Symbolic approach for measuring temporal ''irreversibility''

C. S. Daw*
Engineering Technology Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831-8088

C. E. A. Finney
College of Engineering, University of Tennessee, Knoxville, Tennessee 37996-2210

M. B. Kennel
Institute for Nonlinear Science, University of California, San Diego, La Jolla, California 92093-0402
(Received 27 August 1999)

We describe a symbolic approach for measuring temporal ''irreversibility'' in time-series measurements. Temporal irreversibility is important because it excludes Gaussian linear dynamics and static transformations of such dynamics from the set of possible generating processes. A symbolic method for measuring temporal irreversibility is attractive because it is computationally efficient, robust to noise, and simplifies statistical analysis of confidence limits. We propose a specific algorithm, called ''false flipped symbols,'' for establishing the presence of temporal irreversibility without the need for generating surrogate data. Besides characterizing experimental data, our results are relevant to the question of selecting alternative models. We illustrate our points with numerical model output and experimental measurements.

PACS number(s): 05.45.Tp

I. BACKGROUND

Temporal irreversibility (TI) has become important in characterizing natural phenomena and observed time series [1–11]. By temporal irreversibility, we refer to statistical properties which differ depending on whether one observes the behavior as time proceeds in its natural direction or in reverse. Confirmation of TI is important because it implies the influence of nonlinear dynamics, non-Gaussian noise, or both [8]. In dissipative nonlinear dynamics, for instance, the positive and negative Lyapunov exponents differ in magnitude, not just in sign, as the physical origin of instability in the forward direction is distinct from dissipation (which appears as a virtual instability under temporal reversal). Significant TI automatically excludes Gaussian linear processes (or static nonlinear transformations of such processes) as possible models for the generating dynamics. All linear surrogate data sets produced via the Fourier-transform methods fall in the class of statistically reversible data, thus rejecting temporal reversibility precludes this entire class. In the present context of stochastic processes, temporal reversibility means that any sequence of measurements is exactly as likely as its time reverse to be observed from a given source.

We quantify TI via symbolization of time-series measurements. Symbolization converts continuous-valued time-series measurements into a stream of discrete symbols. Typically, the range of the observed variable is partitioned into a finite number of bins, such that all measurements falling within a given bin are transformed into the same symbolic value. The objective in making such a transformation is to faithfully preserve dominant dynamical features while simplifying and speeding up subsequent computations. Empirical symbolic techniques are known to be useful for chaotic and/or turbulent data and for reconciling noisy data with nonlinear models, among other applications [11–22].

In the following discussion, we describe our general approach for symbolizing data and then identifying and quantifying temporally irreversible portions of the symbol stream. We also describe techniques for evaluating confidence intervals that appear to give good discrimination for known data. Finally, we illustrate how TI can be used to detect noisy bifurcations and characteristic dynamical transitions.

II. SYMBOLIZATION

Figure 1 illustrates how symbolization is achieved by partitioning the data range and assigning a symbol to each measurement based on the region into which it falls. Here we use a single partition line to simplify the discussion. The number of partitions defines the symbol set, often called an alphabet (in this case 0 or 1). When using more than two symbols, our usual convention is to define symbols as integers 0 through $n-1$, where n is the size of the alphabet.

Theoretically, there is a generating partition for each deterministic process that provides a unique, one-to-one mapping of trajectories into symbol sequences [12,13]. However, generating partitions do not generally exist for realistic experimental data because of the effects of noise. As Crutchfield and Packard observed [12], generating partitions do not exist for noisy systems, even in principle. Nevertheless, lack of a generating partition does not prevent the characterization of dynamical complexity using reasonable empirical partition choices, as demonstrated by others [15–19]. For our purposes, we usually define the symbolic partition boundaries such that the scalar measurement range is divided into equiprobable regions. This choice of boundaries is not required for TI detection, but it is convenient because all pos-

*Present address: Engineering Technology Division, Oak Ridge National Laboratory, P.O. Box 2009, Oak Ridge, TN 37831-8088. Electronic address: dawcs@ornl.gov

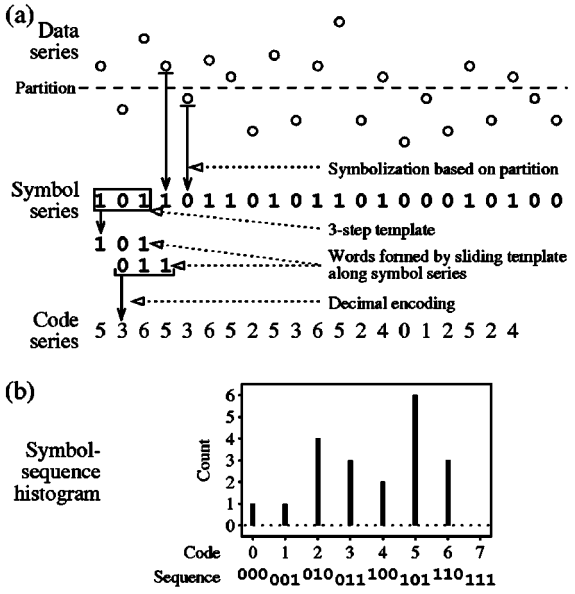

FIG. 1. Diagram illustrating data symbolization (a) and symbol-sequence histogram (b). Data are symbolized based on their value relative to boundaries that partition the data range to produce a symbol series. Words are formed by defining a finite-length template to group consecutive symbols, and this template is shifted along the symbol series. Each word is encoded into its decimal (base-10) equivalent to form a code series. A symbol-sequence histogram is a tally of the absolute counts or relative frequencies of each word.

sible symbol sequences become equally probable in the limit for independent, identically distributed data.

Sequences of symbols, often called words, contain the information about dynamics. One chooses the number of sequential symbols to group together (the word length) and then observes the specific words that are formed by moving a template of this size stepwise along the symbolized time series. This is conceptually similar to time-delay embedding, with discrete symbols instead of continuously valued original measurements. Unfortunately, there does not appear to be any obvious analogue of the geometric construction underlying time-delay embedding, so there is not necessarily any finite word length at which ''all'' information about the process has been retrieved, even with a generating partition. Note that the time-delay reconstruction theorems promise faithful dynamical reconstruction only up to smooth transformations of the original state space; in ''symbol space,'' continuity and differentiability are not useful concepts.

As with time-delay embedding, converting the original time-series data to symbol sequences requires specification of transformation parameters, specifically, the number of symbols in the alphabet, the number of symbols in each word, and the time interval between symbols in each word. For purposes of this discussion, we assume that the data of interest are generated by discrete maps, so that the appropriate intersymbol interval is one step or iterate.

Alphabet size and word length combine to determine the total possible number of sequences N_{seq} that can be represented. In general,

$$N_{\mathrm{seq}}=n^{L}, \tag{1}$$

where n is the number of symbols and L is the fixed word length. With finite data, there is a tradeoff between using large n and L, which improve resolution of the state space and dynamical evolution, and small n and L, which reduce statistical fluctuations in any word bin. Each word represents a unique permutation of symbols and can be represented as a specific number in base n. This convention for designating words makes it possible to display frequency patterns as depicted in Fig. 1(b), which we refer to as a symbol-sequence histogram (SSH).

