

## Observing and modeling nonlinear dynamics in an internal combustion engine

C. S. Daw\*

*Engineering Technology Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831-8088*

M. B. Kennel

*Institute for Nonlinear Science, University of California, San Diego, La Jolla, California 92093-0402*

C. E. A. Finney

*College of Engineering, University of Tennessee, Knoxville, Tennessee 37996-2210*

F. T. Connolly

*Ford Motor Company, Dearborn, Michigan 48121-2053*

(Received 26 September 1997)

We propose a low-dimensional, physically motivated, nonlinear map as a model for cyclic combustion variation in spark-ignited internal combustion engines. A key feature is the interaction between stochastic, small-scale fluctuations in engine parameters and nonlinear deterministic coupling between successive engine cycles. Residual cylinder gas from each cycle alters the in-cylinder fuel-air ratio and thus the combustion efficiency in succeeding cycles. The model's simplicity allows rapid simulation of thousands of engine cycles, permitting statistical studies of cyclic-variation patterns and providing physical insight into this technologically important phenomenon. Using symbol statistics to characterize the noisy dynamics, we find good quantitative matches between our model and experimental time-series measurements. [S1063-651X(98)08903-X]

PACS number(s): 05.45.+b

### I. INTRODUCTION

Under constant nominal operating conditions, spark-ignited internal combustion engines can exhibit substantial cycle-to-cycle variation in combustion energy release. This phenomenon has been observed and studied since the 19th century (e.g., Clerk [1]), and research has continued until the present [2–10]. Extensive reviews of previous works are given in [11,12]. Previous researchers have characterized *cyclic variability* (CV) in terms ranging from stochastic to linear determinism to deterministic chaos. To date there has been no experimental confirmation of deterministic chaos underlying CV, but there has been experimental support presented for both stochastic and linear deterministic features. This ambiguity has created a continuing debate in the engineering community about the true nature of CV [13].

We propose a simple, discrete engine model that explains how both stochastic and deterministic features can be observed. Our model is unique in that it combines stochastic and nonlinear deterministic elements to provide a global combustion description consistent with the underlying physics. Our approach ignores complex spatial details and instead focuses on the cylinder-average mass balance and energy release. The result is a simple nonlinear map that produces cycle-resolved combustion time sequences statistically similar to a real engine. Because our model is physically realistic, we expect that it can be used to predict CV trends with as-fed fuel-air ratio and to provide fundamental insight into the

causes of combustion irregularities.

Economic and regulatory pressures are pushing engine manufacturers to operate with lean fueling and exhaust-gas recirculation (EGR) to increase fuel economy and minimize  $\text{NO}_x$  emissions. CV increases with lean fueling and EGR and actually limits the potential benefits which can be derived from these operating modes. Thus understanding the ultimate causes of CV can have important economic and environmental consequences. Our goal is to provide engine designers with insight that could lead to design improvements and active control methods for improving engine performance.

### II. FOUR-STROKE ENGINE CYCLE

Most gasoline-fueled automobiles use engines operating with the four-stroke, spark-ignition (Otto) cycle. Initially, fuel and air are inducted through the intake valve into the cylinder, and the resulting mixture is compressed. At a point typically just before maximal compression, the discharge of a high-voltage spark initiates combustion. Combustion and expansion of the hot gases proceed following ignition, and work is transferred through the piston, connecting rod and crankshaft to the load. As the piston moves upward again following the power stroke, exhaust gases are vented through the exhaust valve. Following exhaust, a fresh fuel and air charge is inducted into the cylinder to begin the next cycle.

In real engines, not all of the gases in the cylinder are expelled during the exhaust process. This *residual fraction*, which includes combustion products and typically some unreacted fuel and air, is a function of engine design. Residual fraction is affected by several design parameters, one of the most important being *valve overlap*, the brief period that the

---

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

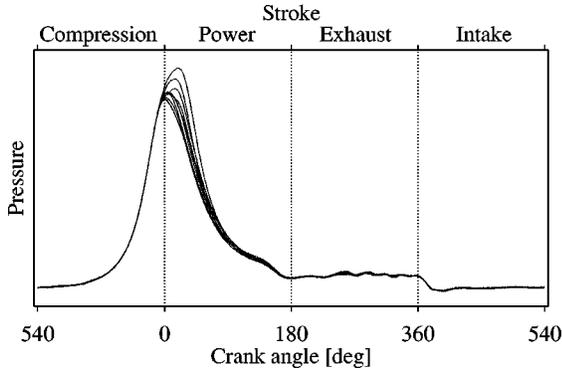


FIG. 1. Variability in cylinder pressure from cycle to cycle. Pressure-trace segments from ten consecutive engine cycles are overlaid for visibility.

intake valve is open before the exhaust valve closes. Valve overlap is generally helpful at higher engine speeds in producing power but at lower, near-idle speeds tends to degrade combustion (i.e., increases CV). The detailed physical mechanisms governing residual fraction are quite complicated, involving turbulent mixing in the cylinder and intake and exhaust ports. For our model, it is the net effect of these flows that is important, as discussed in Sec. III.

The crankshaft transfers power from the piston to the driveshaft, which transfers power to the load (e.g., wheels, in automotive applications). It is typical to describe piston location in terms of *crank angle degrees* (CAD), the angle of the crankshaft relative to *top dead center*, which is at “the top” of its stroke and where the piston extension into the cylinder is maximal.

The course of each combustion event can be followed by monitoring the internal cylinder pressure versus crank angle, as depicted in Fig. 1. As energy is released during combustion, cylinder pressure exceeds that which occurs without combustion, thereby producing useful work. CV causes non-repeating pressure traces during successive engine cycles. The region in Fig. 1 in which the pressure traces diverge is the combustion region. In some cycles, combustion is incomplete, leaving residual unburned fuel. In other cycles, excess energy is produced from combustion of both residual and fresh fuel. Both extremes of combustion are undesirable because they produce alternating pulses of fuel and nitrogen oxides in the exhaust and a feeling of “roughness” to the driver.

