Chaos

Detecting hidden transient events in noisy nonlinear time-series

Cite as: Chaos **32**, 073131 (2022); https://doi.org/10.1063/5.0097973
Submitted: 03 May 2022 • Accepted: 05 July 2022 • Published Online: 28 July 2022

A. Montoya, E. Habtour and F. Moreu

Detecting hidden transient events in noisy nonlinear time-series

Cite as: Chaos **32**, 073131 (2022); doi: 10.1063/5.0097973
Submitted: 3 May 2022 · Accepted: 5 July 2022 ·
Published Online: 28 July 2022

A. Montoya,[1,2,a)] E. Habtour,[3] and F. Moreu[2]

AFFILIATIONS

[1]Sandia National Laboratories, Albuquerque, New Mexico 87185, USA
[2]Department of Civil, Construction, and Environmental Engineering, University of New Mexico, Albuquerque, New Mexico 87131, USA
[3]Department of Aeronautics and Astronautics, University of Washington, Seattle, Washington 98195, USA

[a)]**Author to whom correspondence should be addressed:** acmont@sandia.gov

ABSTRACT

The information impulse function (IIF), running Variance, and local Hölder Exponent are three conceptually different time-series evaluation techniques. These techniques examine time-series for local changes in information content, statistical variation, and point-wise smoothness, respectively. Using simulated data emulating a randomly excited nonlinear dynamical system, this study interrogates the utility of each method to correctly differentiate a transient event from the background while simultaneously locating it in time. Computational experiments are designed and conducted to evaluate the efficacy of each technique by varying pulse size, time location, and noise level in time-series. Our findings reveal that, in most cases, the first instance of a transient event is more easily observed with the information-based approach of IIF than with the Variance and local Hölder Exponent methods. While our study highlights the unique strengths of each technique, the results suggest that very robust and reliable event detection for nonlinear systems producing noisy time-series data can be obtained by incorporating the IIF into the analysis.

Transient events in time-series data refer to a cluster of points that appear suddenly and decay quickly, often leaving little evidence of their existence in the record. Detection and characterization of these elusive events in nonlinear, noisy, and complex signals is an important scientific topic with applications across many engineering fields. In a physical structure, the presence of transient events can indicate undesirable or unexpected sources of energy. This paper explores three fundamentally different approaches for the initial detection of transient events in simulated dynamic response data exhibiting mixed random and chaotic behavior. This type of dynamic behavior was chosen to stress the capabilities of each technique and represent a boundary of behavior that is possible in field data measurements. The information impulse function (IIF), running Variance, and local Hölder Exponent (HExp) were used to examine the data for localized changes in information content, statistical variation, and point-wise smoothness, respectively. These methods were evaluated by how well they can identify an event in numeric experiments that vary the background noise, the size of the event, and the location of the event. Each experiment consisted of 500 iterations to obtain an assessment of reliability. In this study, the results showed that the information-based approach of the IIF is more effective than the running Variance and the local Hölder Exponent techniques in all but a few cases. In the more extreme test cases, the uniqueness of each approach becomes apparent; each method can identify at least some events that the others could not. This study highlighted the value of the IIF as a new time-series analysis tool while stressing the importance of a diverse analysis strategy.

I. INTRODUCTION

Observation of transient events in time-series data is a commonly encountered engineering problem. Specific applications vary widely, but all have high impact. Some examples include the detection of precursors to anomalies in power transmission

data,[1,2] diseases,[3] and earthquakes;[4] malicious attacks on computer networks;[5] financial market unrests; intermittent or rare atmospheric events;[6–8] rogue oceanic waves;[9] sudden transient behaviors in ecological systems;[10,11] and phase transitions in chemical reactions.[12] In mechanical systems, it has been shown that the presence of transient response in the measured acceleration data is indicative of internal components colliding, loose joints, or acoustic emission from cracks forming in the material.[13] Furthermore, changes in response may be indicative of more subtle changes in the system such as evolving friction forces present in bolted joints.[13,14] These subtle changes can act as precursors to damage[15,16] based on different excitations in the operating environment.[17]

Event detection in the above examples is addressed with various signal processing techniques. Many utilize time-dependent statistics inside a machine learning or multi-step detection scheme.[1,2,5] Notably, the timing and frequency of multiple earthquakes are frequently identified manually by a combination of high-pass filtering and time-frequency analysis.[4] In practice, analysis of structural response data for transient detection is handled similarly by an experienced analyst. This laborious approach has become increasingly impractical in vibration testing of complex structures due to the amount of data that needs to be examined.

Automated event detection would be highly advantageous but requires robust feature extraction methods for the initial detection of transient response. Structural health data collected in-use or as part of field tests are often subjected to unknown random[18,19] or mixed random and harmonic excitation.[20,21] The technical challenge of identifying transient responses in nonlinear systems under these conditions is nontrivial. Random excitation can produce a seemingly unpredictable high response due to momentary forays into unstable regions in phase space.[22] This typically produces heavy-tailed distributions[23] that can obscure the detection of a transient event resulting from internal damage precursors[24–26] or an external impact. For example, flexible structures, which are ubiquitous in aerospace, automotive, wind turbine,[27] soft robotic,[28] and electronic[29] systems, are known to exhibit significant nonlinear behavior and are currently the subject of may significant investigations.[30–32] Such systems require the use of structural health monitoring devices since their structures are often exposed to mixed random and periodic excitation during their operation due to combined vibratory, turbulence,[33] acoustic,[34,35] and electromagnetic[36] forces.

In response to the challenge of identifying transient responses in nonlinear systems, the information impulse function (IIF) was developed to examine time-dependent characteristics in chaotic, nonlinear systems.[37] The development of the IIF, published previously, demonstrated that it could find very small disturbances introduced into the parameters of a harmonically driven Duffing system. The results showed that the IIF adapts to the apparent complexity of the time-series being examined. This ability could be leveraged as a tool for identifying anomalous transient events in nonlinear time-series as more consistent behaviors are given less weight. In this paper, however, we provided a rigorous examination of the IIF sensitivity to noise and uncertainty, which is key for implementing the IIF in real-world applications. The sensitivity examination consisted of a numerical analysis in a parametric study to determine the suitability of the IIF for identifying transient events in a chaotic Duffing system driven by combined random and sinusoidal excitation. The objectives of introducing random excitation were to simulate the environmental noise and to gain theoretical insights of uncertainty in the simulation.

Additionally, this study included two well-established time-series analysis methods, the running Variance and local Hölder Exponent, to benchmark the IIF results and highlight its unique approach. The running (or moving window) Variance can be found in many change point detection algorithms[38,39] and denoising methods.[40] The running Variance used in this paper served as a stand-in for similar statistical methods that implement a moving window. The parameters for all three techniques were chosen to have equivalent time resolution for event detection. The Hölder Exponent, as implemented herein, is a measure of the differentiability or smoothness in the time-series. It has been used successfully in the structural health monitoring community for event detection in applications ranging from structural damage detection in buildings, identification of the onset of corrosion electrochemical leads, to vibration monitoring of rotary tools.[41–43] The mathematical mechanics behind the IIF, running Variance, and Hölder Exponent are detailed in Subsections II A–II C, respectively.

Three numeric experiments were presented, each designed to isolate parameters of the system that impact the efficacy of the event detection methods; detailed in Sec. III. In the first experiment, the noise level was varied, and the parameters of an input pulse were held constant. In the second experiment, the pulse energy was varied at the same time location in a steady noise. For the third experiment, the input signal consisted of a single pulse size and one noise level at three different pulse locations in time. To determine the rate of positive event detection under random excitation, each test case was evaluated using 500 randomly generated sequences pulled from a normal distribution.

For each experiment, the displacement, velocity, and acceleration responses of the dynamical system were evaluated individually to emulate different instrumentation schemes used to acquire time-series data such as piezoelectric transducers (acceleration), microphones (velocity), and linear variable differential transformers (displacement). Detailed analysis of the results is included in Sec. IV.

II. EVENT DETECTION TECHNIQUES

A. The information impulse function

The IIF is an attractive impulse detection technique due to its efficiency in displaying the temporal behaviors of a signal in terms of the information content localized to specific regions in time.[37] Specifically, abrupt changes in time due to disturbances, regardless of amplitude, are compared autonomously to the rest of the record and display sharp spikes associated with those events. The principal concept is that the complex portions of a signal will be more difficult to compress or approximate with minimal error.[44] The effort required to compress a signal is calculated through a process of incremental compression. This section presents the theoretical principles and formulation of the IIF.

