portion of detected events are due to random noise fluctuations. The low recall for all noise levels is likely due to the range of different frequency-shift magnitudes generated by sampling absorption events across the full surface of the NEMS device. As the absorption position of the analyte approaches the clamped edge of the NEMS device, the frequency-shift decreases in magnitude and falls beneath the detection threshold. This results in many undetected events and, hence, a low recall.

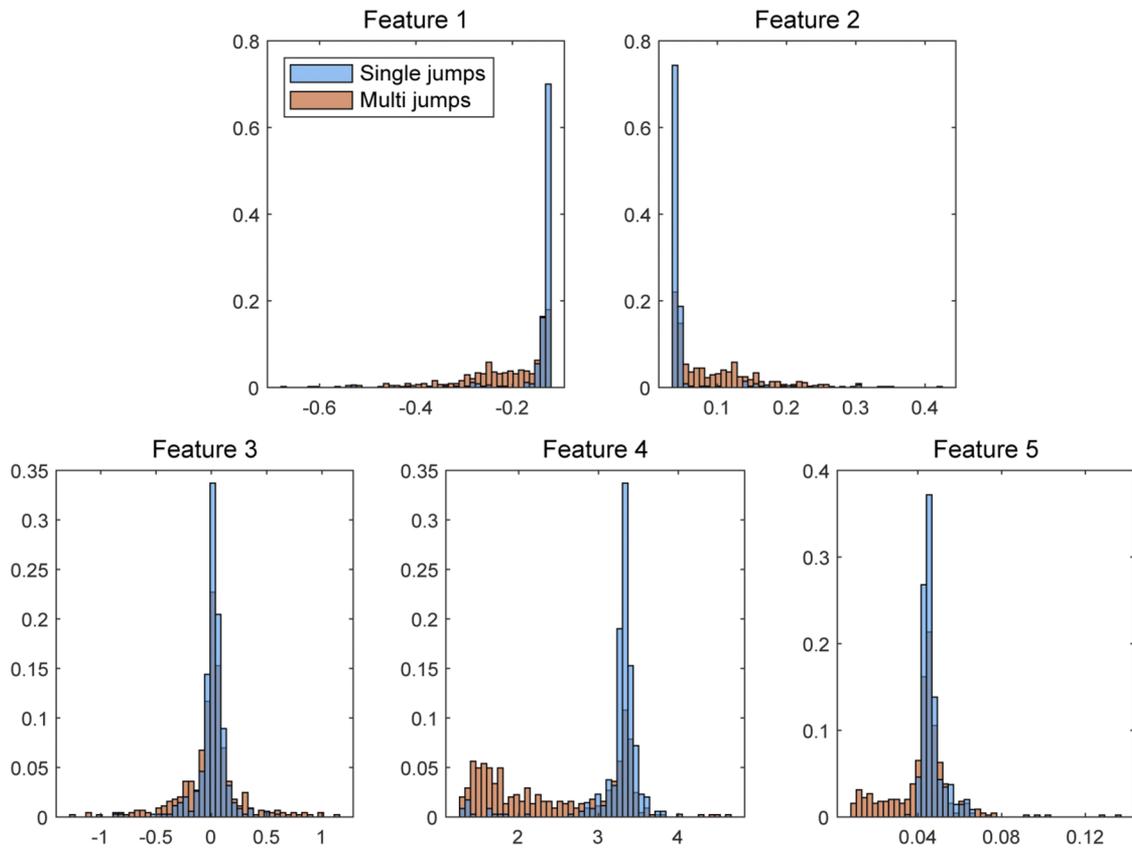


FIG. 4. F-statistic peak features. Detected events for a simulated dataset ($10 \times \text{SNR}$) are classified as either single- or multiple-jump occurrences based on the known jump times input for the simulation.

TABLE III. Classification Phase—Confusion Matrix. Categorization of clustering outcomes (i.e., classifying events as single- or multiple-event jumps after they have already been detected) by considering it as a binary classification problem.

Classification	Event in cluster	Event not in cluster
Single-event jump	True positive (TP)	False negative (FN)
Not single-event jump	False positive (FP)	True negative (TN)

TABLE IV. Classification precision for different feature combinations. ($1\times$ SNR) Clustering results indicate improved single-event identification for more restrictive choices of ϵ . The choice of features generally does not significantly impact precision results. Similar observations are drawn with $10\times$ SNR and $100\times$ SNR.