### III. MODEL DEVELOPMENT

The primary deterministic element we focus on is the presence of retained fuel and oxygen from one engine cycle to the next. Retained fuel and oxygen influence succeeding cycles because of a strong nonlinear dependence of combustion rate on the in-cylinder gas composition at the time of spark. We assume that other dynamical effects can be represented as stochastic fluctuations in one or more key parameters, such as injected fuel-air ratio, residual-gas fraction, and the lower ignition limit (the minimal fuel-air ratio to achieve combustion). By introducing these random parametric fluctuations, we intend to account for complex, high-dimensional processes such as turbulent mixing, fuel-droplet

vaporization, and fuel deposits on the cylinder wall. By the central-limit theorem, we expect these noisy parametric inputs to be Gaussian distributed because they arise from the combined contribution of many processes. Although the fluctuations are input randomly, their ultimate effect is filtered through the global nonlinear determinism.

Our model is discrete in time, representing each full engine cycle (including intake and exhaust) as a single event. The dynamical variables which define the two-dimensional state space are the masses of fuel and air present in the cylinder at the time of spark,  $m[i]$  and  $a[i]$ , respectively.

#### A. Intake phase

The total mass of gas in the cylinder at the time of spark is equal to the residual mass from the previous cycle plus new intake,

$$m[i] = m_{\text{res}}[i] + m_{\text{new}}[i], \quad (1)$$

$$a[i] = a_{\text{res}}[i] + a_{\text{new}}[i], \quad (2)$$

where  $m_{\text{res}}[i]$  and  $a_{\text{res}}[i]$  are the masses of unreacted fuel and air remaining from the previous cycle, and where  $m_{\text{new}}[i]$  and  $a_{\text{new}}[i]$  are the masses of fuel and air introduced during the intake stroke.

The new fuel and air masses fed to the cylinder in each cycle are controlled by two constraints. First, the newly fed fuel and air are externally maintained in a fixed ratio (the *equivalence ratio*)  $\phi_o$ . The high-dimensional dynamics associated with fuel vaporization, fuel-injector variations, and air-fuel mixing are accounted for by stochastically perturbing this ratio about its nominal (mean) value,

$$\frac{m_{\text{new}}[i]}{a_{\text{new}}[i]} = \phi_o [1 + \sigma_\phi N(0,1)], \quad (3)$$

where  $\sigma_\phi$  is a scaling factor and  $N(0,1)$  is a random deviate drawn each cycle from a zero-mean, unit-variance Gaussian distribution.

Second, we assume that immediately before combustion, the total number of moles of fuel and air combined in the cylinder is a constant. In nondimensional units,

$$\frac{m[i] + WRa[i]}{1 + WR} = 1, \quad (4)$$

where  $W = w_f/w_a$  is the ratio of the molecular weights of the principal fuel and air fractions, and  $R$  is the air-fuel mass ratio at stoichiometric burning, that is, the condition at which every fuel molecule is fully oxidized and no excess oxygen remains.

This latter constraint [Eq. (4)] can be reasonably justified assuming ideal-gas behavior and constant input pressure, throttling conditions, temperature, and cylinder volume. In our simulations, we use  $w_f = 114$  g/mol and  $w_a = 29$  g/mol, and  $R = 14.6$ , appropriate for common hydrocarbon fuels in air. Equations (3) and (4) implicitly define the amount of new fuel and air injected into the cylinder each cycle as a function of the amount of residual gas from the previous cycle.

### B. Combustion efficiency

Combustion efficiency in any given cycle is defined as the fraction of the fuel present that burns. In the current version of our model, we assume net combustion efficiency  $C$  is a function only of the in-cylinder equivalence ratio,  $\Phi[i]$  =  $m[i]/a[i]$ , at the time of spark:

$$C[i] = C(\Phi[i]) = C_{\max} [1 + 100^{-(\Phi[i] - \phi_m)/(\phi_u - \phi_l)}]^{-1}. \quad (5)$$

As represented by Eq. (5), the relationship between combustion efficiency and equivalence ratio has a sigmoidal shape, converging to  $C_{\max}$  (near 1) as  $\Phi$  approaches stoichiometry (i.e.,  $\Phi = 1$ ) and converging to 0 as  $\Phi$  becomes very small. We parametrize the position of the knee by  $\phi_l$  and  $\phi_u$ , the conditions where the efficiency is approximately 10% and 90% of  $C_{\max}$ , defining  $\phi_m = (\phi_u + \phi_l)/2$ . In this work, we only consider stoichiometric to lean combustion ( $\Phi \leq 1$ ).

The exact functional form of the combustion efficiency is somewhat arbitrary, but its general shape reflects the experimental observation and the physics of combustion. Specifically, it is known that as  $\Phi$  drops below a critical value called the *lean limit*, the burning rate and combustion efficiency decrease exponentially [14,15]. Also,  $C_{\max}$  cannot exceed 1 by definition. For hydrocarbon fuels, the critical equivalence ratio is typically 0.5–0.6 [14]. The steepness of this curve is a consequence of the sensitivity of flame-front propagation speed to small changes in gas composition near the lean limit.

### C. Combustion and exhaust phase

The heat released in each combustion event is proportional to  $Q[i] = C[i]m[i]$ . The physical mechanism for cycle-to-cycle coupling is that a fraction  $F$  of the unreacted fuel and air remains in the cylinder for the next cycle, thus affecting the next cycle's combustion:

$$m_{\text{res}}[i+1] = Fm[i](1 - C[i]), \quad (6)$$

$$a_{\text{res}}[i+1] = F(a[i] - RC[i]m[i]). \quad (7)$$

We model fluctuations in  $F_o$  by perturbing it each cycle with a random number,

$$F = F_o [1 + \sigma_F N(0,1)], \quad (8)$$

where  $\sigma_F$  is a scaling factor and  $N(0,1)$  is a random deviate drawn each cycle from a zero-mean, unit-variance Gaussian distribution. Experimental measurements suggest that  $F_o$  can vary from 0 to 0.3 depending on engine design and operating conditions [15].

### D. Summary of nondimensionalized model

We have nondimensionalized all air and fuel masses in units of what the fuel and air mass would each be at perfectly combusting stoichiometric conditions with no residuals ( $C = C_{\max} \approx 1$ ,  $m_{\text{res}} = a_{\text{res}} = 0$ ).