Let $x(t)$ be a discrete time-series signal of length N that is uniformly sampled by Δt: for $0 \leq \Delta t \leq \Delta t(N-1)$. The IIF employs the singular value decomposition (SVD) as a simple compression algorithm.[45] The SVD requires that $x(t)$ decompose the signal into a matrix form. It should be noted that the form of the matrix

representation is inconsequential to the IIF theoretical development, and other representations, such as the continuous wavelet transform (CWT), can be utilized. Future study with the IIF using different signal representations including intrinsic mode functions[46] is planned. In this study, the short time Fourier transform (STFT), Eq. (1), was applied to produce matrices for event detection,

$$\Upsilon(\tau,\omega) = \sum_{-\infty}^{\infty} x(t)\,W(\mathrm{t}-\tau)e^{-i\omega t}. \tag{1}$$

Since the IIF is a measure of information content in Υ, the parameter choices that change the content of Υ, such as windowing and overlap processing, will affect the result. It is recommended that a common set of parameters including block size should be considered with any IIF analysis. Herein, the STFT is modulated with a Hanning window, W. Equation (2), with an overlap processing of 50% is as follows:

$$W(t) = 0.5\left(1-\cos\left(2\pi\frac{t}{L-1}\right)\right) \quad 0 \le t \le L. \tag{2}$$

The decomposed signal $x(t)$ is represented by $\Upsilon(\tau,\omega)$ as a function of discrete time step τ and frequency ω. Thus, Y is an m × n matrix, where n and m are the total number of discrete time steps and the number of frequency bins, respectively. The time resolution, $\Delta\tau$, is always greater than Δt and is governed by the choice of segment size L. The signal Υ is decomposed by SVD in Eq. (3),

$$\Upsilon(\tau,\omega) = USV^T. \tag{3}$$

The matrices U and V contain the left and right singular vectors, respectively. The diagonal matrix S contains the singular values. For $n \gg m$, the IIF potential function, Eq. (4), is defined from right singular vectors V, where $V \in \mathbb{C}^{nxn}$,

$$\Phi_{ij}^{R} = \sum_{q=1}^{j} |V_{iq}|^2, \quad \begin{matrix} i = 1,\ldots,n, \\ j = 1,\ldots,C, \end{matrix} \quad C \le m. \tag{4}$$

Note that this function is two-dimensional. The ith index corresponds to the singular vector indexes. Index j corresponds to changes in the rank and is only defined up to C. The rank C defines maximum information[47] or the location of minimum potential. The dummy index q represents the cumulative sum over j.

A size disparity between n and C is needed for optimal observability of a discontinuity or sudden event. A low n produces a poor resolution in the potential curve, which can make differences between vectors indistinct. For an adequate resolution in the time domain with the IIF, L in Eq. (2) must be chosen such that $n \gg m$. The final form of the IIF is given in Eq. (5),

$$IIF(i) = \frac{1}{C(C+1)}\sum_{j=1}^{C}\Phi_{ij}^{R}, \quad \begin{matrix} i = 1,\ldots,n, \\ j = 1,\ldots,C, \end{matrix} \quad C \le m, \tag{5}$$

The IIF is the sum of the potential function, Eq. (4), normalized such that the sum of the IIF over i is always 1/2. Therefore, the IIF can be thought of as a representation of the work done by the SVD to describe the signal information for each singular value at i. While the IIF has a similar form factor to other time-series analysis techniques, IIF values are not independent from one another. IIF results are always relative to the internal informational relationships present in the signal Υ.

For the test cases presented in this paper, the window size in Eq. (2) was chosen to be 16 points with an overlap of 8 points (50%). This resulted in a time and frequency resolution in the STFT of 0.8 s and 0.625 Hz, respectively. These values correspond to an m of 16 and an n of 511 in Eq. (4), which satisfies the requirement $n \gg m$.

B. Running Variance

Running or moving window statistics are commonly used for change point detection.[48] Moving windows are necessary for examining the evolution of a sequence because statistical measures are time independent. To calculate the running Variance with a 50% overlap, $x(t)$ is divided into overlapping segments of length n, and an overlap size $n/2$. Equation (6) defines the starting and ending points for each segment,

$$\begin{gathered} s_j(\tau),\ t\left(\frac{n}{2}(j-1)+1\right) \le \tau \le t\left(\frac{n}{2}(j-1)+n\right), \\ 1 \le j \le \frac{N-n}{(n/2)}+1. \end{gathered} \tag{6}$$

Equation (6) calculates the total number of segments for a given overlapping scheme. With the segment mean represented by $\overline{s_{j(\tau)}}$, Eq. (7) computes the Variance per segment,

$$V_j(\tau) = \frac{\sum_j^n |s_j(\tau) - \overline{s_{j(\tau)}}|^2}{(n-1)}. \tag{7}$$

For the test cases presented in this paper, the segment size was chosen to be 16 points with an overlap of 8 points (50%). Initial test cases showed that this widow scheme was effective for the size of target events in this study.

Aside from using a statistical approach vs an information theoretical approach, the running Variance differs from the IIF in another important way. Each Variance value is independent of the others in the record aside from the influence of adjacent segments from the overlap.

C. Hölder Exponent

The local Hölder Exponent (HExp) measures the local smoothness of a time-series. Sohn *et al.*,[49] following the description given by Struzik for financial time-series analysis,[50] were the first to apply HExp to structural response data for detection of discontinuities.

The exponent m is found by calculating the slope across wavelet modulus maxima, $\Omega(f,\tau)$, over frequency in log space [Eq. (8)],

$$\mathrm{HExp}(t) = -\frac{\log(\Omega(f,\tau))}{\log(f)}. \tag{8}$$

The wavelet transform is provided in Eq. (9). The mother wavelet,Ψ, used in this analysis was the generalized Morse wavelet,

$$\begin{gathered} \Omega(f,\tau) = \int f(t)\Psi_{f,\tau}^{*}(t)dt, \\ \Psi_{f,\tau} = \frac{1}{\sqrt{f}}\Psi\left(\frac{t-\tau}{f}\right). \end{gathered} \tag{9}$$

The HExp value is increasingly negative for less smooth regions. Sharp decreases in value correspond to possible discontinuities in the time-series. Both Struzik[50] and Sohn *et al.*[49] noted that a Fourier transform (STFT) can be used in place of the continuous wavelet transform (CWT), but Sohn *et al.* found the STFT version to be much less effective. Consequently, the CWT is utilized in this analysis.

The HExp derivation has some similarities with the IIF. It is derived by considering an approximation to the signal and utilizes a time-frequency method as a starting point. There are three differences between them. First, the IIF only considers the right singular vectors, a normalized space, while HExp uses the entire signal. Second, the value of HExp per unit time is independent of other values, which is not the case with the IIF. Third, unlike the HExp, the IIF does not require data filtering and smoothing to be effective.[49]

The scheme proposed by Sohn *et al.* is to apply a high-pass filter and then perform a running average. For this analysis, the applied filter was an IIR (zero-phase) with a 0.05 rad/sample cutoff frequency. The default MATLAB setting in the high-pass filter function used has a stopband attenuation of 60 dB and a 0.1 passband ripple. The edge effects from the filter impact the evaluation method and need to be removed, as discussed in Sec. III C. After filtering, a tapered cosine window (Tukey) was applied to the HExp results. The ratio of the taper section to the length of the window was 0.1. After the Tukey window was applied, a 16-point running average 50% overlap was applied. This was chosen to match the windowing for STFT and the running Variance from Eqs. (2) and (6), respectively.

III. NUMERICAL EXPERIMENT DESIGN

A. Target events

A realistic transient response was generated by adding a single sided pulse into the input of the nonlinear process. A pulse in the input of a single-degree-of-freedom harmonic system produces an oscillatory decaying event.[51]

Equation (10) defines the pulse $P(t, A, d, t_p)$ as a function of time t, amplitude A, and width d in the samples. The start location t_p is in seconds. P is expressed in dimensionless units of acceleration. The shape and peak of the added pulse depend on d. For a unit pulse, the form is a Dirac Delta, $\delta(t - t_p)$. For wider pulses, a Hann window is used to create a more realistic pulse shape,

$$P(t, A, d, t_p) = \begin{cases} A\delta(t - t_p), & d = 1, \\ 0.5A\left[1 - \cos\left(2\pi \frac{t}{d-1}\right)\right], & d > 1, t_p \le t \le t_p + (d-1)dt, \\ 0, & \text{otherwise.} \end{cases} \tag{10}$$

Figure 1(a) shows the resulting pulse shapes as a function of time for seven different combinations of A and d. Pulse widths and peak magnitudes overlap to highlight any strong sensitivities of each metric to A or d. The range of pulse widths was chosen to fit inside a 16-point window, which was used by all three detection methods to sample and evaluate each time-series. Figure 1(b) shows the corresponding area under each pulse, demonstrating a smoothly increasing momentum for each pulse.