III. SYMBOL-SEQUENCE IRREVERSIBILITY

We quantify TI by comparing the observed number of occurrences of particular symbol words to the frequencies of their counterparts in reverse time. An important caveat is that for our purposes TI is meaningful only in the context of stationarity: nonstationary dynamics are trivially irreversible, linear or not. Thus one should verify statistical stationarity in addition to TI (e.g., see Refs. [23–25]).

The choice of symbolization parameters affects the ability to distinguish irreversible patterns. For example, a binary alphabet combined into two-symbol words has the following possibilities: 0 0, 0 1, 1 0, and 1 1. The words 0 0 and 1 1 are symmetric and cannot contribute to a TI statistic. The relative fraction of symmetric words for any given n and L is

$$\mathrm{f}_{\mathrm{sym}}=n^{m-L}, \tag{2}$$

where $\mathrm{f}_{\mathrm{sym}}$ is the fraction of the total possible words which are symmetric and $m=\lfloor (L+1)/2 \rfloor$, where the floor function $\lfloor \cdot \rfloor$ evaluates the next lowest integer. The fraction of asymmetric words which can contribute to detecting TI is thus $1-\mathrm{f}_{\mathrm{sym}}$. The fraction of words able to reveal TI increases significantly with both n and L. Consequently, we find it desirable to use larger alphabet sizes (e.g., $4 \leqslant n \leqslant 10$), in contrast to the usual binary alphabets used by other investigators such as Tang *et al.* [15] and Kurths *et al.* [17]. We illustrate the impact of larger alphabet sizes in our examples below.

Temporal irreversibility becomes visible by comparing the symbol-sequence histograms for forward and backward realizations of the observed data. A simple example is illustrated in Fig. 2, which depicts the forward and backward histograms for 3000 observations of x from the Hénon map:

$$x[i+1]=y[i]+1-ax[i]^{2}, \tag{3}$$

$$y[i+1]=bx[i]. \tag{4}$$

The selected parameter values are $a=1.4$ and $b=0.3$, which produce chaos. The symbolization parameters are $\{n=4,\ L=3\}$. Differences between the two histograms (forward realization frequency minus backward realization frequency for each word) reveals that a few specific words contribute most to irreversibility, as illustrated in Fig. 2(b). Note that both positive and negative versions of each difference appear, as each histogram accounts for every word and its time inverse. For statistical analyses shown later, we avoid this double counting.

Figure 2(b) also shows the difference results for a temporally symmetric surrogate of the Hénon example. The surro-

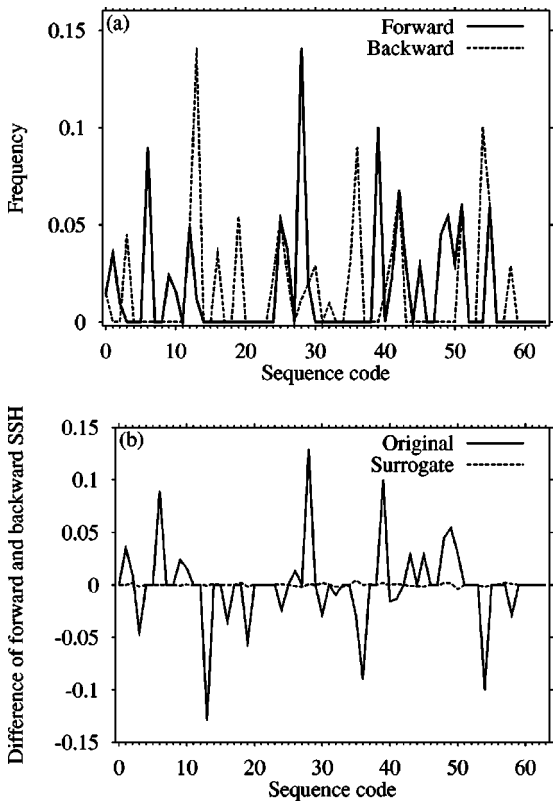

FIG. 2. (a) Forward- and backward-time symbol-sequence histograms for chaotic output from the discrete Hénon model. Symbolization parameters are $\{n=4,\ L=3\}$. (b) Differences between the forward and backward symbol sequences for the chaotic Hénon map. Large peaks (positive or negative) indicate a strong contribution to irreversibility. The dashed line illustrates the result for a temporally reversible surrogate.

gate is constructed by adding (in the continuous scalar observable space) independent orbit segments read in forward and reverse time, as suggested by Diks *et al.* [6]. A stationary linear stochastic Gaussian source is invariant under such transformation. Other standard Fourier-transform surrogate-generation methods such as those suggested by Theiler *et al.* [26] and Schreiber and Schmitz [27] work equally well since they also force temporal symmetry. The small differences in Fig. 2(b) for the surrogate are due to finite-sample statistics.

One may quantify forward-reverse differences with simple statistics such as

$$T_{\mathrm{fb}}=\sqrt{\sum_i\ (P_{\mathrm{f},i}-P_{\mathrm{b},i})^2} \tag{5}$$

and

$$\chi^2_{\mathrm{fb}}=\sum_i \frac{N(P_{\mathrm{f},i}-P_{\mathrm{b},i})^2}{P_{\mathrm{f},i}+P_{\mathrm{b},i}}. \tag{6}$$

In Eq. (5), T_{fb} is the Euclidean norm between the forward and backward histograms and is analogous to the comparative statistic used by Tang *et al.* for fitting nonlinear models to observed data [15]. In Eq. (6), χ^2_{fb} is a similar chi-square statistic applied to the two symbol-sequence histograms.

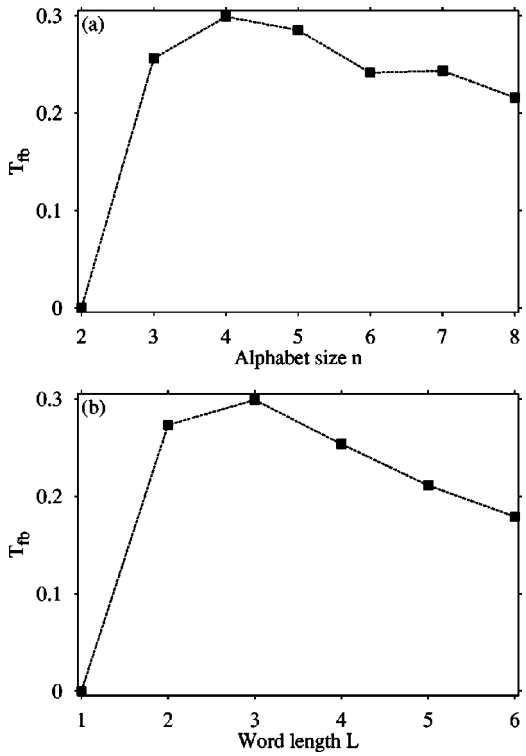

FIG. 3. Variation of observed irreversibility (T_{fb}) with (a) alphabet size (for fixed $L=3$) and (b) word length (for fixed $n=2$) for the Hénon-map data.