The overall model is thus characterized as a two-dimensional dynamic map, taking the state variables  $m$  and  $a$  one cycle forward in time:

$$m[i+1] = A(m[i], a[i], \phi_o, \sigma_\phi, F_o, \sigma_F, \dots), \quad (9)$$

$$a[i+1] = B(m[i], a[i], \phi_o, \sigma_\phi, F_o, \sigma_F, \dots) \quad (10)$$

for mapping functions  $A$  and  $B$ . The key features are the nonlinearity produced by the sharp change in combustion efficiency with  $\Phi$  and the amplification of the random dynamical perturbations in  $\phi_o$  and  $F_o$  by the nonlinear mapping.

The constraint imposed by Eq. (4) reduces the effective degrees of freedom to 1, so that the mapping can actually be written in the form

$$Q[i+1] = f(Q[i], m[i], a[i], \phi_o, \sigma_\phi, F_o, \sigma_F, \dots). \quad (11)$$

We keep the two-equation form [Eqs. (9) and (10)] to simplify computation and to facilitate future improvements.

## IV. MODEL PREDICTIONS

Bifurcation diagrams of combustion heat release based on changes in  $\phi_o$  and the model parameters are shown in Fig. 2. The parameter changes illustrated were selected to be within the expected ranges described in Sec. III. Each plot is produced by iterating the mapping [Eqs. (9) and (10)] for a fixed  $\phi_o$  and the indicated parameter values beginning with arbitrary initial values for  $m$  and  $a$  and discarding start-up transients.

Although the bifurcation details change with parameter values, certain general trends are apparent.

(i) Near stoichiometric conditions, the amount of fuel burned in each cycle stabilizes to a fixed point.

(ii) For a decrease of  $\phi_o$  below a critical value, the amount of fuel burned undergoes a period-2 bifurcation.

(iii) For still lower  $\phi_o$ , combustion oscillations become more complex, leading to multiperiodic or chaotic patterns.

(iv) For very low  $\phi_o$ , combustion effectively ceases.

(v) When noise is added to either  $\phi_o$  or  $F_o$  or both, the detailed bifurcation structure becomes fuzzy but still reflects the underlying sequence of fixed point, period-2 bifurcations and/or chaos.

(vi) Noise also causes the initial bifurcation to occur at a higher  $\phi_o$  (i.e., higher than when no noise is added) and maintains combustion in the extreme lean limit because of occasional spikes of additional fuel.

Briefly stated, our model predicts that combustion becomes unstable near the lean limit due to the onset of period-doubling bifurcations. This instability is enhanced by random perturbations in parameters such as injected equivalence ratio and residual fraction. The prediction of a period-doubling instability is important because it provides a unique signature that can be experimentally verified, and because it demonstrates the relevance of nonlinear dynamics to real-world engine behavior.

## V. EXPERIMENT

To provide a set of detailed observations which could be compared with our model, we made a series of CV measurements at fueling conditions ranging from near stoichiometric to very lean using a highly instrumented and well-

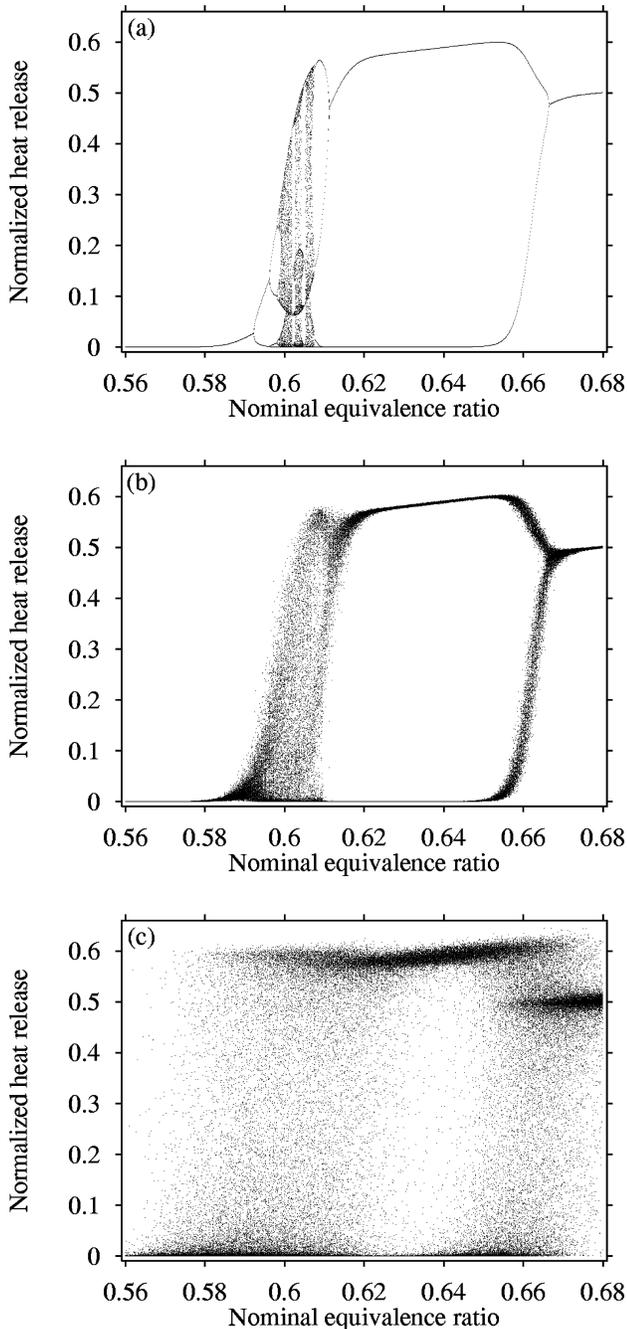


FIG. 2. Model equivalence-ratio ( $\phi_o$ ) bifurcation plots with  $\sigma_\phi=0$  (a),  $\sigma_\phi=0.001$  (b), and  $\sigma_\phi=0.01$  (c). Fixed model parameters are  $\phi_l=0.59$ ,  $\phi_u=0.60$ ,  $F_o=0.25$ , and  $\sigma_F=0$ .

characterized engine. Although the engine was highly instrumented, it was basically a production V8 engine with standard port fuel injection connected to a dc motoring dynamometer. Thus we expect that our observations are likely to be relevant in a practical engineering context.