Feature 1	Feature 2	100% data (No. clustering)	60% data (Looser clustering)	30% data (Tighter clustering)
Moment 1	Moment 2	0.32	0.45	0.57
Moment 1	Moment 3	0.32	0.44	0.53
Moment 1	Moment 4	0.32	0.46	0.54
Moment 1	FWHM	0.32	0.43	0.57
Moment 2	Moment 3	0.32	0.46	0.53
Moment 2	Moment 4	0.32	0.46	0.56
Moment 2	FWHM	0.32	0.43	0.55
Moment 3	Moment 4	0.32	0.44	0.53
Moment 3	FWHM	0.32	0.43	0.49
Moment 4	FWHM	0.32	0.46	0.54

B. Step 2—Reduce to single-event jumps

For the classification phase of the methodology, we now examine the effects of using different features and clustering parameters on the simulated multimodal time-series.

For each F-statistic peak identified in the detection phase, the time-domain features given in Eq. (4) are extracted. A histogram of the five features chosen for clustering is shown in Fig. 4, with the data colored according to whether one or several adsorption events occur in the window of interest. These histograms illustrate that the so-called single and multiple jump events are distributed differently across all features. Single-event jumps are tightly clustered around a single value, while multi-jump events exhibit significantly larger variations. This observation motivates our working principle that events with feature values deviating significantly from the distribution mode are less likely to be associated with single events. However, the overlapping feature distribution between single and multiple jump events poses a challenge for precise classification. This pattern is consistently observed in both the $1\times$ SNR and $10\times$ SNR simulated datasets.

To quantify the performance of the classification phase and the effectiveness of the methodology overall, we consider the binary classification problem of identifying single-event jumps only. The confusion matrix is shown in Table III.

To analyze the effects of using different feature combinations and clustering hyperparameter values, we consider the precision of single event classification for a variety of scenarios. Table IV reports the precision for different pairs of features with increasing cluster tightness on the $1\times$ SNR time-series. The classification has a similar precision regardless of the combination of features,

TABLE V. Effect of adding features on precision. Reported is the median precision for all combinations for each choice of between 2 and 5 features for the three simulated datasets, $1-100\times$ SNR.

Dataset	$1\times$ SNR				$10\times$ SNR				$100\times$ SNR			
	2	3	4	5	2	3	4	5	2	3	4	5
No clustering	0.32	0.32	0.32	0.32	0.38	0.38	0.37	0.38	0.41	0.41	0.41	0.41
Looser clustering	0.45	0.47	0.51	0.52	0.52	0.51	0.70	0.70	0.53	0.55	0.72	0.78
Tighter clustering	0.56	0.56	0.56	0.56	0.69	0.70	0.71	0.72	0.77	0.77	0.78	0.80

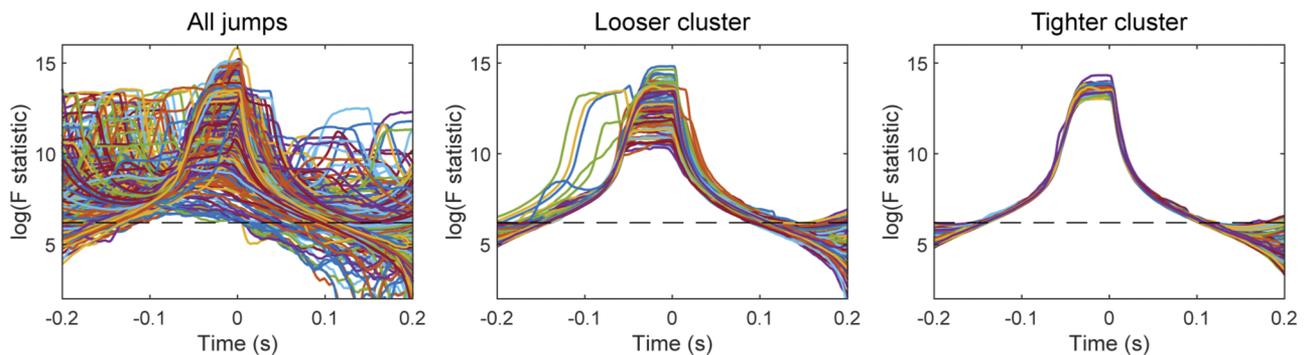


FIG. 5. All detected jumps, before and after clustering, for a simulated dataset. F-statistic time series surrounding each jump detected in the $10\times$ SNR simulated dataset plotted relative to their detected jump time t_{jump} shown with the looser and tighter clustering as indicated in Fig. 6.