B. Nonlinear random process

A random process is a linear or nonlinear function of one or more random variables, where the random variable forms a sequence in time. Any given sequence will produce a deterministic outcome due to the function's operation. This constitutes a specific realization of the random process. A collection of realizations with randomly generated sequences is termed an *ensemble* of the process, and each value in time will be stochastic.[19]

A sequence of 4096 points, as such $X(t) \in \{x_1, x_2, \ldots, x_n\}$, $1 \le n \le 4096$, and sampled at 10 Hz ($\Delta t = 0.10$). The realization length was chosen to emulate a typical block size used in IIF and HExp analysis. The lower sample rate, determined by inspection, ensured that the chaotic system can oscillated about both stable locations without aliasing the results. Each x_n is drawn from the sample space defined by the probability density function of a zero-mean Gaussian process. The resulting random variable $X(t)$ represents zero-mean Gaussian white noise with a standard deviation given by σ. $X(t)$ by itself or used as an input to a deterministic function is an instantiation of a random process. By itself, it is both stationary and ergodic, meaning that the statistical characterization does not vary with time (within a realization) or across realizations (per instance in time).[19]

Equation (11) is a nonlinear, deterministic function of a Gaussian random variable $X(t)$. The output response is given in terms of state variables $Y(t)$ and $\dot{Y}(t)$. Similar to Duffing's equation, this system is bistable and exhibits chaotic behavior under certain excitation levels.[52] Equation (11) represents a nonlinear random process perturbed by the pulse defined in Eq. (10). The parameters of the Duffing system were chosen to match the chaotic system studied in previous work by the authors with the IIF.[37] For theses stiffness and damping parameters, the force amplitude of 0.33 produces a chaotic response. Unit mass, customarily, is assumed. Equation (10) is expressed in a dimensionless form in order to represent a general set of dynamic behavior. The system's initial state is at a stable equilibrium point,

$$\begin{aligned} \ddot{Y}(t) + 0.3\dot{Y}(t) + Y^3(t) - Y(t) &= 0.33\, sin(1.2\, t) + X(t, \sigma) \\ &\quad + P(t, A, d, t_p), \\ Y(t = 0) &= 1, \\ \dot{Y}(t = 0) &= 0. \end{aligned} \tag{11}$$

Numeric solutions for Eq. (11) were produced in MATLAB by the ODE89 solver that was run inside a loop iterating at 1/10 of a second with the following set of parametric equations [Eq. (12)]:

$$\begin{aligned} y' &= \dot{Y}, \\ y'' &= \ddot{Y}, \\ y''' &= -0.3\ddot{Y} - 3Y^2\dot{Y} + \dot{Y}(t) + 0.33(1.2)\cos(1.2t). \end{aligned} \tag{12}$$

Despite having a functional form for P, the noise and pulse terms are modeled as an impulse to the system that occurs at each 1/10 s and is zero otherwise. The derivative of $X + P$ at each loop iteration is zero in this paradigm. When the initial conditions are updated at each loop iteration, the initial acceleration accounts for the additional force from $X + P$.

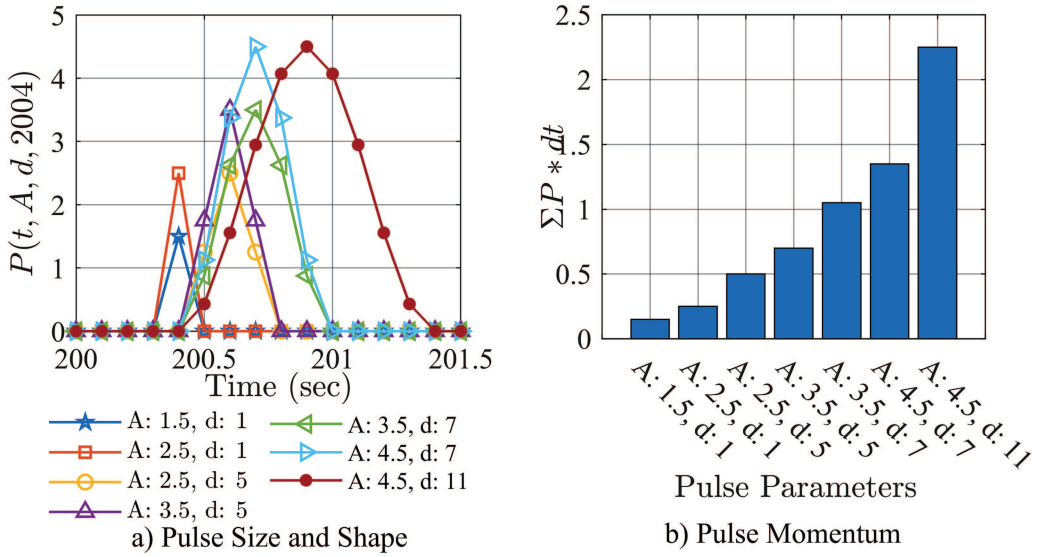


FIG. 1. Different combinations of *A* and *d* producing pulses with variable momentum. (a) Pulse size and shape in time. (b) Momentum per pulse as calculated by the pulse area.

Given a Gaussian random variable input, a nonlinear process will produce non-Gaussian values. Additionally, the process may be nonstationary and, therefore, non-ergodic. In general, periodic functions are nonstationary. However, they can be periodically stationary if windowed with the period.[19] The Duffing oscillator has been shown to be periodic stationary for the period matching the forcing function.[53] In this case, the period is approximately 5 s. The target events are much smaller (less than one second), necessitating examination of this scenario as nonstationary data.

The behaviors of a Duffing system exposed to random vibration can be described as both being stochastic and deterministic. Characterization of those behaviors leads to coupled equations.[54] Dey and Sekh[55] observed that added random excitation changes the Poincaré section, displaying the attractor. Specifically, they noted that the compact shape of the Poincaré section diffuses in a random field.[55] This effect can be readily observed in simulations of Eq. (11) for different levels of noise ($\sigma = 0.1, 0.9$) and no pulse ($A = 0$) (Fig. 2).

Equation (11) without a random variable ($\sigma = 0$) or pulse input ($A = 0$) was simulated to 65 536 points (109.23 min) to obtain a Poincaré section with high resolution, as shown in Fig. 2(a). Bivariate histograms of the Poincaré sections are for five hundred 4096-point realizations with no pulse input ($A = 0$), but added random excitations with $\sigma = 0.1$ and $\sigma = 0.9$ are shown in Figs. 2(b) and 2(c), respectively. The distribution of points tends to follow the shape of the zero-noise attractor shown in Fig. 2(a).

The images in Fig. 2 confirm the findings of Dey and Sekh[55]and Anh and Hieu[54] and suggest that the simulated data that include random excitation are still very much chaotic and heavily influenced by the attractor. This is important since the background coupled stochastic and chaotic response precludes the use of simple filters and denoising methods to simplify the problem. It also implies that the added pulse is subject to these same dynamics, while simultaneously being obscured by random noise. Variable noise and variable pulse amplitudes may, thus, produce different trends in detectability and were examined separately, as explained in Subsection III C.

C. Procedure

Three numeric experiments were conducted for this paper. Each experiment contains three or more test cases. Experiment I has five test cases, experiment II has seven, and experiment III has three. The general procedure for each test case is described in Fig. 3.

For each test case, an ensemble of 500 realizations was generated from a zero-mean normal distribution. The ensemble is represented by a red dotted arrow in Fig. 3. A pulse and a sine wave were added to each realization in the ensemble before propagating it through the Duffing system. The blue box in Fig. 3 represents the evaluation procedure for IIF, Variance, and HExp. The same seed in the pseudo-random number generator was used for each test case to make equitable comparisons across all test cases.

The effectiveness of each method to time-localize the input pulse was evaluated by collecting the maximum value (or minimum for the HExp) within $\pm 2\Delta\tau$ for the resampled time vector ($\sim\pm$ 1.6 s). The resulting range was larger than any of the tested pulse durations, such that small variations in peak response do not affect the outcome. The value in this narrow range is the *local peak*. A *global peak* is then collected for the full record. The two peak values were compared. If they were equal, a success was counted toward a total success rate: number of successes over the total in the ensemble.