Direct estimation of confidence limits using the above simple statistics is problematic because of correlations between word frequencies. However, when such statistics are repeatedly evaluated for many temporally reversible surrogates, it is possible to develop confidence estimates by comparing the distribution of the surrogate statistics with the statistics for the original data. Such comparisons are further strengthened if many realizations of the original data (repeated observations) are also available. As we discuss later, a more preferable situation would be to define a statistic for which confidence limits are available *a priori*, thus avoiding the effort of constructing surrogates.

We typically determine ''optimal'' symbolization parameters for a given data set through systematic evaluation of irreversibility statistics as alphabet size and word length are varied. Figure 3 illustrates this process for the Hénon-map data described earlier. Symbolization parameters chosen this way are not in any sense universal because they reflect case-specific factors such as noise level, the symbol partitioning scheme, the dominant dynamical features present, and data-set size. With the same data, other features such as entropy may be most effectively revealed with different choices of parameters. Thus, to a great extent, the choice of symbolization parameters depends on one's specific needs and constraints. Because selection of a generating partition is not possible, we do not expect any direct connection with geometric properties of the time-delay phase space (e.g., minimum embedding dimension).

IV. FALSE FLIPPED SYMBOLS—DIRECT ESTIMATION OF SIGNIFICANCE

Although surrogates can be useful for relative comparisons of T_{fb} and χ^2_{fb}, it is desirable to have statistics that do

not require generation of surrogates to evaluate significance. There may be technical issues in generating truly faithful surrogates or, equivalently, questions about the actual null hypothesis represented by the surrogate generator. Moreover, a direct test, especially in symbol space, is computationally simpler and could be embedded in an online instrument or control system.

Here we propose an *a priori* statistic [referred to hereafter as false flipped symbols (FFS)]. Our statistic is constructed using a classical binomial model to find the likelihood of obtaining the observed distribution of words under the null hypothesis that the forward and backward realizations of each word are equally probable. If then we find that the observed distribution is highly unlikely, we can reject the null that the original data are temporally reversible.

For the moment we concentrate on a single word class; the class encompasses the word and its time inverse. We count the frequency that this word was observed in the forward direction c_f and the backward direction c_b. Intrinsically reversible symbol words (e.g., 1 0 0 1) trivially maintain $c_f = c_b$ and are excluded from further calculation. For the rest, under the null hypothesis of reversibility, it is as likely to observe any given word in the forward as in the backward direction. This suggests using the classical two-sample binomial test to provide the likelihood of seeing the observed values of c_f and c_b given that each event is equiprobable.

Given N total draws with probability of an event θ, the binomial probability of seeing exactly k events is

$$b(k,N,\theta) = \frac{N!}{k!(N-k)!}\,\theta^k(1-\theta)^{N-k}. \tag{7}$$

For our purposes, we are interested in the tail probabilities of observing extreme cases, i.e., where the number of forward or backward observations of one or more words is unexpectedly high or low. Given the null, the likelihood of observing k or fewer occurrences of a word (or $N-k$ or more occurrences of its time inverse) is

$$B(k,N,\theta) = \sum_{j=0}^{k} b(j,N,\theta) + \sum_{j=N-k}^{N} b(j,N,\theta). \tag{8}$$

Thus the likelihood L of accepting the null when it is actually true is

$$B(c_f - 1,N,0.5) \leqslant L \leqslant B(c_f,N,0.5), \tag{9}$$

with $N = c_f + c_b$. The inequality is on account of the discrete nature of the distribution. If one were simulating the accuracy of the null by generating reversible time series and evaluating L for each, one would select a particular value of L randomly and uniformly between the limits above. In the limit of many reversible time series, the distribution of these L's converges to uniform (0,1), as expected. Given but a single test on a single data set, it is conventional to be conservative against rejecting the null by using the upper bound, though one should recognize this would not converge to a uniform distribution under the null distribution, and is a heuristic bias.

The binomial test described above assumes that the events occur independently of one another. With realistic data, this assumption is unlikely always to be valid because of temporal correlations between successive observations. If temporal correlations exist, the actual degrees of freedom for evaluating the significance are unknown. Thus, we perform another test, looking for sequential correlations in the occurrences of forward and backward versions of a given word. If there exist statistically significant correlations here, the result of the reversibility test is deemed to be unfaithful. Thus, a designation that an observed time series is irreversible requires that the reversibility test reject its null and also that the sequential independence test accept its null.

We see if forward and backward events appear to occur without sequential correlation. Specifically, whenever an eligible target word class occurs, we recall the immediately previously observed orientation of that same word and increment counts c_s and c_d, depending on whether that word occurred in the same (s) or different (d) orientation from its predecessor. By construction, $c_s + c_d + 1 = c_f + c_b$ for any word class. If the relative frequency of forward words is $\theta_f = c_f/(c_f + c_b)$, then under sequential independence the expected proportion of same-to-same transitions is $\theta_s = \theta_f^2 + (1-\theta_f)^2$. We thus find the likelihood L_c, under the null of no correlation, of observing this particular c_s and c_d as

$$B(c_s - 1, c_s + c_d, \theta_s) \leqslant L_c \leqslant B(c_s, c_s + c_d, \theta_s). \tag{10}$$

Thus we require that L_c not reject its null.

What if it does reject? Our approach to this problem is selectively to down-sample the total set of events so that correlations are minimized. We have considered two methods for implementing selective sampling.

(1) Random subsampling of all events with increasing sparseness, e.g., randomly subsample a fraction $p < 1$ of the words in the symbolized time series to qualify for any statistic.

(2) Given any word class, ignore subsequent occurrences of that class which occur sufficiently closely in time. This decorrelation-interval approach is common in many nonlinear time-series algorithms.

For both approaches, one excludes more data until the correlation indicator statistic indicates the null of independence is sufficiently likely, and then examines the result of the reversibility test.

V. TARGETED FALSE FLIPPED SYMBOLS

We have described the detection of TI for a single word class. We want to produce a statistic for the whole data set: as seen in Fig. 2(a), some words much more strongly indicate irreversibility than others. We give two alternatives: combining the result of many tests of different words, or selection of a particular word which best demonstrates irreversibility.

Assume we have computed likelihoods L_i regarding reversibility for the various word classes enumerated by i, with particular L_i selected uniformly between the bounds shown above. With N_w independent tests of a null hypothesis, the quantity

$$X^2=\sum_{i=1}^{N_w} -2\ln L_i \tag{11}$$

is χ^2 distributed, which with standard algorithms we compute our final $\mathcal{L}$, again uniform in (0,1) under the null. Especially small values of $\mathcal{L}$ imply a small likelihood that this apparent level of irreversibility would have been observed had the symbol stream been generated by a reversible dynamical process. Small $\mathcal{L}$ signals irreversibility.