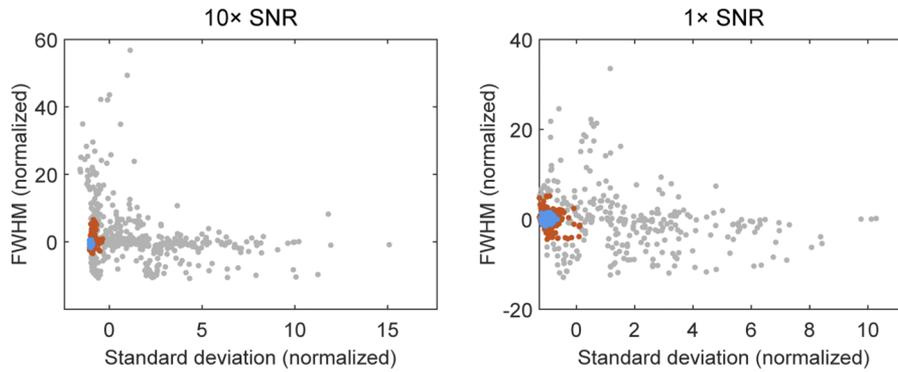


FIG. 6. Clustering algorithm for identifying single-event jumps. DBSCAN applied to the extracted features of the F-statistic vs time curve excluded a significant proportion of noise fluctuations and multiple-jump events. The tightness of the DBSCAN clustering is controlled by the neighborhood radius parameter ϵ . Loose and tighter clusters produced for different ϵ values are shown in orange and blue, respectively.

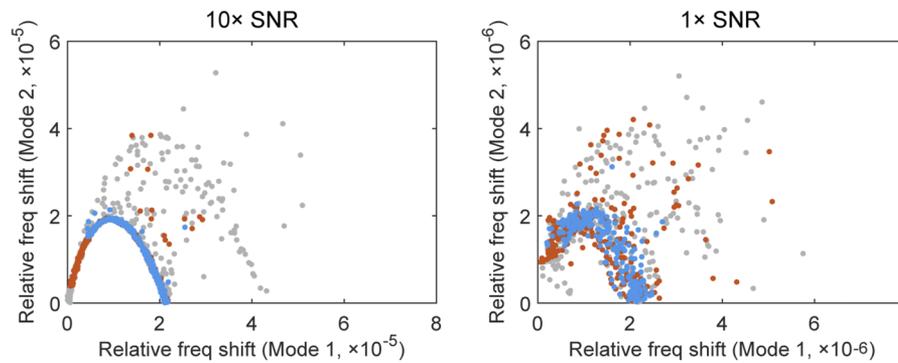


FIG. 7. Resulting frequency shifts of single-event jumps. Once DBSCAN has been applied to the data, the corresponding frequency shifts may be measured and reported. Shown in orange and blue is the choice of ϵ corresponding to somewhat looser and tighter clustering, respectively, than that estimated to belong to single-event jumps.

a result also observed for the 10× SNR and 100× SNR time-series data.

Clustering is then performed by adding more features and determining the precision for each combination of features. Adding

more features generally improves the performance of clustering for any chosen neighborhood radius parameter, ϵ . Moreover, it is observed that increasing the number of features decreases the necessity of precise clustering—in other words, performance is similar

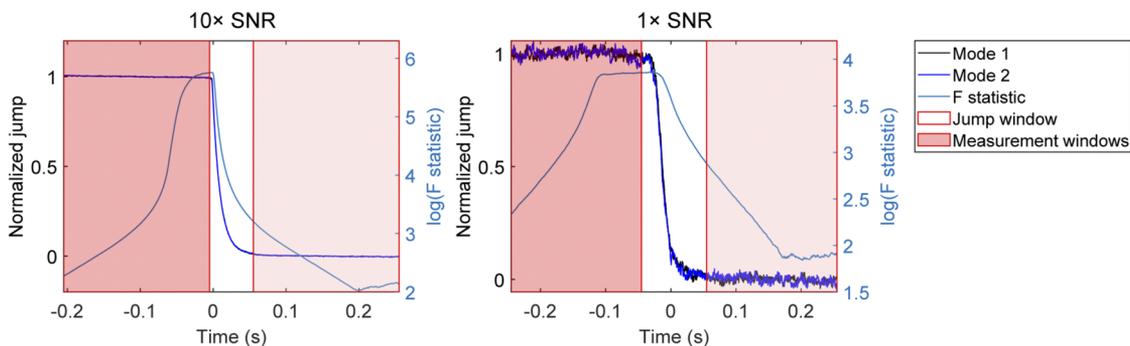


FIG. 8. Change Point Detection Evaluation. Post-processing: the jump signatures exhibit expected behavior similar to that of purely single-event jumps. Moreover, the predefined jump window does not include any values of the transient jump behavior and, thus, will not impact final measurements within the measurement windows. This is observed in 10× SNR and 1× SNR.