Positively identified events were evaluated by how easily they could be observed above the background. A threshold was set to be the top 95th percentile of the peaks for each segment as determined by the empirical cumulative distribution (ECDF). The 95th

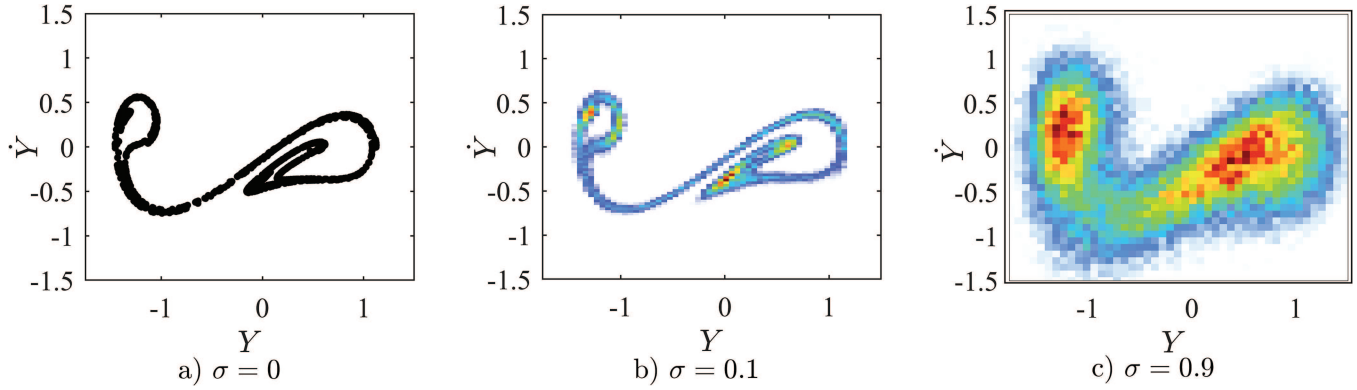


FIG. 2. The combined stochastic and deterministic behavior of the nonlinear random process of Eq. (11). Points on the Poincaré plot appear to diffuse away from the attractor. (a) A 65 536-point solution without noise. (b) Five hundred 4096-point realizations with noise level 0.1 plotted as a bivariate histogram. (c) Five hundred 4096-point realizations with noise level 0.9 plotted as a bivariate histogram.

percentile was chosen to capture the heavy-tailed character of the output data while providing a relatively stable datum compared to results obtained with 99th percentile. The resulting peak-to-threshold ratio(PTR_{95}), expressed in dB, is given in Eq. (13),

$$PTR_{95}(\text{peak}) = 20\log_{10}\left(\frac{\text{peak}}{T_{95}}\right). \tag{13}$$

For each of the following three numeric experiments, selected values for σ, t_p, and the tuple $[A, d]$ were applied to sample the possibility space created by Eq. (11). An example output is given to guide the reader through the summarized results.

IV. RESULTS AND DISCUSSION

A. Example output

Figure 4 shows the results of three metrics for a time-series. The left column displays the full record, and the right column shows the behaviors surrounding the pulse location within 15 s. The top row displays the source time-series containing an event and a corresponding baseline result, where no event is present. Results for the IIF, Variance, and HExp for the pulsed response only are given in the rows below.

Figure 4(a), top left, displays two acceleration responses from the numeric solution of Eq. (11). Plotted in solid black is the result when the pulse added to input has the parameters $[A, d] = [3.5, 5]$ at $t = 200.3$ s. The baseline, plotted with a dashed blue line, results when the pulse input is absent: $[A, d] = [0,0]$. In both cases, the applied random realization was $X(t, \sigma = 0.5)$. Although significant, the pulse [Figs. 4(a) and 4(e)] is only readily visible when the baseline is included due to several large spike-like peaks in the acceleration response.

The running Variance, shown in Fig. 4(c), interprets the response similarly to spikes that can easily be mistaken for an event. As a result, the running Variance cannot successfully identify the event. On the other hand, both the IIF and HExp [Figs. 4(b)

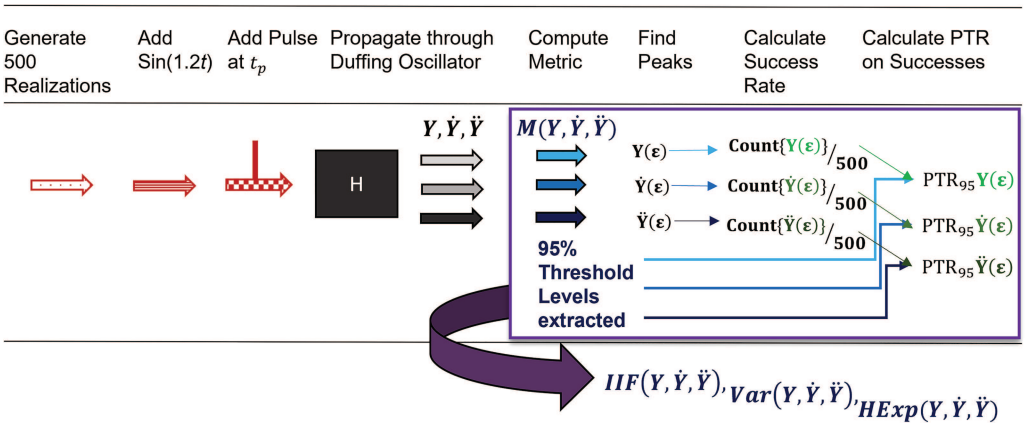


FIG. 3. Simulation and evaluation methodology for each test case.

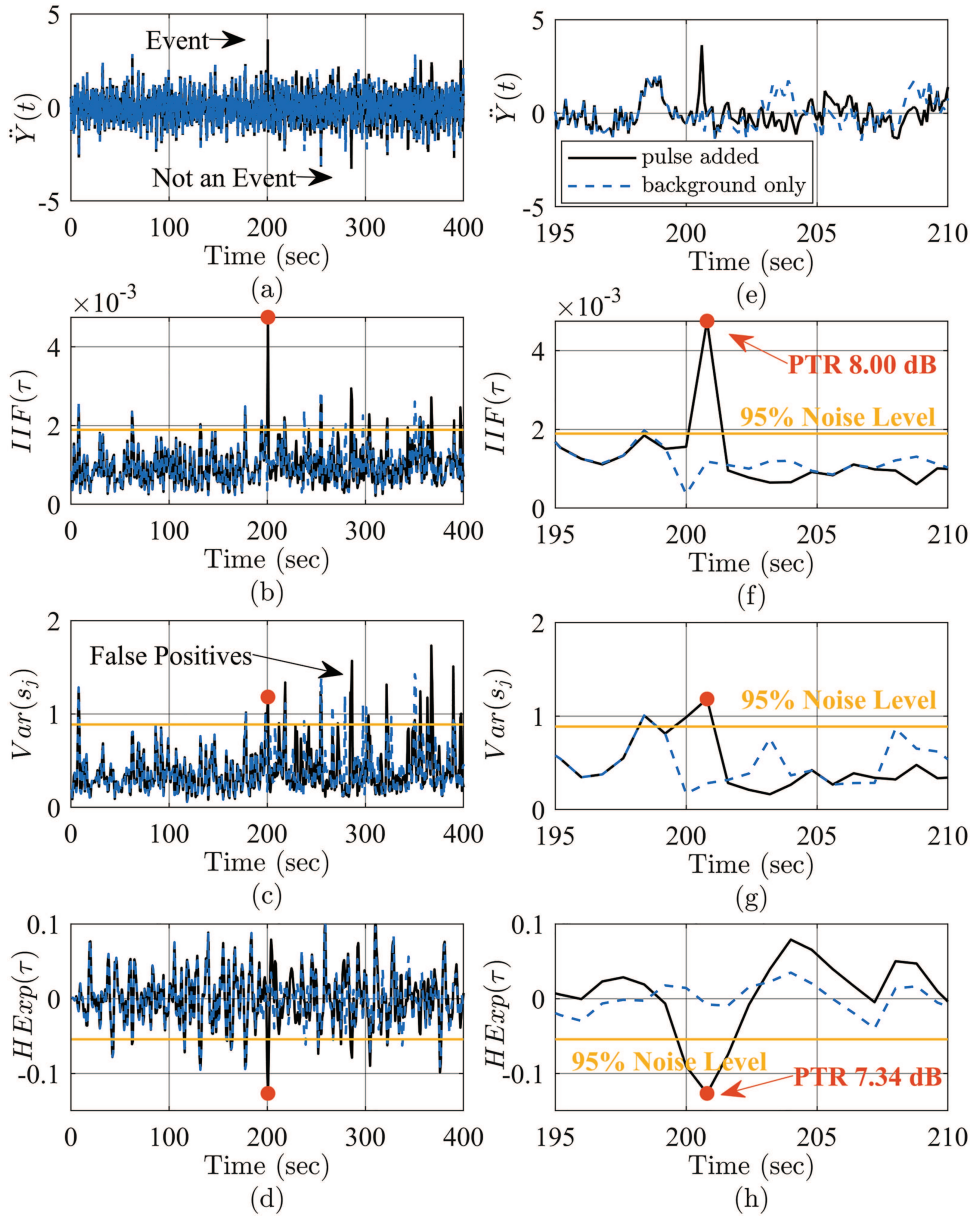


FIG. 4. Example test case at $\sigma = 0.5$ showing the acceleration results for all three metrics. (a) Two example time-series with a pulse (black) and without a pulse (blue). (b) IIF results for both time-series in (a), with a pulse (black) and without a pulse (blue). (c) Variance results for both time-series in (a). (d) HExp results for both time-series in (a). (e) 15 s surrounding the pulse location. (f) IIF results 15 s surrounding the pulse location with peak-to-threshold (PTR) value for the identified pulse. (g) Variance results 15 s surrounding the pulse location. The event is above the 95% noise level, but no PTR value is given due to the false positives visible in (c). (h) HExp results 15 s surrounding the pulse location within corresponding PTR value.

and 4(d)] can easily identify the event. The IIF results are slightly more conclusive in that the distance above the next largest peaks is more significant than in the HExp case, which is reflected in the large PTR value for IIF.