The nominal engine operating condition was 1200 RPM, 27.1-N m brake torque (engine load), 20 CAD before top center spark. The dynamometer was operated in speed-control mode to maintain a nearly constant engine speed despite erratic combustion at very lean conditions. Without the dynamometer, the engine speed fluctuates significantly when large numbers of misfires occur, making engine behavior nonstationary, and a comparison of the engine and model

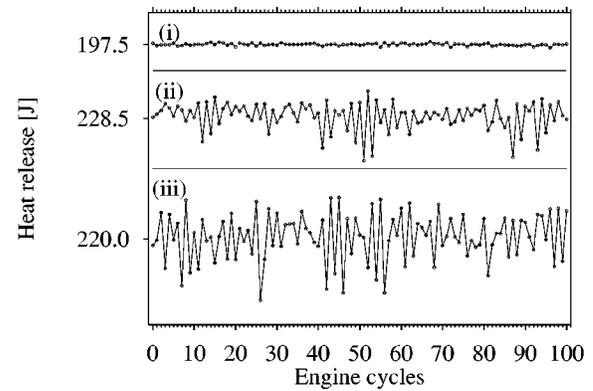


FIG. 3. Segments of measured engine heat-release time sequences at nominal operating conditions for  $\phi_o=0.91$  (i), 0.59 (ii), and 0.53 (iii).

problematic. Feedback engine controllers were engaged to achieve an operating condition; once the condition was achieved, the feedback controllers were shut off, and the engine was run in open-loop mode, except for the dynamometer speed control. This strategy assured that combustion was minimally influenced by feedback controllers while the engine ran at a constant speed.

We recorded combustion pressure once per CAD from a single cylinder and nominal operating conditions at a 50-Hz rate for over 2800 contiguous cycles. Combustion pressure measurements were made with a piezoelectric pressure transducer mounted in the cylinder head. To provide a dynamic measurement that could be compared with the model, we calculated the combustion heat release for each cycle by integrating the cylinder pressure data using a method equivalent to the Rassweiler-Withrow method [15]. As a result, for each engine experiment we produced time sequences of over 2800 heat-release values.

Figure 3 shows short segments of heat-release time sequences from the engine at three equivalence ratios. At the nearly stoichiometric condition [Fig. 3(a)], combustion is variable, but the range of variations is small. For lean conditions [Figs. 3(b) and 3(c)], the range of combustion heat release increases. As seen in the figure, the mean value of the heat release for the three conditions shifts slightly, but the main difference is in the increase in variance.

Although many measurements of CV have been made previously, we believe that the experimental protocol described above is unique. Specifically, we took great pains to eliminate noncombustion dynamic effects from the standard engine controllers, and we collected much longer sequences of combustion measurements than is usual. The additional data provided us with much greater confidence in the existence of consistent dynamic patterns.

One complication in our experimental procedure which made comparisons with the model more difficult was that we controlled the injected fuel-air ratio by adjusting the throttle. Because the throttle changes intake pressure, factors such as in-cylinder mixing and residual gas fraction are also changed. Thus it was not possible to make a series of runs changing only one parameter. Nevertheless, we were able to vary the degree of CV enormously from very small amounts near stoichiometric fueling ( $\phi_o=1.0$ ) to very high amounts at very lean fueling ( $\phi_o<0.55$ ).

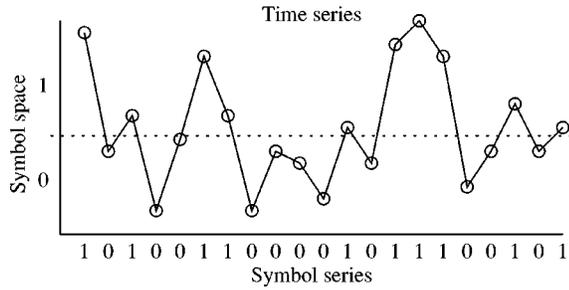


FIG. 4. Procedure for conversion of a time series to a symbol series based on a defined partition.

## VI. APPLICATION OF SYMBOL STATISTICS

Contending with the effects of dynamical parameter noise is a key issue in characterizing the predicted model dynamics and comparing the model with experimental observations. As seen in the model bifurcation diagrams, we expect such noise to blur but not completely obscure the deterministic signature. To observe these noisy dynamical patterns, we employ the symbol-statistics approach suggested by Tang *et al.* [16]; an alternate but analogous approach may be found in Ref. [17]. Although this method is motivated by symbolic dynamics theory, it is not completely rigorous, mainly because generating partitions are undefined in the presence of noise. The reader is referred to Crutchfield and Packard [18] for a detailed discussion. Like Tang *et al.*, we use a practically motivated approach to depict the dynamic patterns in model-produced or observed measurements and to compare the goodness of fit between the model and experimental data. In the latter case, we have been able iteratively to adjust the model parameters to obtain good agreement with the observations.

Our basic idea in using the symbol-statistics approach is to discretize the predicted or measured heat-release values into a finite set of discrete values, as predicted in Fig. 4. Depending on the value of a given heat release, it is assigned one of  $n$  symbolic values (e.g., 0 or 1 for  $n=2$ ; in the terminology of symbol dynamics,  $n$  is the *alphabet size*). Typically, we define discretization partitions such that the individual occurrence of each symbol is equiprobable with all others. We do this to obtain ready discrimination between random and nonrandom symbol sequences, recognizing that the resulting partition is not generating. Since a generating partition is undefined in the presence of noise [18], some such practical approach is required in order to proceed.

Once a heat-release time series is symbolized, we evaluate the relative frequency of all possible symbol sequences in the data defined by a symbol-sequence vector of  $L$  cycles length (in the terminology of symbolic dynamics, this sequence vector constitutes a *word* and  $L$  is the *word size*). For example, if we let  $L=5$ , we determine the relative frequency of occurrence for each possible sequential combination of five symbols. A simple way to keep track of symbol-sequence frequencies is to assign a unique number to each possible sequence by evaluating the equivalent base-10 value of each base- $n$  sequence; we call this number the *sequence code*. For example, a sequence of 010101 occurring with a binary partition ( $n=2$ ) would have a sequence code of 21. This is very similar to the approach used by Rechester and

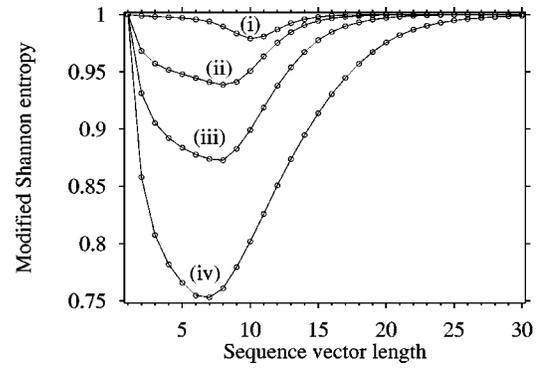


FIG. 5. Determination of a suitable sequence vector length  $L$  by a minimum in the modified Shannon entropy. Data are from the model at four equivalence ratios ( $\phi_o$ ): 0.91 (i), 0.67 (ii), 0.63 (iii), and 0.59 (iv).