This sort of full implementation of FFS is potentially very powerful because it can reflect the cumulative indication of TI in all symbol words. However, a general scheme for collecting representative uncorrelated samples of all symbol words is difficult to implement using the fixed-word-length symbolization approach described here, because not all the events entering counts corresponding to *different* word classes are independent. We are developing a version of this sort of FFS using a more sophisticated state-selecting algorithm, an adaptive variable-order ''context tree'' model suggested by data-compression methods, which does in fact produce nearly independent counts. This construction has already been used for a symbolic stationarity test by one of the authors [28].

At present, we discuss a modified FFS method, which we call ''targeted FFS'' (TFFS). Initially we divide the time series into two equal halves. In the front half of the data set, we note which symbol word occurs most frequently relative to its time inverse, the word class for which the difference c_f-c_b is largest, where for each word class $c_f>c_b$. This word then serves as our first target. Note that it emphasizes words which are both asymmetric and occur frequently. Other choices, such as maximizing $(c_f-c_b)^2/(c_f+c_b)$ are reasonable.

We then note $c_{f,1}$ and $c_{b,1}$ for this same symbol word in the second half of the data set. For this second counting, we eliminate from consideration any repeated occurrences of the word class within a specified decorrelation interval. To begin, we estimate this interval from the autocorrelation function for the data. Later, we apply our correlation indicator statistic (described below) to confirm that our decorrelation interval is sufficient.

We next identify a second target word class using the second half of the data in a manner analogous to our initial step. In some cases this second target word may be the same as the first target sequence, or it may be different on account of random fluctuation. The latter is typically true for temporally reversible data.

We now note $c_{f,2}$ and $c_{b,2}$ for this second target word class when it occurs in the first half of the data. As before, we eliminate any repeated occurrences within the decorrelation interval.

A coin analogy helps to explain the rationale behind the above procedure. Suppose one has a collection of coins and wishes to test the null hypothesis that all the coins are balanced (corresponding to reversible data). One could begin with a trial run in which all of the coins were initially flipped a number of times. A good candidate for testing the null in a second flipping session would be that coin which produced the greatest difference between the numbers of heads and tails. If the selected coin still acts unbalanced in this second set of trials, we can confidently reject the null.

With the c_f and c_b results for both target words, it is then possible to evaluate the case for TI by considering the likelihood that the observed number of forward and backward counts would be seen under a time-symmetric null. Note that for complete temporal symmetry, the probability of observing the forward occurrence of both target words is 0.5. Unlikely counts for either or both targets are grounds for rejection of the null. We consider both the counts for the front and back targets separately and then the combined counts. Specifically, we evaluate the likelihood L_1 of observing the forward target count in the front half of the data as

$$B(c_{f,1}-1,N_1,0.5)\leqslant L_1\leqslant B(c_{f,1},N_1,0.5), \tag{12}$$

where $N_1=c_{f,1}+c_{b,1}$, this time with the one-sided B,

$$B(k,N,\theta)=\sum_{j=N-k}^{N} b(j,N,\theta). \tag{13}$$

Typically, we reject the null if $L\leqslant 5\%$.

When N_1 is large (e.g., >100), the normal approximation for the binomial distribution can be used, and we estimate the normal deviate statistic

$$Z_{\mathrm{irr},1}=\frac{c_{f,1}-(1/2)N_1}{\sqrt{(1/4)N_1}}. \tag{14}$$

In this case, we define our confidence level based on the one-tailed, normal-deviate probabilities and reject the null hypothesis when $Z_{\mathrm{irr},1}$ exceeds some limit (e.g., 1.645).

Because an unlikely value for $c_{f,2}$ is also grounds for rejection of the null, we repeat the previous evaluation to estimate L_2 and $Z_{\mathrm{irr},2}$, where $c_{f,1}$ is replaced by $c_{f,2}$, $c_{b,1}$ is replaced by $c_{b,2}$, and N_1 is replaced by $N_2=c_{f,2}+c_{b,2}$. We again reject the null if L_2 is small or $Z_{\mathrm{irr},2}$ exceeds the specified limit.

One possible approach for combining the separate target word counts into a single test would be to apply Eq. (11) to the results of the front and back target tests above. This would produce a chi-square statistic that could be evaluated to decide on rejection of the null. Another approach we have found to be very effective is to consider the total number of forward target events, regardless of which target word is involved. Using this constraint, the expected value for the sum of both forward target counts is $N_{f,t}=\frac{1}{2}(N_1+N_2)$ and the variance of this sum is $N_t/4$ when the null is true. If we can establish that the observed total forward count is unexpectedly high, we can also reject the null. Thus we evaluate

$$B(N_{f,t}-1,N_t,0.5)\leqslant L_t\leqslant B(N_{f,t},N_t,0.5), \tag{15}$$

where $N_{f,t}=c_{f,1}+c_{f,2}$ and $N_t=N_1+N_2$, and

$$Z_{\mathrm{irr},t}=\frac{N_{f,t}-(1/2)N_t}{\sqrt{(1/4)N_t}}. \tag{16}$$

We find that the combined count statistics L_t and $Z_{\mathrm{irr},t}$ typically give the strongest rejection of the null for truly irreversible data due to the larger number of target data points included.

We confirm sufficient decorrelation for the samples used to estimate the above statistics by considering the transitions in sample orientation (runs test) in a similar fashion to the single-target tests. That is, we evaluate

$$B(c_{s,1}-1,N_{c,1},0.5)\leqslant L_{c,1}\leqslant B(c_{s,1},N_{c,1},0.5) \qquad (17)$$

where $N_{c,1}=c_{s,1}+c_{d,1}$ and

$$B(c_{s,2}-1,N_{c,2},0.5)\leqslant L_{c,2}\leqslant B(c_{s,2},N_{c,2},0.5), \qquad (18)$$

where $N_{c,2}=c_{s,2}+c_{d,2}$. If either $L_{c,1}$ or $L_{c,2}$ are sufficiently small, the hypothesis of decorrelation is rejected, the decorrelation interval is increased, the targets are resampled, and the runs test is repeated.

To summarize, we choose particular word classes from each half of the data, measure the occurrences of forward and backward words in the other half, apply the appropriate decorrelation interval, and combine into irreversibility tests [Eqs. (15),(16)] and runs tests [Eqs. (17),(18)]. Significant irreversibility requires small L_1, L_2, and L_t and sufficiently large L_c to validate the test.

If one preselected a particular target word class to examine, this algorithm could be easily implemented as an online measurement, as it only needs to accumulate specific word counts. With a small word buffer, one could sequentially discard the most distant previous observation, providing an easily updated measurement of irreversibility for the recent past.

VI. A NUMERICAL EXAMPLE

We next illustrate the usefulness of TFFS using a more challenging model, the Ikeda ring laser (IRL) [29]. Briefly, this model is a discrete time map coupled to a differential equation, producing a delay differential system. The resulting time-asymptotic dynamics is chaotic (with multiple positive Lyapunov exponents) and has sufficiently high dimension that standard time-series analysis tools such as correlation dimension and false nearest neighbors fail to give useful results [29].