In both cases, the IIF and HExp were less sensitive to the obfuscating character of the background response, but the mechanisms responsible for success are entirely different. The HExp success is because the pulse produced a sharp enough response that the continuity in the signal was disrupted at a greater rate than the disruption created by the random background. The IIF is successful because the information content at the pulse location is sufficiently higher than the rest of the signal. The IIF is a measure of the

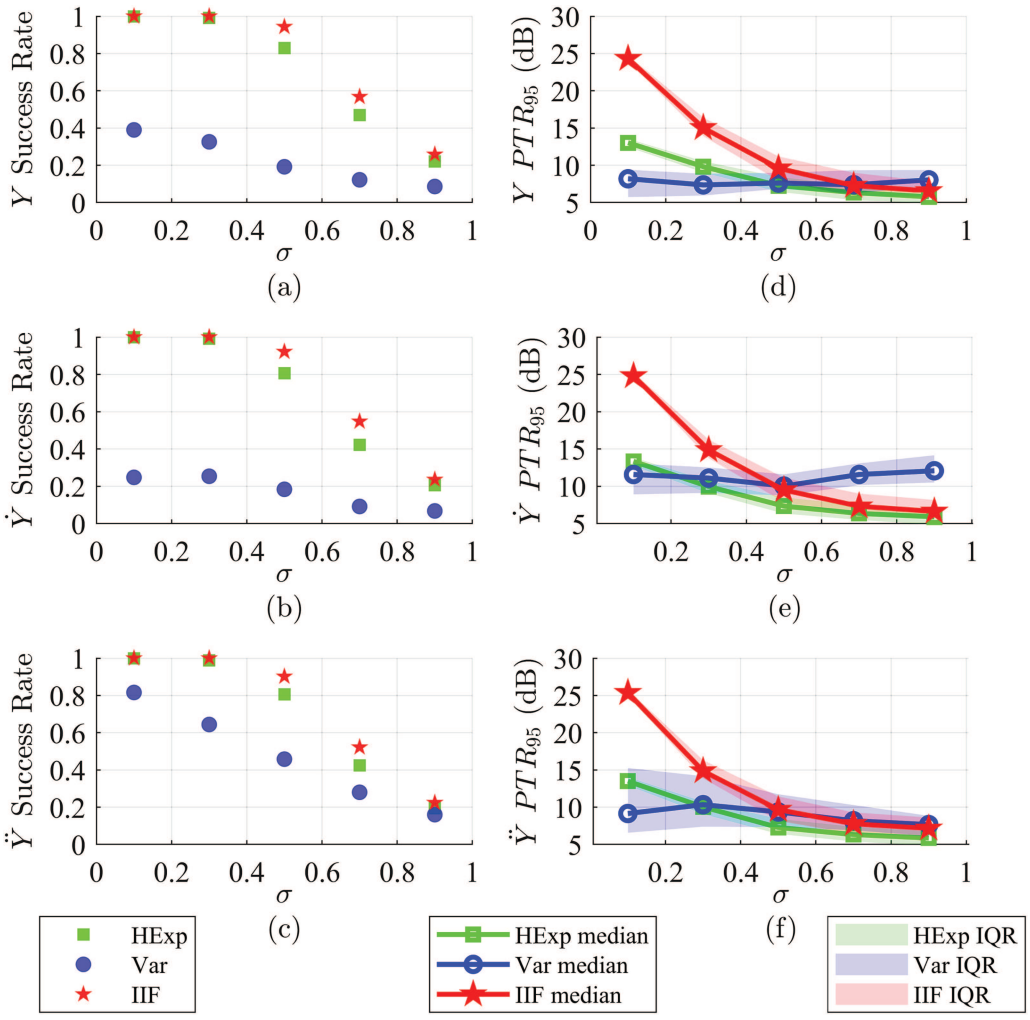


FIG. 5. Success rates and resulting PTR values for all five test cases in experiment I. (a) Success rates for displacement. (b) Success rates for velocity. (c) Success rates for acceleration. (d) Peak-to-threshold (PTR) values for the successful displacement results in (a). (e) PTR values for the successful velocity results in (b). (f) PTR values for the successful acceleration results in (c).

event's uniqueness, with each value weighted against the rest of the record.

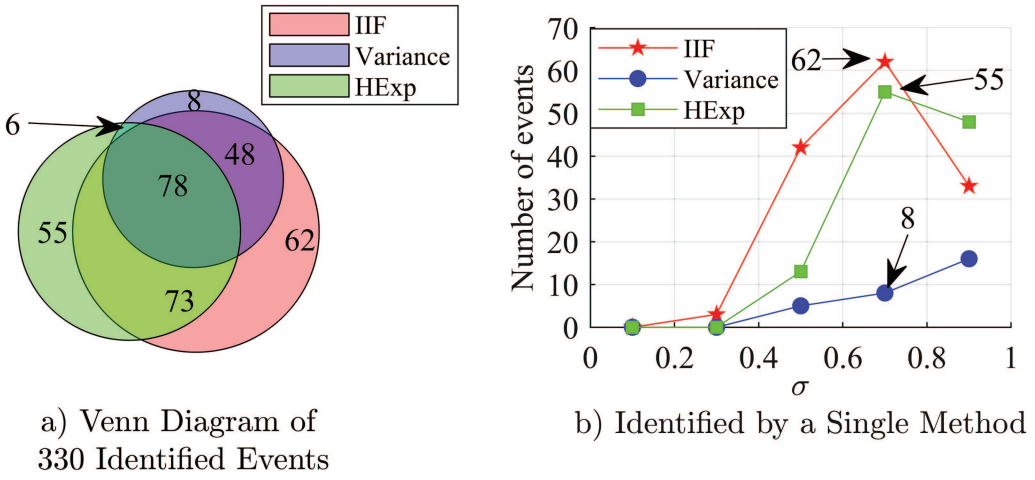
B. Numeric experiment I: Variable noise

The amplitude of the random variable was varied by choosing a different value of σ to define the originating normal distribution. Five different values of σ were used: 0.1, 0.3, 0.5, 0.7, and 0.9. The pulse width, amplitude, and starting location were held constant at $d = 5$ points, $A = 3.5$, and $t_p = 200.3$ s.

Figure 5 presents the results of experiment I. The left column [Figs. 5(a), 5(b), and 5(d)] displays the success rates as a function of sigma for displacement, velocity, and acceleration, respectively. The right column [Figs. 5(d)–5(f)] displays the median PTR values as a function of σ. The shaded regions are the interquartile range (IQR) defined by the 75th and 25th percentiles of the PTR values.

As expected, the success rates for all metrics are inversely proportional to noise. The PTR values for the IIF and HExp also continually decreased as noise increases. In terms of success rate, the IIF outperforms the other methods. In terms of PTR, the IIF beats the HExp in all cases and the running Variance in most.

The success of the Variance requires some scrutiny. Variance appears to not decrease in confidence as σ increases. In fact, it has a positive slope before overtaking both the IIF and HExp at $\sigma > 0.5$ on both the velocity and displacement records. However, the success rate is much lower than the two other metrics at 10% or less. This means that at $\sigma = 0.9$, the PTR is calculated off to 50 or fewer points.



a) Venn Diagram of 330 Identified Events

b) Identified by a Single Method

FIG. 6. (a) Identified events by method for experiment I, $\sigma = 0.7$, acceleration. (b) Events identified by a single method only for experiment I, acceleration.