White [19]. Using this method of identifying each symbol sequence also allows us to observe the relative-frequency histograms as two-dimensional plots (see Sec. VII).

We refer to the tally of symbol-sequence frequencies versus sequence code as a *symbol-sequence histogram*. Because of our partitioning rule, the relative frequency of each possible sequence for truly random data will be equal (subject to the availability of sufficient data). Thus any significant deviation from equiprobability is indicative of time correlation and deterministic structure. Similar to Tang *et al.*, we define a modified Shannon entropy as

$$H_S(L) = \frac{1}{\ln n^{\text{seq}}} \sum_i p_{i,L} \ln p_{i,L} \quad (12)$$

where  $n^{\text{seq}}$  is the total number of sequences with nonzero frequency,  $i$  is a string-sequence index of sequence vector length  $L$ , and  $p_{i,L}$  is the probability of string sequence  $i$ . The only difference between Eq. (12) and the definition used by Tang *et al.* is that we use the number of non-zero-frequency sequences instead of the total number of possible sequences. This choice of  $n^{\text{seq}}$  reflects the fact that many possible sequences may not be realized because of finite data-set length. The result is to bias  $H_S$  upward when the number of possible sequences becomes large relative to the available data. For random data  $H_S$  should equal 1, whereas for nonrandom data it should be between 0 and 1.

One approach we found useful for selecting an appropriate sequence vector length ( $L$ ) involves using the modified Shannon entropy. Specifically, we find that  $H_S$  typically reaches a minimum value as vector length is increased from 1. This trend is illustrated in Fig. 5 using data generated with the model at four operating conditions. As the bifurcation progresses, the nonrandom part of the dynamics becomes more evident, even though significant parametric noise is present. We explain this minimum in  $H_S$  as reflecting the symbol-sequence transformation which best distinguishes the data from a random sequence. Sequence vectors that are too short lose some of the important deterministic information. Sequence vectors that are too long reflect noise and data depletion (i.e., there are not enough data to obtain reliable

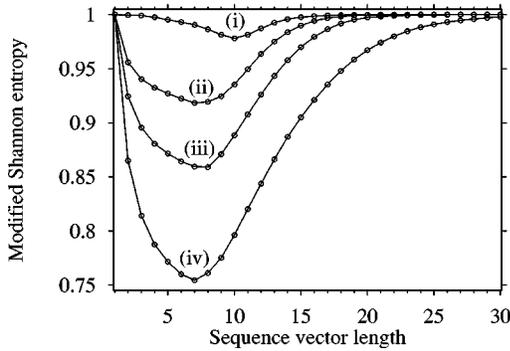


FIG. 6. Modified Shannon entropy as a function of sequence vector length  $L$  for experimental engine data at four equivalence ratios ( $\phi_o$ ): 0.91 (i), 0.67 (ii), 0.63 (iii), and 0.59 (iv).

statistics for such long sequences). One can thus argue that the  $L$  value for which  $H_S$  is minimum is an “optimal” choice for the given data.

## VII. SYMBOL-SEQUENCE COMPARISONS

In comparing the experimental heat-release patterns with those predicted by the model, we begin by evaluating the general trends for the modified Shannon entropy. As illustrated in Fig. 6, we find that the real engine exhibited similar trends in  $H_S$  with degree of leanness in fueling and symbol-sequence length (compare with Fig. 5 for the model). The similarity in the figures suggests that the model and real engine behave consistently in response to increasingly lean fueling.

Figure 7 illustrates a more detailed comparison of the model and experiment through a sequence of symbol-sequence histograms. The abscissa for all plots is the sequence code for the six-member binary sequence (i.e.,  $n = 2$ ,  $L = 6$ ). Note that in both cases, peaks reflecting high frequencies of 010101 (21) and 101010 (42) combinations begin to emerge from a flat profile as the noisy period-2 bifurcation begins. The high visibility of these peaks, even when there is a high noise level, suggests that symbol-sequence histograms may be generally useful for detecting the onset of noisy bifurcations.

## VIII. FITTING THE MODEL

As mentioned previously, we were not able to control our experimental engine such that all of the parameters remained constant while the injected fuel-air ratio was reduced. We also had no way of directly measuring the residual fraction  $F_o$ , or the noise amplitudes for  $F_o$  and the injected fuel-air ratio  $\phi_o$ . Thus in making detailed comparisons between the model and experiment, we were limited to evaluating how well the model could reproduce the observed behavior as the unknown parameters were adjusted over physically reasonable ranges.

To fit the model against observed data, we adjusted the model free parameters ( $\phi_1, \phi_u, \phi_o, \sigma_\phi, F_o, \sigma_F$ , and a multiplicative scale factor relating nondimensional to experimental heat-release units). We assumed the structure of the parametric noise terms to be Gaussian. We optimized the fit by iteratively adjusting the parameters to give the best agree-