The IRL model output consists of a complex electric-field amplitude $\zeta(t)$ and a spatially averaged population inversion $w(t)$, coupled according to

$$\zeta(t+\tau_R)=E_I e^{i(\omega_I-\omega_0)t}+Be^{i\kappa}\zeta(t)e^{(\beta+i\alpha)w(t)}, \qquad (19)$$

$$\frac{dw(t)}{dt}=Q-2\gamma\{w(t)+1+|\zeta(t)|^2(e^{Gw(t)}-1)/G\}, \qquad (20)$$

where t is time and τ_R is propagation time around the optical ring. Definitions of α, β, γ, κ, ω, G and Q may be found in Ref. [29]. For illustration purposes, we choose the canonical parameter values ($E_I=1$, $B=0.9$, $\kappa=0.4$, $\beta=0$, $\alpha=6$, and $\omega_I=\omega_0$) [29].

We create a discretized measurement by recording the real portion of the field amplitude each time the population inversion changes from negative to positive; we denote this strobed measurement ζ'. A first return map of 10 000 points from the resulting sampled time series is illustrated in the top left corner of Fig. 4. Note that in a low-dimensional projection of this type, most of the high-dimensional dynamics is obscured, and the map is visually almost featureless. Further, there is no obvious asymmetry about the diagonal to indicate TI. It is clear, however, that TI is present for the IRL model output when we apply TFFS to the sampled data. For a symbolization of $\{n=4,\ L=3\}$, we obtain $Z_{\mathrm{irr},t}=9.07$, indicating a very strong rejection of the reversible null. We obtain similarly strong rejections of the null even for other reasonable choices of symbolization parameters such as $\{n=8,\ L=2\}$, $\{n=7,\ L=3\}$, and $\{n=3,\ L=4\}$. We can also use the IRL model to illustrate the robustness of TFFS to the effects of additive noise. In Fig. 4, we have plotted the $Z_{\mathrm{irr},t}$ statistic for the above IRL data after adding successively higher levels of Gaussian measurement noise (with zero mean). The noise levels indicated on the horizontal axis represent the ratio of the additive noise variance to the variance of the model output. Results for 25 noise replicates are shown at each level to illustrate the variability in $Z_{\mathrm{irr},t}$ due to sampling variations in the noise. The solid line indicates the average $Z_{\mathrm{irr},t}$ value for the 25 replicates. Note that rejection of the reversible null continues to be strong up to noise levels of 40%. We attribute the noise robustness of TFFS to the fact that our statistic is based on the most frequent irreversible word, which we expect to be one of the last dynamical features to be obscured as noise is increased.

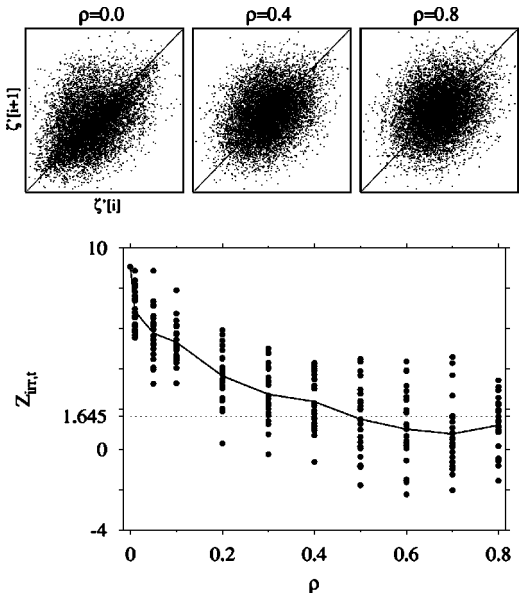

FIG. 4. TFFS results for strobed output from the Ikeda ring laser with 25 realizations of additive noise. Model parameters were set to the canonical values (see text). Noise levels are expressed as the dimensionless ratio, denoted ρ, of the noise variance to the model output variance. Symbolization parameters are $\{n=4,\ L=3\}$ with a decorrelation interval of three iterates. Each return-map scale has been adjusted for maximal visibility.

VII. AN EXPERIMENTAL EXAMPLE

We now illustrate the application of our symbolic method for TI detection to a physical problem involving noisy dynamics. Specifically, there can be significant variations in the combustion output of an internal combustion engine from one cycle to the next. This combustion variation is particu-

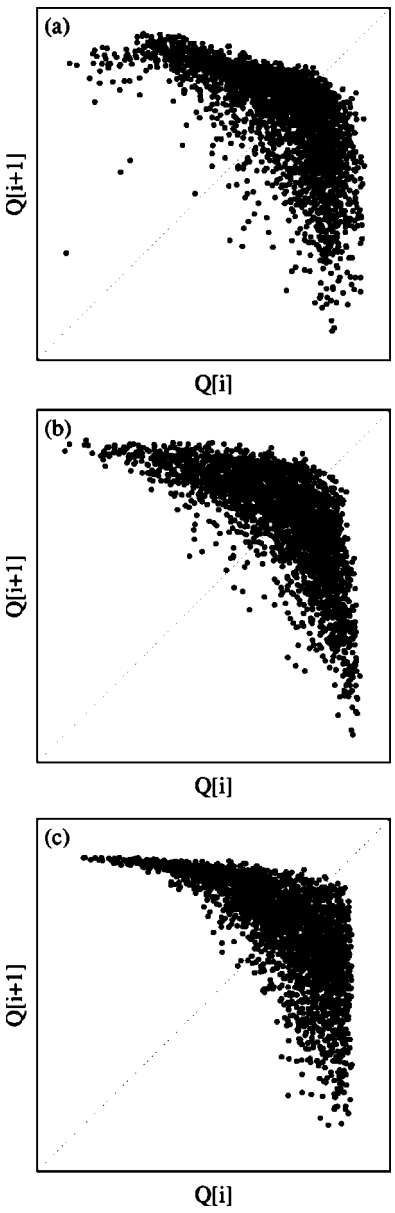

FIG. 5. Typical return maps for engine combustion variations as observed experimentally (a) and predicted by two alternative map models, the ACO model (b), and the NND model (c). The plotted data represent standardized (dimensionless) combustion heat release (Q) for 2800 consecutive engine cycles.

larly important for advanced, low-emission, high-efficiency engines that are designed to operate under lean-fueling conditions (i.e., with an excess of air relative to the amount of fuel injected). Experimentally, such variations are measured in terms of the amount of combustion energy released for each successive power stroke in the engine. As described by Daw *et al.* [21], the observed combustion variations [Fig. 5(a)] can be described in terms of a noisy, nonlinear map [Fig. 5(c)]. A key feature of the map is that high-energy combustion events tend to be followed by low-energy combustion events, and vice versa. The fuzzy appearance of the map is conjectured to arise from noise in both engine parameters and combustion measurements.