The interquartile range for Variance when the success rate was high (80% for $\sigma = 0.1$ on the acceleration) was wide. It is possible that results with fewer points cannot capture the variation. Comparatively, the PTR values for Variance are steady compared to the IIF and HExp.

An important factor to consider when comparing these metrics is that each method has a unique focus. Success in one metric does not imply success in another. Figure 6(a) shows a Venn diagram for 330 positively identified peaks (out of 500 possible) in the acceleration responses corresponding to $\sigma = 0.7$ noise level from Fig. 5 by identification method. Figure 6(b) shows the number of peaks for each metric identified by only one metric as a function of noise level. To guide the reader, the corresponding values from Fig. 6(a) are labeled in Fig. 6(b) with arrows. Note that the overlap and separation in event detection by a method is only viewable when all have less than 100% success rate.

The HExp was able to find more events that were not found otherwise in exactly two cases: $\sigma = 0.9$, acceleration as shown in Fig. 6(b), and by a smaller margin $\sigma = 0.9$, velocity. In all other cases, the IIF was superior in this respect.

While the IIF captures most of the events, there is a benefit to using all three. For the example shown in Fig. 6(a), the success rates for the IIF, Variance, and HExp are 52%, 28%, and 42%, respectively. Including all three increases, the success rate is 66%.

B. Numeric experiment II: Variable pulse

For this experiment, the amplitude of the random variable was held constant at $\sigma = 0.5$, the pulse location remained at 200.3 s, and the pulse width and amplitudes were varied. Seven different combinations of the tuple $[A, d]$ were used: $[A, d] \in \{[1.5, 1], [2.5, 1], [2.5, 5], [3.5, 5], [3.5, 7], [4.5, 7], [4.5, 11]\}$. The pulses corresponding to the tuple set above were integrated to obtain the momentum imparted to the system that each pulse represents. The momentum values increased monotonically and are plotted

TABLE I. Notional time-series characteristics that predict the success or failure of one event detection method over another.

	Time-series characteristics			
	Successful detection	Possible failure mechanisms		
IIF	Event has comparative uniqueness	Random excitation mimics the event elsewhere in the record		
Running Variance	Event contains values with comparatively wide variation		Large amplitude spikes due to nonlinear behavior exist elsewhere in the record	
HExp	Event has a comparatively sharp slope			Event shape is smoothed from superposition of nearby random points

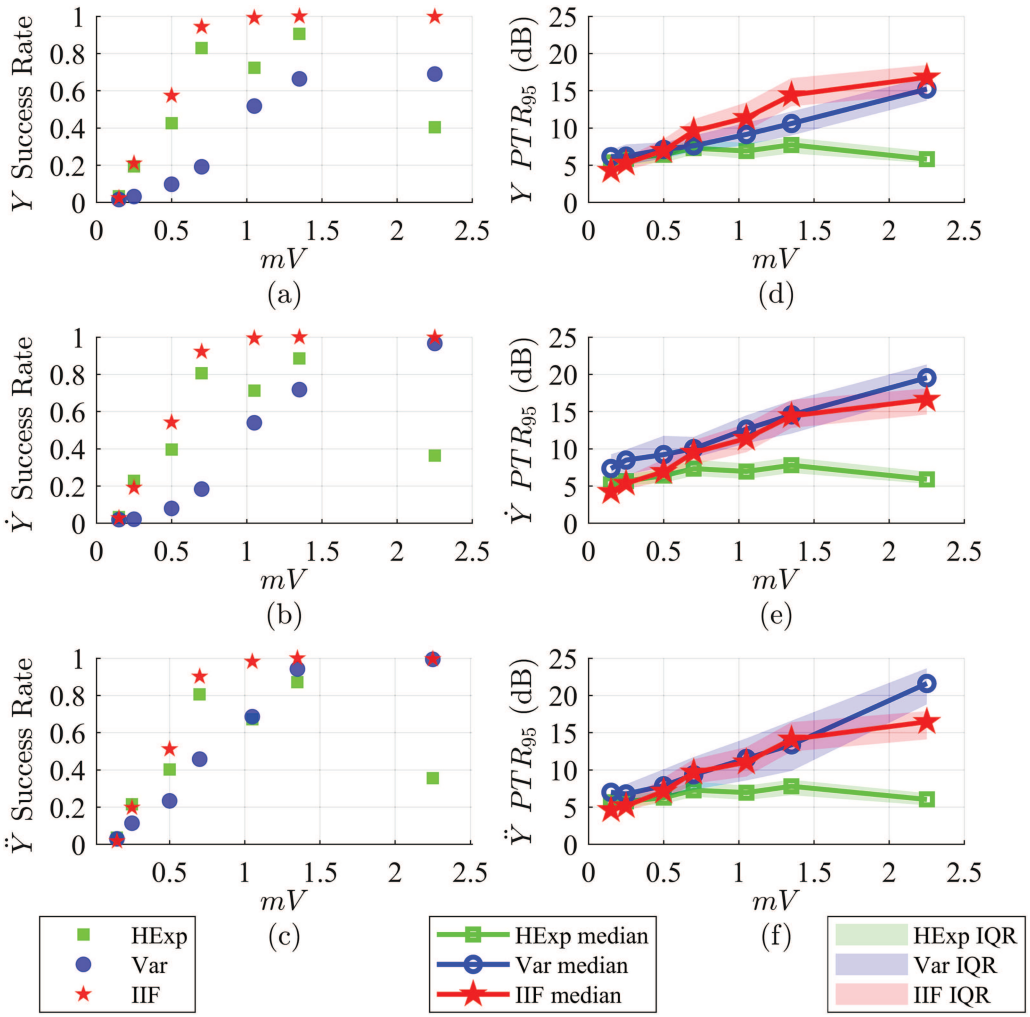


FIG. 7. Success rates and resulting PTR values for all seven test cases in experiment II. (a) Success rates for displacement. (b) Success rates for velocity. (c) Success rates for acceleration. (d) Peak-to-threshold (PTR) values for the successful displacement results in (a). (e) PTR values for the successful velocity results in (b). (f) PTR values for the successful acceleration results in (c).

against the success rate and PTR for each pulse in Fig. 7. The format is otherwise the same as in Fig. 5.

Both the IIF and Variance have increasing success rates with increased pulse momentum, with the IIF displaying higher success rates. The PTR values are very close for most test cases with overlapping IQR. As expected, Variance takes a much clearer lead over the IIF at larger pulse values. Because the IIF is a normalized metric, it is less sensitive to amplitudes.

The HExp shows a notable drop in performance at two locations: $[A, d] = [3.5, 7]$ and $[4.5, 11]$. The momentum values are $mv = 1.05$ and 2.25, respectively. Both these correspond to instances where the pulse shape widened from the previous case. The HExp, in theory, is a measure of smoothness in the signal. Wider pulses produce shock response with a comparatively lower rise time. This should make the response at the event smoother and, therefore, more difficult for the HExp to identify. PTR for HExp is comparatively steady, not trending strongly with pulse size.

The HExp has the highest success rate in both the velocity and acceleration results at the two smallest (and sharpest) pulse test cases. The margin is small. For the velocity results on $[A, d] = [2.5, 1]$, the success rates for HExp are 22.8% vs 19.2% for the IIF. The PTR values between methods are difficult to differentiate for these cases. As with the Variance in Experiment I, the success rate is $< 20\%$. Thus, a PTR comparison only represents a handful of points and should be interpreted with caution.

C. Numeric experiment III: Variable location

Three different pulse locations were tested in this experiment to explore any effect the location has on a nonstationary time-series.

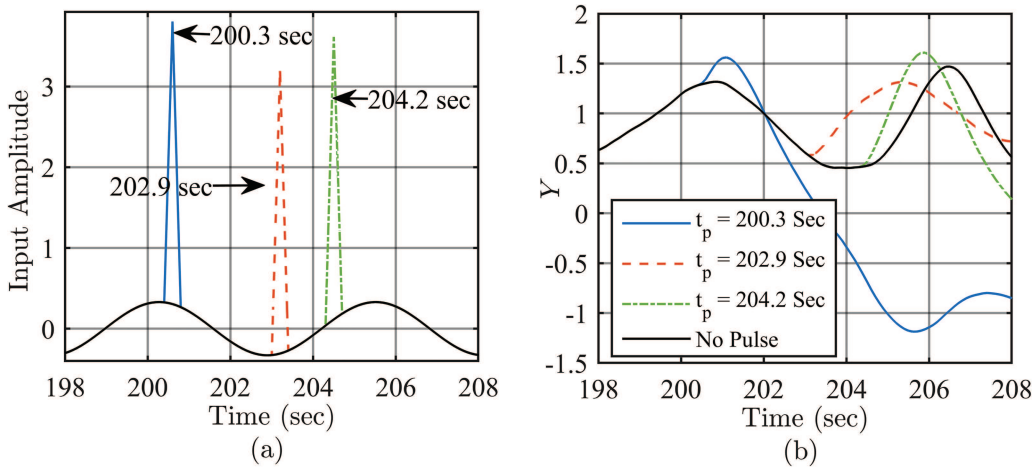


FIG. 8. (a) Tested pulse location with labeled start times plotted on $0.33\sin(1.2t)$. (b) Resulting displacement results from Eq. (11) for the same random realization and the pulse locations in (a).