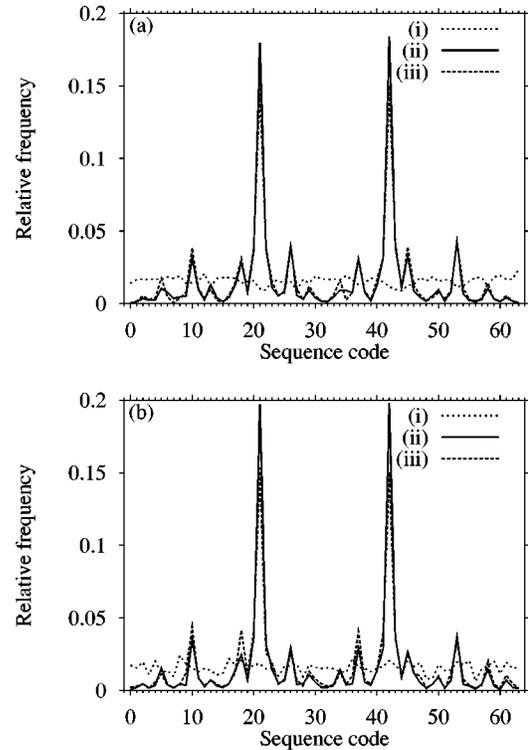


FIG. 7. Symbol-sequence histograms, with a binary partition and sequence vector length of 6, for model (a) and engine (b) at three equivalence ratios ( $\phi_o$ ): 0.91 (i), 0.59 (ii), and 0.53 (iii).

ment between symbol statistics (as represented by the symbol-sequence histogram) for iterations of the model and experimental data.

In Ref. [16], the Euclidean distance between vectors whose elements are the occupations of all possible sequences served as the minimization target, but in this work, we use a two-sample  $\chi^2$  statistic as the criterion for minimization:

$$\chi^2 = \sum_i \frac{(N_i^{\text{obs}} - N_i^{\text{model}})^2}{N_i^{\text{obs}} + N_i^{\text{model}}}. \quad (13)$$

Using this method, we can also evaluate the statistical significance of a trial model fit, using the standard  $\chi^2$  probability inference with  $n^{\text{seq}} - 1$  degrees of freedom, with  $n^{\text{seq}}$  the total number of sequences with non-zero frequency for either model or data.

With a good fit, the model converges to a  $\chi^2$  value which will accept the null hypothesis that the same process generated the histogram-bin occupations for model and experiment, thus providing some quantitative assurance that the model is a valid description of the experimental observations. The minimization algorithm [20] was a hybrid of simplex and genetic methods designed for continuous parameter spaces without requiring derivatives.

Whereas a binary partition is sufficient for detecting the onset of bifurcations, we find that higher-level partitions are needed to obtain the best fit of our model to experimental data. A vivid example of the insufficiency of a binary partition for fitting is illustrated in Figs. 8 and 9. Figure 8(a) is the first return map for measured engine heat-release values at

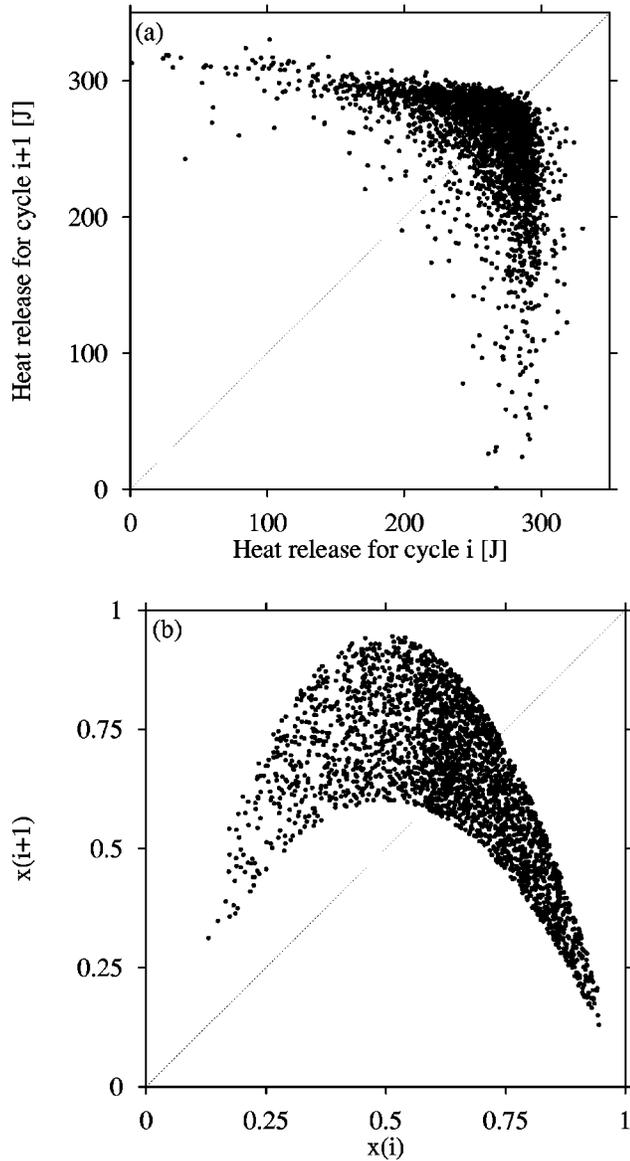


FIG. 8. Lag-1 return maps for engine data at  $\phi_o = 0.63$  (a) and a noisy period-2 logistic map (b). The noisy logistic map is introduced to show that a simple binary partition using our equiprobable convention does not distinguish two very different models.

$\phi_o = 0.63$ , whereas Fig. 8(b) is a first return map for the logistic map with Gaussian noise added to the feedback parameter. In the latter case, the mean value of the feedback parameter has been adjusted to produce a period-2 oscillation. These return maps clearly reveal that the shapes of the underlying deterministic maps are different. Indeed, we expect the logistic map to be less accurate than our engine model in matching our experimental data. Nevertheless, as illustrated in Fig. 9(a), the symbol-sequence histograms using binary symbolization are nearly identical, making discrimination of the two data sets on the basis of these histograms very difficult. Our point here is that a binary partition is clearly insufficient for distinguishing a more- from a less-appropriate model, given that the dynamics is dominated by a noisy period-2 oscillation. Even in the absence of noise, an equiprobable binary partition cannot distinguish period 2 from period 4, period 8, etc.