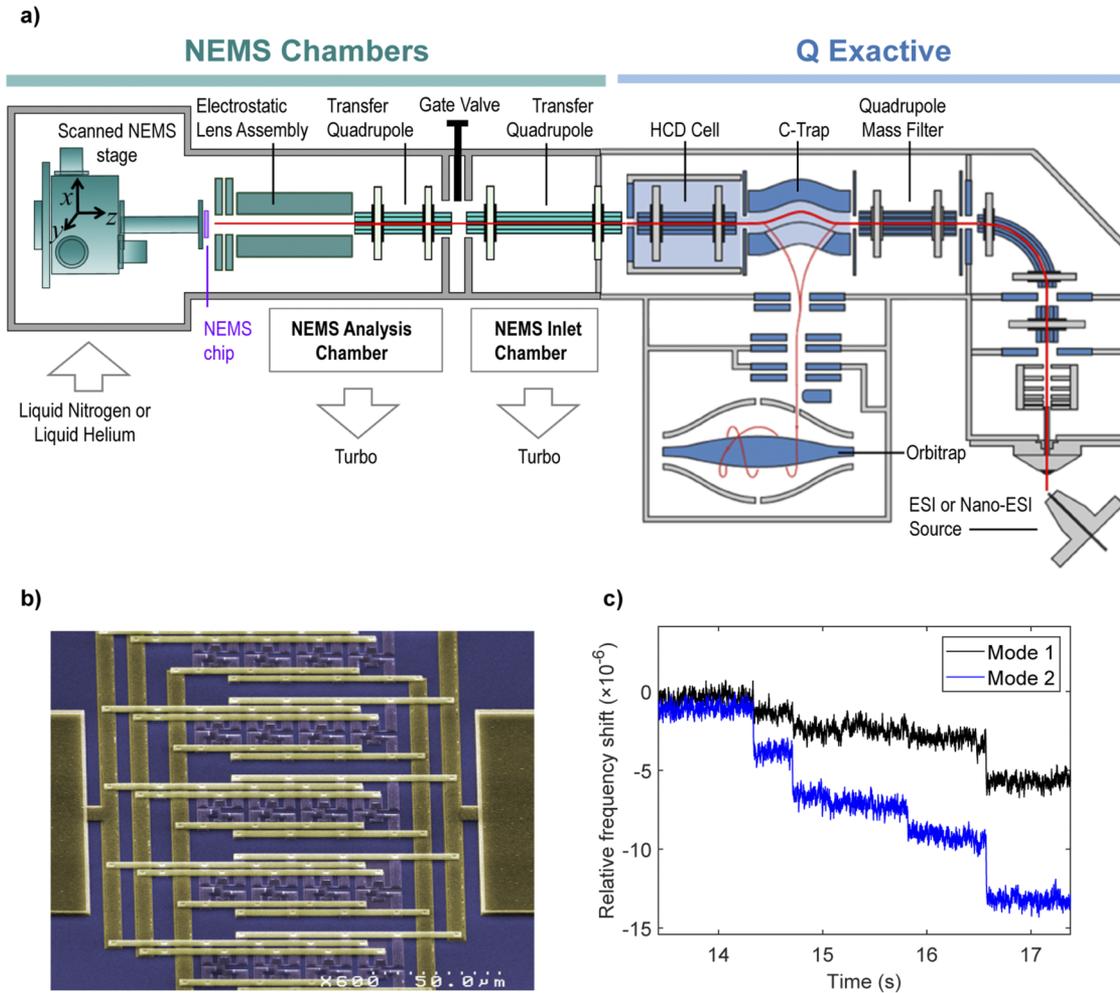


FIG. 9. Measurement apparatus of GroEL molecules with a doubly clamped NEMS beam in high vacuum. (a) Architecture of the Hybrid Q Exactive-NEMS System that delivers intact proteins to the orbitrap chamber for analysis of mass-to-charge ratio and then onto the NEMS for single molecule analysis. Taken from Neumann, Doctoral Dissertation, California Institute of Technology (2020).⁷ (b) SEM image of a 20-device array of doubly clamped beams showing their metallization layers, AlSi (colored in yellow), used to interconnect the electrical connections of each resonator. (c) As GroEL molecules physisorbed into a single NEMS resonator, the resonant frequency of each tracked flexural eigenmode abruptly shifts.