Two leading physical models have been proposed to explain the noisy map behavior. The first model explains the origin of combustion variations as the result of linear, anticorrelated oscillations (ACO) in the fueling system. It is hypothesized that Gaussian-distributed flow noise in the air intake is amplified by the linear acoustical characteristics of the intake system to produce anticorrelated pressure waves. These pressure waves subsequently entrain alternately more and less fuel from the injector into the cylinder on succeeding intake strokes. Under this model, the observed combustion changes are then the result of the in-cylinder combustion responding to the fueling variations. The nonlinear appearance of the map [Fig. 5(b)] is due to the (static) nonlinear variation of combustion with fuel level.

The ACO model can be written explicity in terms of the amount of fuel that is injected into the cylinder on each cycle i:

$$Q[i] \propto C[i]m[i], \tag{21}$$

where $Q[i]$ is amount of combustion heat release in cycle i, $m[i]$ is the amount of fuel present in the cylinder in that cycle, and $C[i]$ is the efficiency of the combustion process (i.e., the fraction of fuel burned). Combustion efficiency C is a function of ϕ, the ratio of injected fuel m to injected air a:

$$\phi[i] = R\frac{m[i]}{a[i]}, \tag{22}$$

where the proportionality R defines the stoichiometric equivalence of air to fuel.

A key feature is that C depends in a highly nonlinear way on ϕ because of the large effect of fuel-air ratio on flame speed as the so-called lean limit of combustion is approached [21]. This nonlinearity can be approximated by a sigmoidal function of the form

$$C[i] = C(\phi[i]) = C_{\max}[1 + 100^{-(\phi[i]-\phi_m)/(\phi_u-\phi_l)}]^{-1}, \tag{23}$$

where $C_{\max}$ is the maximal combustion efficiency achieved when fuel and air are stoichiometrically balanced (in this notation, when $\phi = 1$). The quantities ϕ_u and ϕ_l are fitting parameters for the sigmoidal function and $\phi_m = (\phi_u + \phi_l)/2$. The quantity C thus varies between $C_{\max}$ (typically close to 1) and 0, depending on ϕ.

The remaining aspect of the ACO model is that the injected fuel-air ratio is assumed to vary according to

$$\phi[i] = \phi_0 + \phi'[i], \tag{24}$$

where ϕ_0 is a nominal average fuel-air ratio and $\phi'[i]$ is a perturbation in the inlet system. The fuel-air perturbation occurring in each cycle is given by

$$\phi'[i] = (1-\alpha)\phi'[i-1] + \alpha\mathcal{N}(0,\psi^2), \tag{25}$$

where α is a linear filter constant and $\mathcal{N}(0,\psi^2)$ is a random Gaussian deviate of zero mean and standard deviation ψ. Physically, Eqs. (24) and (25) mean that the combustion oscillations are caused directly by variations in the air-fuel ratio, which in turn come from the effects of linearly filtered flow noise in the engine intake system. The filter constant α characterizes the linear response, and the magnitude of the

driving noise is characterized by ψ. Because $\phi'[i]$ is proportional to $\phi'[i-1]$, the combustion oscillates alternately high and low.

Because the nonlinear combustion response described by Eq. (23) is static (i.e., depends only on the present), there can be no onset of bifurcations or chaos. This is a physical example of the general class of processes described by Schreiber and Schmitz [27], and thus is well represented by Fourier-transform surrogates. From the earlier discussion, we also expect that the ACO process should not exhibit significant TI.

The second leading model is based on the hypothesis that the dominant physical process is associated with residual-fuel effects in the cylinder (see Daw *et al.* [21]). Specifically, it is hypothesized that low-energy combustion events leave residual fuel behind in the cylinder, which then contributes to the net fuel charge supplied to the next power stroke. Such a residual-fuel effect creates a nonlinear ''memory,'' which also produces a nonlinear return map [Fig. 5(c)]. When dynamic and measurement noise are suppressed, this map exhibits classical nonlinear dynamics including the onset of period-2 bifurcations as fueling is leaned (i.e., fuel-to-air ratio is reduced). Ultimately, the period-2 bifurcation process recapitulates to a fixed point as fueling is dropped to the point that no combustion occurs. Even when noise is added, the global structure of the fuzzy map is controlled by the noisy bifurcation. The combustion oscillations are thus explained as the result of a noisy bifurcation, and the model is referred to as the noisy nonlinear dynamics (NND) model. As with other bifurcating nonlinear maps, we expect the NND model to exhibit TI.

The fuel and air variation for the NND model can be written explicitly as

$$m[i]=(1-F)m_0+F(1-C[i-1])m[i-1], \qquad (26)$$

$$a[i]=(1-F)a_0+F(1-RC[i-1])m[i-1], \qquad (27)$$

where F is a residual fraction of unburned gas left in the cylinder after each cycle, and m_0 and a_0 are the nominal masses of fuel and air, respectively, injected each cycle. As for the ACO model, combustion heat release is as in Eq. (21), and combustion efficiency depends on ϕ according to Eq. (23). The difference here is that $m[i]$ and $a[i]$ depend in a nonlinear way on past combustion events. Thus, nonlinear memory is introduced, and bifurcations and chaos are possible. Noisy inputs can be introduced into the NND model through Gaussian random perturbations to the injected ϕ and residual-gas fraction, for example,

$$\phi[i]=\phi_0+\mathcal{N}(0,\sigma_\phi^2), \qquad (28)$$

$$F[i]=F_0+\mathcal{N}(0,\sigma_F^2). \qquad (29)$$

The ACO and NND models represent fundamentally different kinds of physical processes. It is important to verify which of these processes dominates actual engine behavior because it could totally change our approach to understanding and dealing with the combustion-variation problem. Specifically, the optimal engineering approaches for controlling/damping linear oscillations versus stabilizing bifurcations and chaos are likely to be very different.