Figure 9(b) illustrates how much easier it is to distinguish

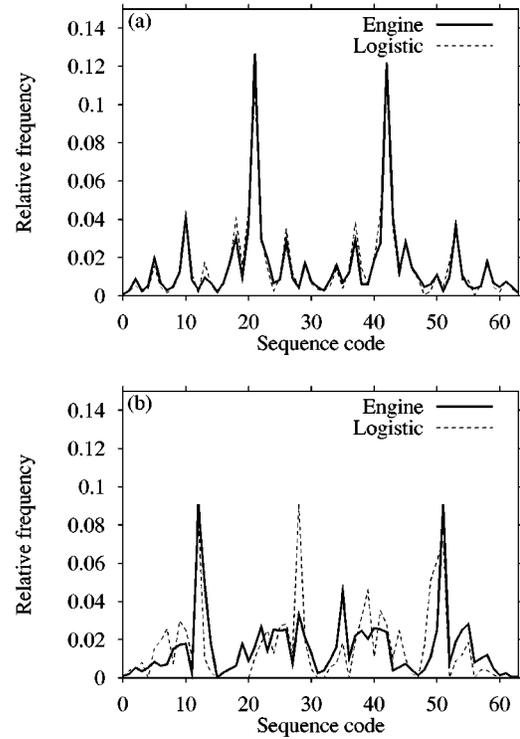


FIG. 9. Symbol-sequence histograms for engine data at  $\phi_o = 0.63$  and a noisy period-2 logistic map. A binary partition and sequence vector length of 6, seen in (a), does not provide the discriminatory power seen with a quaternary partition and sequence vector length of 3, seen in (b).

the engine data from a noisy logistic map by increasing the number of symbols  $n$  from 2 to 4 [the sequence length was reduced from 6 to 3 to produce histograms with the same level of detail ( $2^6 = 4^3 = 64$  elements)]. Since a generating partition is not available in cases such as ours, partition choice must be based on a tradeoff between the number of symbols and the sequence vector length, both of which affect the size and statistical significance of the symbol-sequence histogram. In our case, the high noise level and rapid divergence of nearby points reduces the information memory of the system to be a short time, making symbolizations long in time but coarse in partitioning less desirable, and empirically, less effective for best fitting.

In Fig. 10, we use return maps to illustrate how well our fitted model matches the observed data for a moderately lean fueling case. We used a variable partition of 9 and 8 in a two-cycle-long sequence to achieve this degree of fit, where 9 is the number of partitions applied to the first member of the sequence and 8 is the number of partitions used for the second member. Variable partitioning schemes are by no means essential to the method but may be used without harm. They sometimes provide better fits because the one-dimensional marginal distributions of trial simulations are constrained to match the data more closely when the symbol partitions are not commensurate.

Note that in the first return map comparison [Fig. 10(a)], there is a downward bias in the upper-left portion of the experimental map. This bias is a consistently occurring feature that we expect is due to a real difference between our

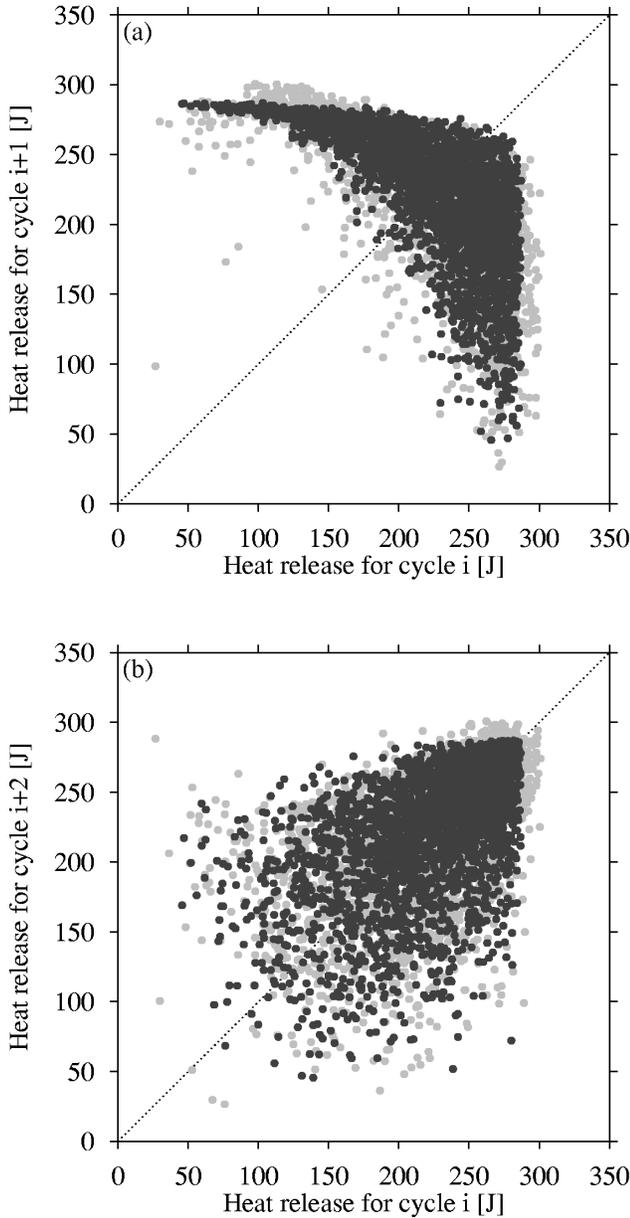


FIG. 10. Return maps for engine data at  $\phi_o = 0.59$  (light points) and optimized-fit model data (dark points) for lag 1 (a) and lag 2 (b).

model and the experimental engine. We attribute this to an additional temperature effect that is currently unaccounted for by the model. Specifically, we conjecture that lower-than-expected heat release occurs following exceptionally poor combustion events because of reduced initial temperature at the time of spark. We plan to include this reduced temperature effect in future model revisions.