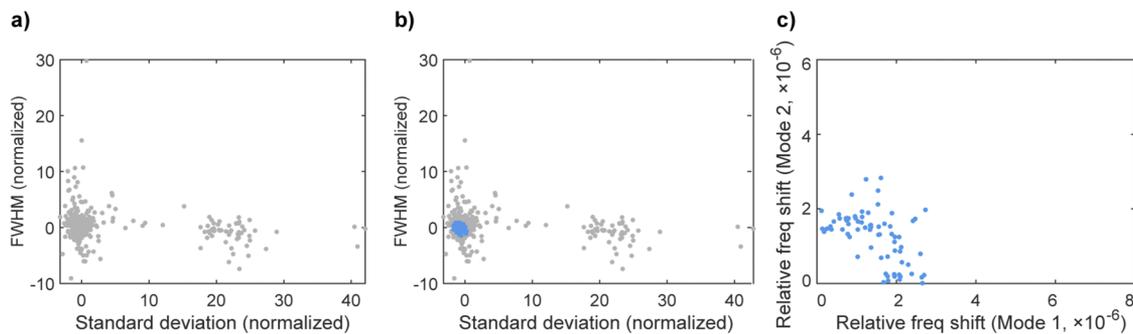


FIG. 10. Single-event jump detection on experimental data (a) Step 1: All identified jump events summarized in a feature space plot. (b) Step 2: Feature space plot depicting clustered single-event jumps. (c) Step 3: Frequency shift space plot of all selected points after clustering.

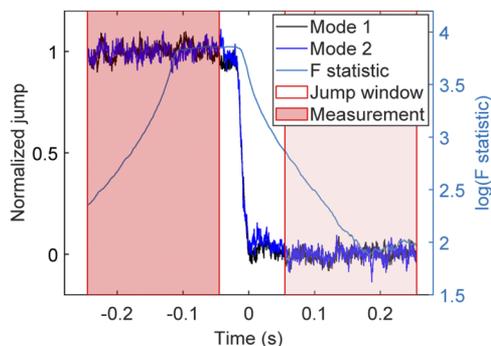


FIG. 11. Frequency shift measurement. Time series depicting jump signature of detected single-event jumps and corresponding F-statistic. For the GroEL data, a fitted exponential gave a time constant of 11.1 ms, close to the PLL time of 10 ms.

across a wider range of choices of ϵ . Table V summarizes the average precision obtained across all possible combinations of features.

The clustering of jump events is shown in Fig. 5 for two different noise estimates: one with 30% (looser cluster) and another with 60% (tighter cluster) of the data classified as noise. The tighter cluster gives a more precise classification, as expected. Figure 6 gives two of the statistical features used for clustering, showing the data have a central tendency, with more central features corresponding to single-event jumps. A significant proportion of multiple-jump events are removed with clustering, but it is not possible to eliminate them entirely. Some multiple-event jumps remain with the looser cluster choice but more events with a lower signal are preserved.

C. Step 3—Output frequency shifts of single-event jumps

As previously described, following the reduction of the data to single-event jumps, the jump magnitude (or frequency shift) is measured and reported. Using the same clustering parameters as

outlined in step 2, Fig. 7 depicts the resulting frequency shifts of the identified single-event jumps. We can readily distinguish single-event jumps from noise because the single-events follow a parabolic-like "backbone", a curve that can be predicted using an Euler-Bernoulli model for the NEMS device in Eq. (1). Multiple-event jumps, and hence outliers, lie outside this parabolic shape.

As a final form of validation, we introduce the concept of a *jump signature*. The jump signature aims to illustrate the prototypical jump with the noisy data of variable SNR events. The algorithm is visually validated by comparing the jump signature before and after removing outliers through filtering and clustering; see Figs. 5 and 6. Compared with the large frequency fluctuations present in individual jumps shown in Fig. 1, the corresponding jump signatures feature significantly reduced noise, demonstrating that the jumps are all effectively aligned. No drift (or slope) in the measurement windows before or after the jump also verifies that multiple-jump events are not a significant portion of the events used for the calculation, c.f. Fig. 8. In addition, for the $10\times$ SNR simulation, fitting an exponential curve to the portion of the jump signature immediately after the jump occurs gives a time constant of 10.02 ms, compared with the 10 ms used for the simulation.