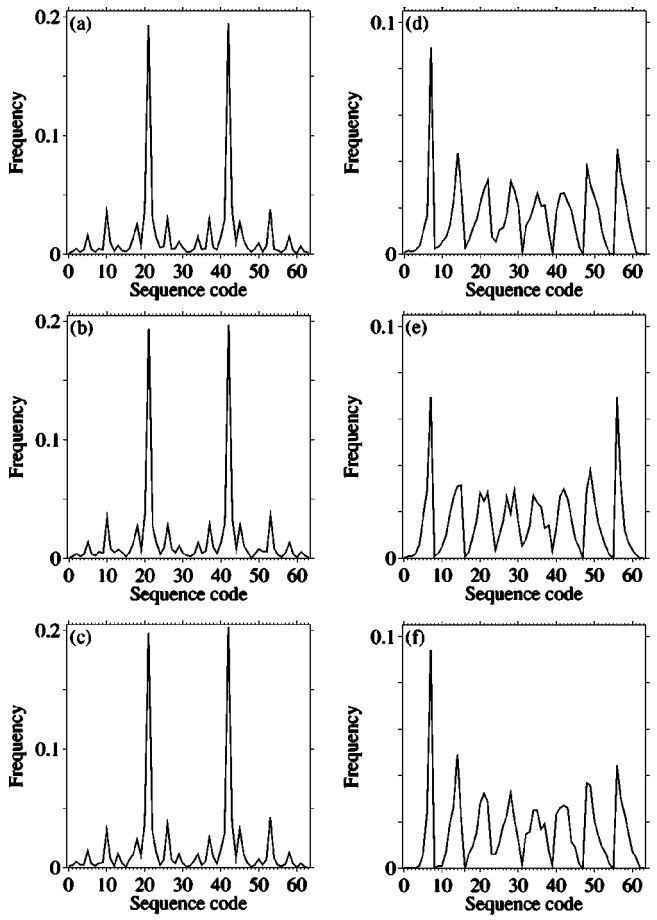

FIG. 6. Comparison of symbol-sequence histograms from the experimental and the ACO and NND model data representing combustion heat release for 2800 consecutive engine cycles. For symbolization with parameters $\{n=2,\ L=6\}$, both the ACO (b) and NND (c) models are nearly indistinguishable from experiment (a). For symbolization with parameters $\{n=8,\ L=2\}$, the discrimination of temporal irreversibility allows better comparison of the ACO (e) and NND (f) models with experiment (d). Within sample-to-sample variability, the ACO model SSH is symmetric, whereas the NND model and experimental SSHs have similar asymmetric features (most notably sequences 7 and 56).

As shown in Fig. 5, it is not possible to readily distinguish the models from return maps alone. Also, binary-alphabet symbolization such as that used by Tang *et al.* [15] is inadequate to distinguish between the models. This inadequacy is illustrated in Fig. 6, which is a comparison symbol-sequence histograms ($\{n=2,\ L=6\}$) for typical realizations of the ACO (b) and NND (c) models fitted with the data (a) depicted in Fig. 5(a). By adjusting the fitting parameters, both models can be made to agree with the experiment within the sample-to-sample error band.

On the other hand, with a choice of symbolization parameters that strongly discriminates TI, we can readily distinguish differences between the models. In Fig. 6, we observe clear differences in the symbol-sequence histograms for the same ACO (e) and NND (f) model outputs compared with the experimental data (d) with symbolization parameters $\{n=8,\ L=2\}$. The biggest differences show up in the relative frequencies of the sequences 0 7 (sequence code 7) and 7 0 (sequence code 56). The sequence 0 7 represents a

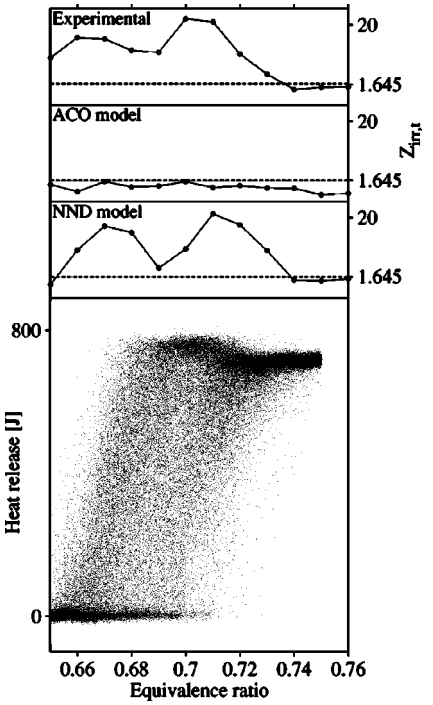

FIG. 7. Temporal irreversibility in engine combustion heat release as the mixture is made more lean. The lower plot was experimentally measured by adjusting the fuel-air equivalent ratio in small increments and recording combustion heat releases for several hundred cycles at each ratio. Corresponding TFFS statistics are plotted for the experimental data as well as corresponding output from the ACO and NND models (see text for parameters). Symbolization parameters are $\{n=4,\ L=3\}$. The broken line at $Z_{\mathrm{irr},t}=1.645$ marks 0.05 significance (the 95% confidence limit).

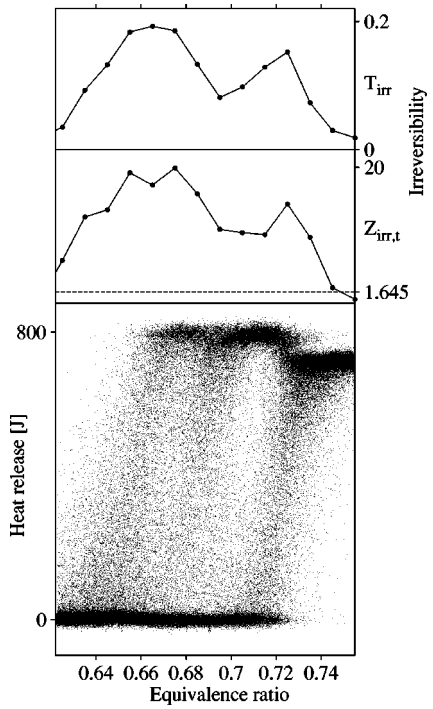

FIG. 8. Temporal irreversibility in combustion heat release as the fuel-air ratio is reduced. The experimental procedure was the same as for Fig. 7 except that the engine was adjusted to have a higher residual gas fraction. The increased level of TI implies a higher degree of bifurcation. Symbolization parameters are $\{n=8,\ L=2\}$. The broken line at $Z_{\mathrm{irr},t}=1.645$ marks 0.05 significance (the 95% confidence limit).

misfire followed by an excessive burn, while 7 0 represents the reverse. For the ACO model, both events occur with equal frequency (within sample variations), but for the NND model, sequence 0 7 is much more likely than 7 0 (revealing a clear temporal bias).

Perhaps the most effective comparison of the models and data is produced by applying TFFS to repeated engine outputs as the fuel-air ratio is reduced in a stepwise fashion [30]. As discussed above, the NND model predicts the onset of a noisy period-doubling bifurcation sequence (with an inherent temporal bias), while the ACO model predicts noisy oscillations with no temporal bias. Figure 7 illustrates a direct evaluation of temporal irreversibility using TFFS for experimental engine data and realizations of each of the models. Parameter values for the ACO model were: $\phi_u=0.70$, $\phi_l=0.68$, $\psi=0.03$, and $\alpha=0.275$; for the NND model: $\phi_u=0.685$, $\phi_l=0.665$, $\psi=0.007$, and $F=0.14$.

As we observe, the TI in the NND model output rises and falls as fuel-air ratio is reduced, closely tracking the experimental engine data. Because we know that the NND model exhibits this variation as the result of noisy bifurcations, we infer that the experimental engine is undergoing this same type of transition. As expected, the ACO model produces no significant TI at any fueling level.