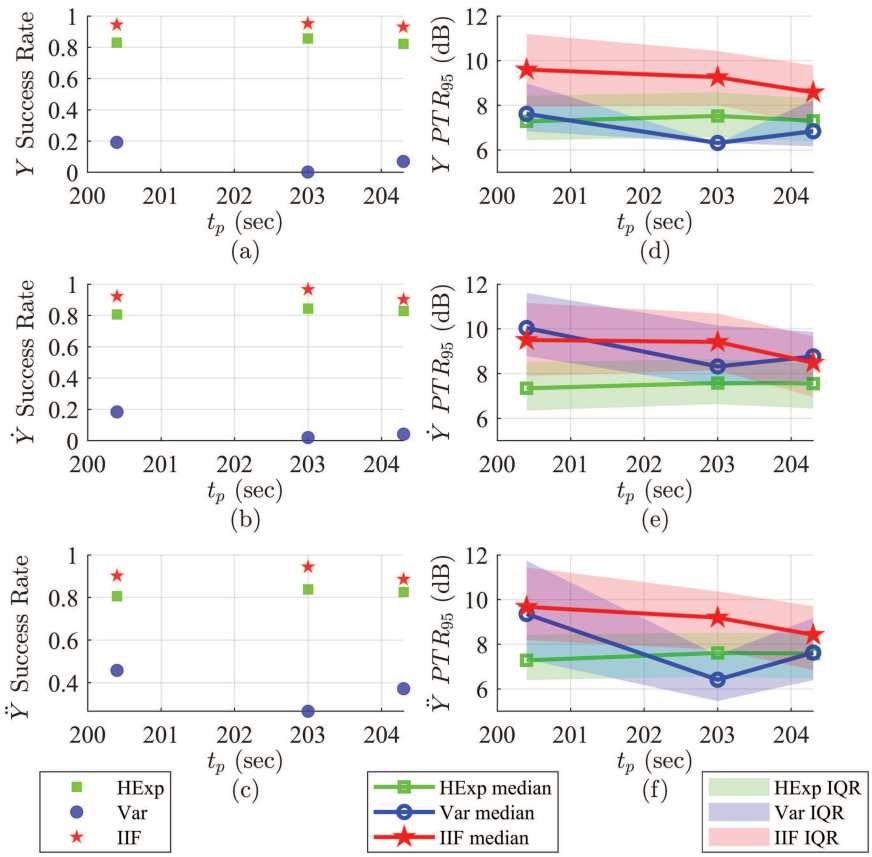


FIG. 9. Success rates and resulting PTR values for all three test cases in experiment III. (a) Success rates for displacement. (b) Success rates for velocity. (c) Success rates for acceleration. (d) PTR values for the successful displacement results in (a). (e) PTR values for the successful velocity results in (b). (f) PTR values for the successful acceleration results in (c).

The amplitude of the random variable is constant at $\sigma = 0.5$, the pulse width at 5 points, and the pulse amplitude at 3.5. The start times, t_p, were chosen to be at different locations in a single period of the sinusoidal driving force. Figure 8 shows the locations plotted on a sine wave with a frequency of 1.2 rad.

Figure 9 displays the results using the same format as Figs. 5 and 7, with the event start time given on the abscissa. It can be observed that the values are affected, but the differences from test point to test point are slight for both the IIF and HExp. Variance appears to be more affected by location. The trend in success rate between IIF, Variance, and HExp remains unchanged. The IIF has higher success rates, which is expected from the results of Experiments I and II. The PTR ratios have significant IQR overlap, showing no clear superiority of any one metric over any other.

Unsurprisingly, all three methods are challenged by high levels and very small events, as shown in Experiments I and II. Under these more challenging conditions, the differences in approach for each metric, as depicted in Fig. 6(a), are apparent. Table I lists notional characteristics that should produce successful detection and the counter characteristic that could cause a particular method to underperform or fail given the same time-series.

These failure mechanisms are notional but supported by the observations given throughout the Results and Discussion section. As an artifact of its derivation, the IIF is sensitive to multiple pulses of similar amplitude because IIF values are weighted against the total. Support for the notion that random excitation can produce loss in performance is in both Experiment I and II, where the IIF did not see significant degradation in performance until the momentum of the event size was less than a factor of 1.5 times the rms of the random noise. Likewise, a relative insensitivity to pulse location supports the idea that the variation in performance can mostly be attributed to the random excitation. Running Variance showed similar trends with a lower performance threshold. Figure 4(c) demonstrated the difference in mechanism. In that case, the event was obscured by the presence of large spikes manifested by nonlinear behavior under random excitation. The HExp and IIF, in that example, were shown not to be as susceptible to the situation. Presumably, the additional spikes in the example had large amplitudes like the target event but did not share other characteristics such as a sharp rise. Results in Experiment II confirmed that the HExp is sensitive to pulse shape.

V. CONCLUSIONS

The IIF easily detects events against a mixed random and chaotic background, for most of the test cases. Experiment I showed that the IIF was less sensitive to random excitation than the HExp or running Variance for similar time resolution. Experiment II showed that, unlike the HExp, the IIF was more sensitive to the pulse momentum than its dimensions. Experiment III showed that the IIF and HExp techniques were less sensitive to pulse location than the Variance. The behavior of the IIF was closer to the Variance in Experiment II, where it was more similar to the HExp in Experiment I. This suggests that the IIF can provide similar event detection coverage in place of the HExp and Variance for most cases.

The success of the IIF could be attributed to a couple of particulars in the IIF concept. Since IIF examines normalized changes in time-dependent behavior, it will produce higher weight to small changes than an amplitude-based metric. The other is that the record's entirety is considered, and consistent or repeated behaviors will have lower IIF values even if they are large or nonlinear. While the HExp also had high success rates, the IIF did not require any filtering or smoothing to achieve these results. The HExp also consistently had low PTR values, which indicates that the found pulses may be less evident than desired.

All three of the methods presented in this paper represent very different approaches to transient signal detection and could be complimentary for some applications. The Variance is a statistical measure that operates on amplitude, so it is sensitive to magnitude changes. The HExp is a measure of smoothness, so it is more sensitive to pulse shape than the others. The IIF measures information change and its values are weighted against the behavior of the full record. The difference between these three methods is only apparent at the more extreme test cases, where the efficacy of all three methods was less than 100%. Including all three methods can drastically improve the success rate in those cases. If the methods were similar, that would not be the case.

Overall, these results illustrate that the IIF holds promise as a mathematical tool for interrogating nonlinear dynamical systems. Our future work includes the development of an analysis methodology using the IIF that can monitor for a range of event sizes and a demonstration of event detection on field data with unknown mixed random excitation. It is our hope that this method will become an essential part of automated monitoring systems for complex systems, streamlining analysis and increasing design confidence.

ACKNOWLEDGMENTS

This work is part of the Advanced Simulation and Computing (ASC) initiative at Sandia National Laboratories. Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under Contract No. DE-NA0003525. This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

A. Montoya: Conceptualization (lead); Formal analysis (lead); Methodology (equal); Resources (lead); Software (lead); Writing – original draft (equal); Writing – review and editing (equal). **E. Habtour:** Conceptualization (equal); Methodology (equal); Supervision (equal); Writing – original draft (equal); Writing – review and editing (equal). **F. Moreu:** Project administration (lead); Supervision

(equal); Writing – original draft (supporting); Writing – review and editing (supporting).