IV. APPLICATION TO EXPERIMENTAL DATA

We now apply the proposed single-event jump detection approach to measure the masses of individual proteins adsorbing to the surface of a doubly clamped NEMS beam in a high vacuum. These macromolecules are of identical (fixed) mass to a degree of high precision and modulated only by the binding of hydrogen, water molecules, etc. Details of these experimental measurements are reported elsewhere,^{7,22,23} for which only a summary is provided here.

A hybrid Orbitrap-NEMS system, illustrated in Fig. 9(a), is used to perform single-molecule nanomechanical mass measurements of *E. coli* GroEL chaperonin, a noncovalent 801 kDa complex consisting of 14 identical subunits.^{7,22,23} GroEL is pre-selected using

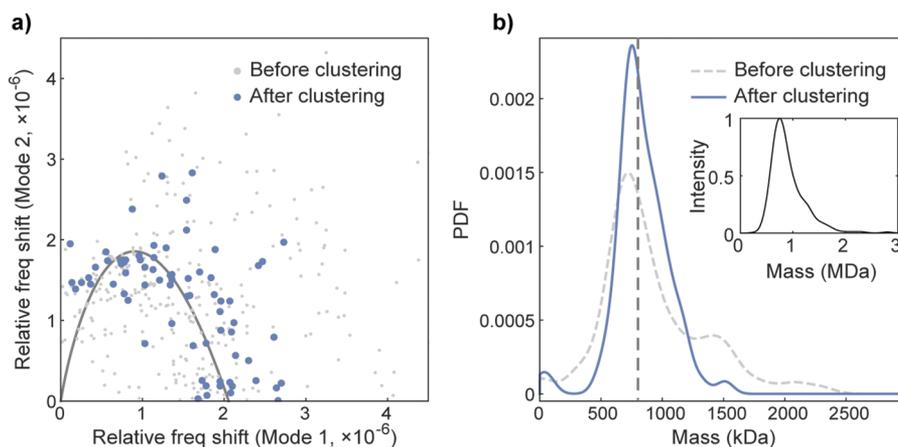


FIG. 12. Mass measurements of GroEL molecules with a doubly clamped NEMS beam in high vacuum. (a) The frequency shift of measurements absorption events of the first two eigenmodes; (solid line) best fit to the E-B model. (b) Smoothed histogram using kernel density estimation showing mass distributions before and after application of clustering to mass adsorption events. The dashed line indicates the expected mass for the 801 kDa GroEL sample. Inset gives data from Neumann *et al.*⁹

the quadrupoles of the orbitrap system, ensuring only intact GroEL molecules are delivered to the NEMS. A 20-device NEMS array of doubly clamped beams is used to localize the focal point of the ion beam, as shown in Fig. 9(b). Subsequently, the smallest NEMS device in the array with the best mass resolution (length of $7\ \mu\text{m}$) is operated in isolation.

Individual molecular adsorption events of intact GroEL molecules abruptly shift the resonant frequency of the first two flexural eigenmodes of the smallest NEMS device; see experimental data in Fig. 9(c). The tracked frequency data collected for each flexural eigenmode are time series, with jump events due to molecular adsorption. These jump events are detected and analyzed using the algorithm described in this article; see Figs. 10 and 11.

After extracting the single mass absorption events from the multimodal time-series data, the frequency shifts are used to predict the masses of the analytes using the method of Dohn *et al.*⁵ The least-squares approach of Dohn *et al.* is formally equivalent to the 2-mode theory of Hanay *et al.*¹ when only two modes are used. These mass measurements are reported in Fig. 12 and are also compared to a previous analysis of the frequency time series, which does not use the present jump detection algorithm.⁹ The data (before clustering) is similar to that reported by Neumann *et al.*⁹ and displays a distinct secondary peak that can be interpreted as a doublet, i.e., two GroEL particles. Clustering removes this peak, suggesting that it is spurious in nature.