The usefulness of TFFS for delineating the noisy bifurcation process in experimental engine data is further illustrated in Fig. 8. Here we observe a more complex variation in TI produced with fuel-air ratio reductions in an engine operat-

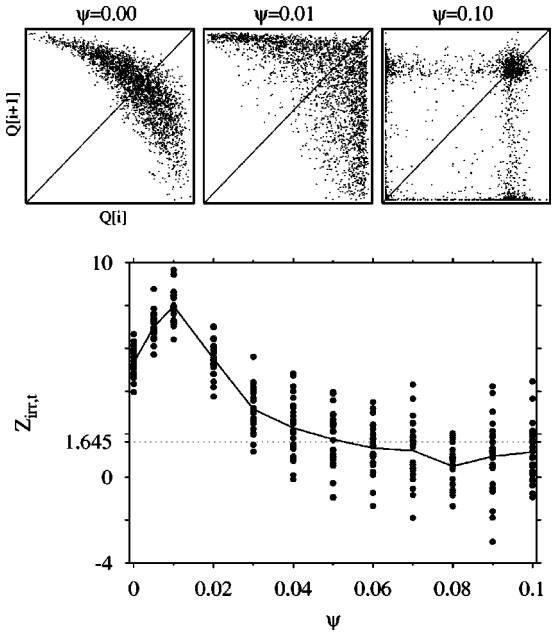

FIG. 9. TFFS results for 25 realizations of the NLD engine model with varying levels of dynamic noise on ϕ (as-injected fuel-air ratio), denoted ψ. Other model parameters were set as defined in the text. Symbolization parameters are $\{n=8,\ L=2\}$ with a decorrelation interval of two iterates. Each return-map scale has been adjusted for maximal visibility. All units are dimensionless.

ing at a higher residual fraction (externally imposed by the experimenter). In particular, we see multiple peaks in TI, and the second peak is significantly larger than the peak observed at the lower residual fraction in Fig. 7. The TFFS statistic, which focuses on only the most irreversible word, follows the same trend as the T_{irr} metric, which evaluates all words, so TFFS quantifies the same changes in TI, with the statistical confidence bound not directly attainable with T_{irr}.

Based on simulations with the NND model at the higher residual fraction, we infer that the larger TI peak results from transitions to higher-order periodicities and chaos. One might generally expect such a trend when nonlinear memory is increased.

Because of the potential importance of parametric pertubations (such as in injected fuel), it is also interesting to observe the impact of increasing dynamic noise levels on the TFFS statistic. Results of numerical experiments in which increasing levels of parametric noise are added to the NND model are illustrated in Fig. 9. In this case, the noise level on the horizontal axis represents the magnitude of the standard deviation in the as-injected fuel parameter relative to its nominal value. At selected noise levels, we have plotted the resulting value of $Z_{\mathrm{irr},t}$ for 25 different realizations of model output (2800 engine cycles for each realization). The plotted solid line represents the mean for the 25 realizations.

We observe that the dynamic noise can reach rather high levels before there is significant attenuation in TI. This implies that one should be able to observe TI in engine combustion experiments even when engine parameters are highly noisy, such as one might expect for realistic driving conditions. We attribute this dynamic noise insensitivity to the fact that dynamic noise is being amplified by the same nonlinearity that generates the important dynamics.

VIII. SUMMARY AND CONCLUSIONS

Our results indicate that detection of temporal irreversibilities in time-series data can be accomplished effectively and efficiently using symbolization. The TFFS statistic in particular appears to be useful for evaluating a time-reversible null hypothesis without the need for generating surrogates.

ACKNOWLEDGMENTS

The authors thank F.T. Connolly of the Ford Motor Company for use of the experimental engine heat-release data and R.M. Wagner, formerly of the University of Missouri–Rolla and presently of the Oak Ridge National Laboratory, for use of the experimental engine bifurcation data and accompanying NND model fits.

[1] G. Weiss, J. Appl. Probab. **12**, 831 (1975).
[2] Y. Pomeau, J. Phys. (France) **43**, 859 (1982).
[3] A.J. Lawrance, Int. Stat. Rev. **59**, 67 (1991).
[4] P. Rothman, J. Appl. Econometrics **7**, 187 (1992).
[5] G.B. Giannakis and M.K. Tsatsanis, IEEE Trans. Signal Process. **42**, 3460 (1994).
[6] C. Diks, J.C. van Houwelingen, F. Takens, and J. DeGoede, Phys. Lett. A **201**, 221 (1995).
[7] J.B. Ramsey and P. Rothman, J. Money Credit Banking **28**, 1 (1996).
[8] L. Stone, G. Landan, and R.M. May, Proc. R. Soc. London, Ser. B **263**, 1509 (1996).
[9] B.P.T. Hoekstra, C.G.H. Diks, M.A. Allessie, and J. DeGoede, Chaos **7**, 430 (1997).
[10] M.J. van der Heyden, C.G.C. Diks, B.P.T. Hoekstra, and J. DeGoede, Physica D **117**, 299 (1998).
[11] J.B. Green Jr., C.S. Daw, J.S. Armfield, C.E.A. Finney, R.M. Wagner, J.A. Drallmeier, M.B. Kennel, and P. Durbetaki, Society of Automotive Engineers Technical Paper No. 1999-01-0221 (1999).
[12] J.P. Crutchfield and N.H. Packard, Physica D **7**, 201 (1983).
[13] J.P. Crutchfield and K. Young, Phys. Rev. Lett. **63**, 105 (1989).
[14] A.B. Rechester and R.B. White, Phys. Lett. A **156**, 419 (1991); **158**, 51 (1991).
[15] X.Z. Tang, E.R. Tracy, A.D. Boozer, A. deBrauw, and R. Brown, Phys. Rev. E **51**, 3871 (1995).
[16] U. Schwarz, O. Benz, J. Kurths, and A. Witt, Astron. Astrophys. **277**, 215 (1995).
[17] J. Kurths, A. Voss, P. Saparin, A. Witt, H.J. Kleiner, and N. Wessel, Chaos **5**, 88 (1995).
[18] M. Lehrman, A.B. Rechester, and R.B. White, Phys. Rev. Lett. **78**, 54 (1997).
[19] X.Z. Tang and E.R. Tracy, Chaos **8**, 688 (1998).
[20] C.E.A. Finney, J.B. Green, Jr., and C.S. Daw, Society of Automotive Engineers Technical Paper No. 980624 (1998).
[21] C.S. Daw, M.B. Kennel, C.E.A. Finney, and F.T. Connolly, Phys. Rev. E **57**, 2811 (1998).
[22] H. Voss and J. Kurths, Phys. Rev. E **58**, 1155 (1998).
[23] T. Schreiber, Phys. Rev. Lett. **78**, 843 (1997).
[24] M.B. Kennel, Phys. Rev. E **56**, 316 (1997).
[25] A. Witt, J. Kurths, and A. Pikovsky, Phys. Rev. E **58**, 1800 (1998).
[26] J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, and J.D. Farmer, Physica D **58**, 77 (1992).
[27] T. Schreiber and A. Schmitz, Phys. Rev. Lett. **77**, 635 (1996).
[28] M.B. Kennel and A. Mees, Phys. Rev. E **61**, 2563 (2000).
[29] H.D.I. Abarbanel and M.B. Kennel, Phys. Rev. Lett. **80**, 3153 (1998).
[30] R.M. Wagner, J.A. Drallmeier, and C.S. Daw (unpublished).