DATA AVAILABILITY

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

REFERENCES

[1]M. Mishra, Int. Trans. Electr. Energy Syst. **29**, e12008 (2019).
[2]D. Zhou, Y. Zhao, Z. Wang, X. He, and M. Gao, IEEE Trans. Ind. Electron. **67**, 2337 (2020).
[3]U. R. Acharya, Y. Hagiwara, J. E. W. Koh, S. L. Oh, J. H. Tan, M. Adam, and R. S. Tan, Biocybern. Biomed. Eng. **38**, 373 (2018).
[4]G. Liu, C. Li, Z. Peng, X. Li, and J. Wu, Geophys. Res. Lett. **44**, 4829, https://doi.org/10.1002/2017GL072511 (2017).
[5]M. Rabbani, Y. Wang, R. Khoshkangini, H. Jelodar, R. Zhao, S. Bagheri Baba Ahmadi, and S. Ayobi, Entropy **23**, 529 (2021).
[6]F. Wolf, J. Bauer, and N. Boers, Chaos **30**, 033102 (2020).
[7]Z. Y. Wan, P. Vlachas, P. Koumoutsakos, and T. Sapsis, PLoS One **13**(5), e0197704 (2018).
[8]R. J. Webber, D. A. Plotkin, M. E. O'Neill, D. S. Abbot, and J. Weare, Chaos **29**, 053109 (2019).
[9]A. Blanchard and T. Sapsis, Ocean Eng. **243**, 110242 (2022).
[10]T. B. Francis, K. C. Abbott, K. Cuddington, G. Gellner, A. Hastings, Y. C. Lai, A. Morozov, S. Petrovskii, and M. L. Zeeman, Nat. Ecol. Evol. **5**(3), 285 (2021).
[11]A. Hastings, K. C. Abbott, K. Cuddington, T. B. Francis, Y. C. Lai, A. Morozov, S. Petrovskii, and M. Lou Zeeman, J. R. Soc. Interface **18** (180), 20210257 (2021).
[12]K. A. Ryu, C. M. Kaszuba, N. B. Bissonnette, R. C. Oslund, and O. O. Fadeyi, Nat. Rev. Chem. **5**, 322 (2021).
[13]N. Manini, O. M. Braun, E. Tosatti, R. Guerra, and A. Vanossi, J. Phys.: Condens. Matter **28**, 293001 (2016).
[14]M. S. Allen, B. J. Deaner, and D. J. Segalman, in *The Mechanics of Jointed Structures: Recent Research and Open Challenges for Developing Predictive Models for Structural Dynamics* (Springer, 2018), p. 255.
[15]E. M. Habtour, D. P. Cole, C. M. Kube, T. C. Henry, R. A. Haynes, F. Gardea, T. Sano, and T. Tinga, J. Intell. Mater. Syst. Struct. **30**, 1355 (2019).
[16]M. Jin, M. R. W. Brake, and H. Song, J. Sound Vib. **453**, 268 (2019).
[17]A. G. Piersol and T. L. Paez, *Harris' Shock and Vibration Handbook* (McGraw Hill Education, 2010).
[18]C. R. Farrar and K. Worden, *Structural Health Monitoring: A Machine Learning Perspective* (Wiley, 2012).
[19]P. H. Wirsching, T. L. Paez, and K. Ortiz, *Random VibrationsâĂŕ Theory and Practice* (Dover Publications, Mineola, NY, 2006).
[20]N. Muhammad, Z. Fang, and M. Shoaib, Microelectron. Reliab. **107**, 113614 (2020).
[21]D. Capriglione, M. Carratù, M. Catelani, L. Ciani, G. Patrizi, A. Pietrosanto, L. Signorini, R. Singuaroli, and P. Sommella, Meas. Sensors **18**, 100104 (2021).
[22]O. Artime, J. Ramasco, and M. San Miguel, Sci. Rep. **7**, 41627 (2017).
[23]V. Belenky, D. Glotzer, V. Pipiras, and T. P. Sapsis, Probab. Eng. Mech. **57**, 1 (2019).
[24]S. Vantadori, R. Haynes, G. Fortese, E. Habtour, C. Ronchei, D. Scorza, and A. Zanichelli, Fatigue Fract. Eng. Mater. Struct. **41**, 20 (2018).
[25]S. Guth and T. Sapsis, "Analytic methods for estimating the effects of stochastic intermittent loading on fatigue-crack nucleation," in *Advances in Nonlinear Dynamics*, NODYCON Conference Proceedings Series, edited by W. Lacarbonara, B. Balachandran, M. J. Leamy, J. Ma, J. A. Tenreiro Machado, and G. Stepan (Springer, 2022), pp. 407–420.
[26]T. Vairo, P. Gualeni, A. P. Reverberi, and B. Fabiano, Sustainability **13**, 6836 (2021).
[27]E. Habtour, D. Di Maio, T. Masmeijer, L. C. Gonzalez, and T. Tinga, J. Nondestruct. Eval. Diagnostics Prognostics Eng. Syst. **5**(2), 021005 (2022).
[28]B. Peele, S. Li, C. Larson, J. Cortell, E. Habtour, and R. Shepherd, Soft Rob. **6**, 142 (2019), see https://Home.Liebertpub.Com/Soro.
[29]M. Ernst, E. Habtour, and A. Dasgupta, IEEE Trans. Compon. Packag. Manuf. Technol. **6**, 561 (2016).
[30]A. Del Carre, P. Teixeira, R. Palacios, and C. E. S. Cesnik, "Nonlinear response of a very flexible aircraft under lateral gust. Modeling for design of highly flexible aircraft II," in *Forum of Aeroelasticity and Structural Dynamics IFASD 2019, 9–13 June 2019*, edited by A. De Gaspari, E. Dowell, and S. Ricci (Journal of Aeorelasticity and Structural Dynamics, 2019), IFASD-2019-090.
[31]E. Kantor, D. E. Raveh, and R. Cavallaro, AIAA J. **57**, 2158 (2019).
[32]Y. Mahgoub and A. El-Badawy, Nonlinear Dyn. **108**, 315 (2022).
[33]D. Qi and J. Harlim, Philos. Trans. R. Soc. London A **380** (2229), 20210205 (2022).
[34]C. T. Nguyen, V. Langlois, J. Guilleminot, F. Detrez, A. Duval, M. Bornert, P. Aimedieu, and C. Perrot, Int. J. Solids Struct. **249**, 111684 (2022).
[35]G. Fasulo, P. Vitiello, L. Federico, and R. Citarella, Aerospace **9**, 327 (2022).
[36]B. Qiu, X. Zhang, S. Xia, T. Sun, Y. Ling, S. Zhou, H. Guang, Y. Chen, Z. Xu, M. Liang, and H. Zou, Compos. Part A **155**, 106811 (2022).
[37]A. Montoya, E. Habtour, and F. Moreu, Entropy **22**, 1199 (2020).
[38]Z. Chen, F. Li, L. Zhu, and Y. Xing, J. Syst. Sci. Complex **1**, 1 (2021).
[39]N. Hao, Y. S. Niu, and H. Xiao, "Equivariant variance estimation for multiple change-point model," arXiv:2108.09431.
[40]Z. Liu, M. Perrodin, T. Chambrion, and R. S. Stoica, "Windowed total variation denoising and noise variance monitoring," arxiv.2101.11850 https://doi.org/10.48550/arXiv.2108.09431
[41]N. Pnevmatikos, F. Konstandakopoulou, B. Blachowski, G. Papavasileiou, and P. Broukos, Soil Dyn. Earthq. Eng. **139**, 106328 (2020).
[42]A. M. Homborg, P. J. Oonincx, and J. M. C. Mol, Corrosion **74**, 1001 (2018).
[43]C. Zhou, Z. Jiang, C. Sun, and Z. Zhu, IEEE Access **8**, 96661 (2020).
[44]A. N. Kolmogorov, IEEE Trans. Inf. Theory **14**, 662 (1968).
[45]H. C. Andrews and C. L. Patterson, IEEE Trans. Commun. **24**, 425 (1976).
[46]A. Stallone, A. Cicone, and M. Materassi, Sci. Rep. **10**, 15161 (2020).
[47]C. Eckart and G. Young, Psychometrika **1**, 211 (1936).
[48]C. Truong, L. Oudre, and N. Vayatis, Signal Process. **167**, 107299 (2020).
[49]H. Sohn, A. N. Robertson, and C. R. Farrar, Struct. Eng. Mech. **17**, 409 (2004).
[50]Z. R. Struzik, Physica A **296**, 307 (2001).
[51]C. Sisemore and V. Babuška, *The Science and Engineering of Mechanical Shock* (Springer International Publishing, Cham, 2020).
[52]W. Y. Tseng and J. Dugundji, J. Appl. Mech. **37**, 292 (1970).
[53]H. Ogura, Y. Ueda, and Y. Yoshida, Prog. Theor. Phys. **66**(6), 2280 (1981).
[54]N. D. Anh and N. N. Hieu, Probab. Eng. Mech. **30**, 27 (2012).
[55]K. K. Dey and G. A. Sekh, J. Stat. Phys. **182**, 1 (2021).