Finally, in Fig. 11, we illustrate how well the general trends of the model and experimental data match. At each of three fueling conditions corresponding to near stoichiometric, moderately lean, and very lean, we fitted the model to the observed data. It is apparent that the same basic patterns are clearly occurring in both cases; namely a transition from (1) very-small-amplitude Gaussian combustion variation near stoichiometric fueling to (2) a noisy period-2 combustion bifurcation at moderately lean fueling to (3) a noisy

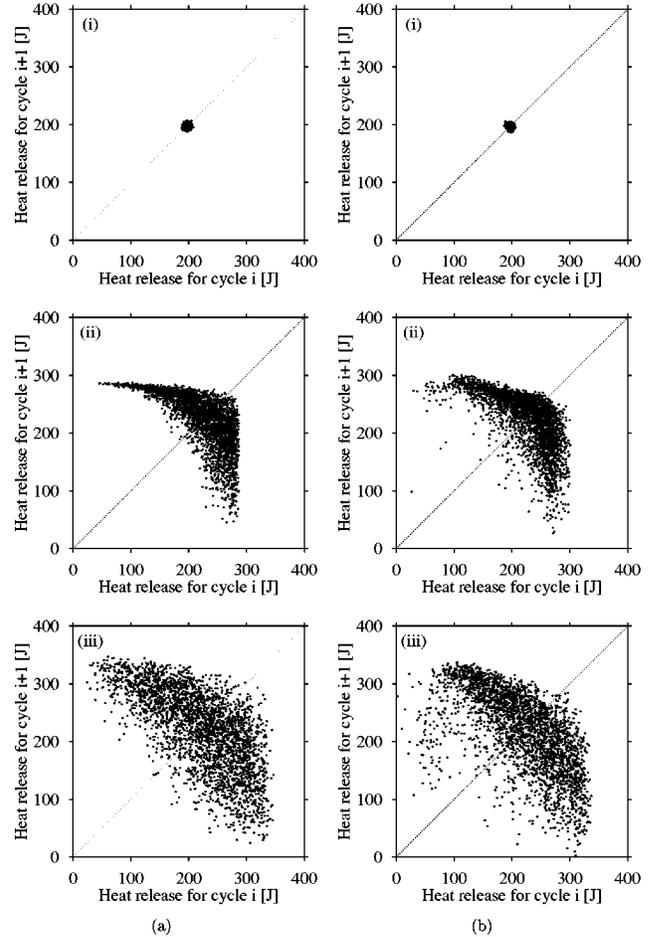


FIG. 11. Return maps for model (a) and engine (b) at three equivalence ratios ( $\phi_o$ ): 0.91 (i), 0.59 (ii), and 0.53 (iii).

multiperiod (possibly chaotic) combustion condition at the leanest condition. For all three fueling conditions the fitted parameter values are well within the range of physical plausibility. The model is specific enough that there is little chance that such similar trends could be caused by overfitting an overly general and unrealistic mathematical function.

## IX. CONCLUSIONS

We believe our model provides a physically reasonable hypothesis that explains the observed time-resolved patterns in cyclic combustion variability. Depending on the injected fuel-air ratio, the behavior can appear to be purely stochastic or a mixture of stochastic and nonlinear dynamics. This range of possible behavior may help explain apparently conflicting observations from previous studies. The ability to describe engine fluctuations with such a simple yet physically plausible model may also aid in the development of cycle-resolved control schemes to reduce or alter the pattern of cyclic fluctuations in order to improve engine performance.

Symbol-sequence statistics are useful for characterizing engine behavior and quantitatively confirm the ability of a low-dimensional nonlinear map to explain the experimental observations. We expect such statistics to be generally useful

for detecting noisy bifurcations and fitting models to noisy data.

We conjecture that the basic modeling approach we used for engine combustion variations may have more general applications. Specifically, we expect that there are many engineering systems involving dominant deterministic nonlinearities and smaller-scale high-dimensional features. In such systems it may often be possible to reproduce observed glo-

bal dynamics with simple deterministic models involving stochastic parametric noise.

#### ACKNOWLEDGMENTS

This work was supported by the U.S. Department of Energy, Office of Energy Research and the Ford Motor Company. The authors express appreciation to E. R. Tracy for invaluable discussions regarding symbol statistics.

- 
- [1] D. Clerk, *The Gas Engine* (Longmans, Green & Co., London, 1886).
- [2] R. K. Barton, D. K. Kenemuth, S. S. Lestz, and W. E. Meyer, SAE Paper No. 700488, 1970.
- [3] J. C. Kantor, *Science* **224**, 1233 (1984).
- [4] M. R. Belmont, M. S. Hancock, and D. J. Buckingham, SAE Paper No. 860324, 1986.
- [5] J. W. Daily, *Combust. Sci. Technol.* **57**, 149 (1988).
- [6] J. K. Martin, S. L. Plee, and D. J. Remboski, Jr., SAE Paper No. 880201, 1988.
- [7] Y. Moriyoshi, T. Kanimoto, and M. Yagita, SAE Paper No. 930066, 1993.
- [8] G. Grünefeld, V. Beushausen, P. Andresen, and W. Hentschel, SAE Paper No. 941880, 1994.
- [9] C. Letellier, S. Meunier-Guttin-Cluzel, G. Gouesbet, F. Neveu, T. Duverger, and B. Cousyn, SAE Paper No. 971640, 1997.
- [10] C. E. A. Finney, K. Nguyen, and C. S. Daw, in *Proceedings of the 32nd Japanese Symposium on Combustion, 1994* (Combustion Society of Japan, Sendai, 1994), p. 400.
- [11] M. B. Young, SAE Paper No. 810020, 1981.
- [12] N. Ozdor, M. Dulger, and E. Sher, SAE Paper No. 940987, 1994.
- [13] C. S. Daw, C. E. A. Finney, J. B. Green, Jr., M. B. Kennel, J. F. Thomas, and F. T. Connolly, SAE Paper No. 962086, 1996.
- [14] I. Glassman, *Combustion*, 2nd ed. (Academic, New York, 1987).
- [15] J. B. Heywood, *Internal Combustion Engine Fundamentals* (McGraw-Hill, New York, 1988).
- [16] X. Z. Tang, E. R. Tracy, A. D. Boozer, A. deBrauw, and R. Brown, *Phys. Rev. E* **51**, 3871 (1995); X. Z. Tang, E. R. Tracy, A. D. Boozer, A. deBrauw, and R. Brown, *Phys. Lett. A* **190**, 393 (1994).
- [17] M. Lehrman, A. B. Rechester, and R. B. White, *Phys. Rev. Lett.* **78**, 54 (1997).
- [18] J. P. Crutchfield and N. H. Packard, *Physica D* **7**, 201 (1983).
- [19] A. B. Rechester and R. B. White, *Phys. Lett. A* **156**, 419 (1991); *ibid.* **158**, 51 (1991).
- [20] R. Storn and K. Price, International Computer Science Institute Technical Report No. TR-95-012, 1995 (unpublished). Available at <ftp://ftp.icsi.berkeley.edu>