V. DISCUSSION

Emerging applications of NEMS mass sensors demand that they operate with lower signal-to-noise and higher event rates for analyte adsorptions compared to instrument response time. In a separate publication describing the instrument design and detection of individual molecules adsorbing to NEMS devices,⁹ the raw data obtained suffered from both issues simultaneously. The feature-based clustering algorithm proposed here demonstrates that multiple-jump events can be reliably removed at the cost of reducing the effective event rate; however, they cannot be eliminated entirely. This emphasizes the need for experimental design with a carefully chosen event rate based on sensor signal-to-noise characteristics. Our jump detection approach provides a framework for determining the upper limit for this event rate based on the response time and signal-to-noise of the instruments. In this way, the protocol described in this study is required to perform high-event-rate experiments, such as those envisaged under high-throughput single-molecule experiments.

Future work could expand upon the approach outlined here in several ways. First, to improve the classification step of the method, additional features of the F-statistic transient surrounding each jump event could be incorporated. This study used the first four moments of the F-statistic transient only. Careful selection of additional higher-order features could allow for the recovery of further single-jump events not captured using the present algorithm. Further work could also consider adaptable window sizes for event detection and measurement, which would provide flexibility to adapt to changing noise conditions and event rates.

Second, future efforts could address the events currently classified as multiple-jump events that are simply filtered out here. In particular, the use of multiple modes to determine mass moments^{4,5}

could allow for accurate characterization of multiple particles landing simultaneously, or equivalently, which are too closely spaced in time to allow for separation into individual jump events. More generally, alternate approaches for either detection or classification may be implemented through this modular framework. For example, given prior information about the jump transient behavior, a model-fitting approach may be used for jump detection.

Finally, this study discusses the importance of adaptability to online detection. Further efforts could make these algorithms fully online by first creating the clustered features using a calibration set and then classifying further events into the appropriate cluster using nearest-neighbor clustering.

ACKNOWLEDGMENTS

The support from the Wellcome Leap Foundation through its Delta Tissue program, Alexander Makarov from Thermo Fisher Scientific, the Heck lab (Utrecht Univ., NL) for providing the GroEL samples analyzed in this work, and the NSF through its MRI and PFI programs is gratefully acknowledged.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

M.L.R. and J.E.S. proposed and supervised the project. A.P.N., A.G., and A.R.N. undertook the project and generated numerical results. A.P.N. developed the moving window F-statistic, significance bootstrapping, and extraction of features. A.G. incorporated clustering with DBSCAN. A.P.N. and A.G. wrote the software, developed simulations, and calculated performance metrics. All authors edited and discussed the manuscript.

Adam P. Neumann: Formal analysis (equal); Investigation (equal); Methodology (equal); Software (equal); Writing – original draft (equal); Writing – review & editing (equal). **Alfredo Gomez:** Formal analysis (equal); Investigation (equal); Methodology (equal); Software (equal); Writing – original draft (equal); Writing – review & editing (equal). **Alexander R. Nunn:** Formal analysis (equal); Investigation (equal); Methodology (equal); Writing – review & editing (equal). **John E. Sader:** Conceptualization (equal); Investigation (equal); Supervision (equal); Writing – review & editing (equal). **Michael L. Roukes:** Conceptualization (equal); Investigation (equal); Supervision (equal); Writing – review & editing (equal).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

APPENDIX: DESCRIPTION OF NUMERICAL PACKAGE

The methodology reported in this study is included in the Matlab and Python libraries for change point detection and classification in jump-detection. This library is written in Matlab and Python

and is available on platforms running Python 3.9 or later. Source code is available from <https://github.com/NEMS-AI/jump-detection> under the MIT license and deployed with complete documentation that includes installation instructions and explanations with code snippets for advance use. The code features the following:

Input. There are two main ways of interfacing with the algorithm. First, by running the entire script, one can simply provide a time-sequenced.csv file, with multivariate populating the proceeding columns.

Modularity. This library is designed such that some aspects may be swapped out for similar algorithms. For example, the user may choose to use a different change point detection algorithm but may want to use the implemented clustering and filtering schemes. For this reason, each step in the algorithm is designed to work with minimal coupling between different modules.

Bootstrapping. This implementation contains code to find preferred parameters to analyze the F-statistic values. Namely, this gives an initial estimate for the Fstat threshold parameter based on the SNR of the provided signal.

Change point detection. This package includes methods to iterate through input signals and calculate F-statistics through a moving window approach. Moreover, for each change point discovered, summary statistics of the F-statics are calculated.

Filtering. This package includes a basic filtering of outlier change points as described in this study.

Clustering. This package includes clustering based on derived features from the F-statistic plot.

Measurement. For our application of NEMS-MS, the primary purpose of the algorithm is to calculate the relative frequency change before and after each event. Knowledge of the precise time of the events is secondary. This package provides methods to perform both calculations.

Evaluation. Evaluation metrics are provided to assess both the alignment of the detected change points as well as the classification of different types of events. These are provided through visual and numerical representations of the event classifications.

REFERENCES

- ¹M. S. Hanay *et al.*, “Single-protein nanomechanical mass spectrometry in real time,” *Nat. Nanotech.* **7**, 602–608 (2012).
- ²J. E. Sader, M. S. Hanay, A. P. Neumann, and M. L. Roukes, “Mass spectrometry using nanomechanical systems: Beyond the point-mass approximation,” *Nano Lett.* **18**, 1608–1614 (2018).
- ³E. Sage, “Neutral particle mass spectrometry with nanomechanical systems,” *Nat. Commun.* **6**, 6482 (2015).
- ⁴M. S. Hanay *et al.*, “Inertial imaging with nanomechanical systems,” *Nat. Nanotech.* **10**, 339–344 (2015).
- ⁵S. Dohn, W. Svendsen, A. Boisen, and O. Hansen, “Mass and position determination of attached particles on cantilever based mass sensors,” *Rev. Sci. Instrum.* **78**, 103303 (2007).
- ⁶E. Sage, “Nouveau concept de spectromètre de masse à base de réseaux de nanostructures résonantes,” Ph.D. Dissertation (Grenoble University, 2013).
- ⁷A. P. Neumann, “Towards single molecule imaging using nanoelectromechanical systems,” Ph.D. Dissertation (California Institute of Technology, 2020).
- ⁸S. Aminikhanghahi and D. J. Cook, “A survey of methods for time series change point detection,” *Knowl. Inf. Syst.* **51**, 339–367 (2017).
- ⁹A. P. Neumann *et al.*, “A hybrid Orbitrap-NEMS instrument for real-time single-molecule analysis of intact proteins,” (unpublished).
- ¹⁰C. Truong, L. Oudre, and N. Vayatis, “Selective review of offline change point detection methods,” *Signal Process.* **167**, 107299 (2020).
- ¹¹W. J. Riley, “Frequency jump detection and analysis,” in *Proceedings of the 40th Annual Precise Time And Time Interval Systems And Applications Meeting* (Institute of Navigation, 2008), pp. 241–254.
- ¹²S. N. Rodionov, “A sequential algorithm for testing climate regime shifts,” *Geophys. Res. Lett.* **31**, L09204, <https://doi.org/10.1029/2004gl019448> (2004).
- ¹³T. Warren Liao, “Clustering of time series data—A survey,” *Pattern Recognit.* **38**, 1857–1874 (2005).
- ¹⁴M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, “A density-based algorithm for discovering clusters in large spatial databases with noise,” in *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining* (AAAI Press, 1996), pp. 226–231.
- ¹⁵H. Hotelling, “The generalization of student’s ratio,” *Ann. Math. Stat.* **2**, 360–378 (1931).
- ¹⁶R. A. Johnson, D. W. Wichern *et al.*, *Applied Multivariate Statistical Analysis*, 405 (Prentice Hall, NJ, 1992).
- ¹⁷N. Martin and H. Maes, *Multivariate Analysis* (Academic Press London, 1979).
- ¹⁸S. Gonçalves and D. Politis, “Discussion: Bootstrap methods for dependent data: A review,” *J. Korean Stat. Soc.* **40**, 383–386 (2011).
- ¹⁹H. R. Kunsch, “The jackknife and the bootstrap for general stationary observations,” *Ann. Stat.* **17**, 1217–1241 (1989).
- ²⁰J.-P. Kreiss and S. N. Lahiri, *Handbook of Statistics* (Elsevier, 2012), Vol. 30, pp. 3–26. Bootstrap methods for time series.
- ²¹D. W. Allan, “Statistics of atomic frequency standards,” *Proc. IEEE* **54**, 221–230 (1966).
- ²²R. J. Rose, E. Damoc, E. Denisov, A. Makarov, and A. J. R. Heck, “High-sensitivity Orbitrap mass analysis of intact macromolecular assemblies,” *Nat. Methods* **9**, 1084–1086 (2012).
- ²³E. Sage *et al.*, “Single-particle mass spectrometry with arrays of frequency-addressed nanomechanical resonators,” *Nat. Commun.* **9**, 3283 (2018).