Detection of Thermoacoustic Instabilities Via Nonparametric Bayesian Markov Modeling of Time-Series Data

Sihan Xiong

Mechanical Engineering Department,
Pennsylvania State University,
University Park, PA 16802-1412
e-mail: sux101@psu.edu

Sudeepta Mondal

Mechanical Engineering Department,
Pennsylvania State University,
University Park, PA 16802-1412
e-mail: sbm5423@psu.edu

Asok Ray[1]

Fellow ASME
Mechanical Engineering Department,
Pennsylvania State University,
University Park, PA 16802-1412
e-mail: axr2@psu.edu

Real-time detection and decision and control of thermoacoustic instabilities in confined combustors are challenging tasks due to the fast dynamics of the underlying physical process. The objective here is to develop a dynamic data-driven algorithm for detecting the onset of instabilities with short-length time-series data, acquired by available sensors (e.g., pressure and chemiluminescence), which will provide sufficient lead time for active decision and control. To this end, this paper proposes a Bayesian nonparametric method of Markov modeling for real-time detection of thermoacoustic instabilities in gas turbine engines; the underlying algorithms are formulated in the symbolic domain and the resulting patterns are constructed from symbolized pressure measurements as probabilistic finite state automata (PFSA). These PFSA models are built upon the framework of a (low-order) finite-memory Markov model, called the D-Markov machine, where a Bayesian nonparametric structure is adopted for: (i) automated selection of parameters in D-Markov machines and (ii) online sequential testing to provide dynamic data-driven and coherent statistical analyses of combustion instability phenomena without solely relying on computationally intensive (physics-based) models of combustion dynamics. The proposed method has been validated on an ensemble of pressure time series from a laboratory-scale combustion apparatus. The results of instability prediction have been compared with those of other existing techniques. [DOI: 10.1115/1.4037288]

1 Introduction

Thermoacoustic instabilities accrue from nonlinear interactions between the unsteady heat release and acoustics in the confined chamber of a combustor [1–3]. Consequently, if the self-sustained pressure oscillations have high amplitudes, the performance and operational life of machineries (e.g., gas turbine engines), which use such combustors, could be adversely affected. Therefore, it is imperative to appropriately design and operate combustors to ensure timely detection and mitigation of thermoacoustic instabilities, which may require model-based and/or dynamic data-driven analysis involving analyses of time series of pressure oscillations and flame images.

The difficulties in handling the complex nonlinear dynamics, as encountered in the analysis of combustion instabilities, often limit the applications of physics-based modeling tools for anomaly detection and decision and control of combustion dynamics at different operating conditions. To this end, there has been much interest in early detection of thermoacoustic instabilities from the perspectives of dynamic data-driven application systems [4,5]. For example, Nair and Sujith [6] have used the local flow test and 0–1 chaos test to study the chaotic structure of pressure time series and concluded that the pressure measurements during stable operation in combustors undergo deterministic chaos and gradually relax their chaotic behavior when the system approaches an unstable condition. It is also claimed that low-amplitude irregular pressure fluctuations therein possess multifractal structures, which may contain valuable prognostic information for early detection of thermoacoustic instabilities. Similarly, Gianni et al. [7] have used a topologically invariant index to recognize the transition mechanism leading to thermoacoustic instabilities, which is shown to be an early precursor. Other researchers believe that time series acquired in combustors can be modeled as a random process, such as a Markov chain, whose model structure may reflect the physical nature of the combustion process. Much research efforts have been expended on Markov chain modeling, and several popular strategies are listed in Refs. [8–11]. Hauser et al. [12] have recently used flame images to detect the onset of combustion instabilities in the framework of symbolic time-series analysis. Neural networks have also been used for such imaging-based analysis [13].

Implementations of dynamic data-driven application systems [4] include predictions of flame lean-blowout [14] and instabilities [5] in laboratory-scale combustors using symbolic time-series analysis. Especially, Sarkar et al. [15] reported an information-theoretic state-splitting and state-merging algorithm to model flame lean-blowout phenomena in the framework of a special class of probabilistic finite state automata (PFSA), called D-Markov machines [16,17], whose entropy rate may also be used to detect the thermoacoustic instabilities. However, this tree-based algorithm lacks a coherent probabilistic interpretation and may not able to accommodate more general interactions between current measurement and measurement history. Moreover, the task of parameter estimation becomes difficult even for moderately high depth (e.g., $D > 1$) in D-Markov machines due to the paucity of time-series data. As for the detection procedure, this method is based on an empirical threshold rather than the statistical detection theory. This method may not be sequentially implementable, thus possibly limiting its use in real-time detection of thermoacoustic instabilities.

The current paper, which is a major extension of the work [18] reported recently by the authors in a conference, proposes a Bayesian nonparametric approach [19] to address the previously mentioned difficulties. This approach automatically selects maximal order and lags (i.e., measurement history) of a D-Markov machine from limited (time series) data to formulate a parsimonious representation. Then, to detect the thermoacoustic instabilities in real time, a sequential online testing algorithm is developed upon the constructed D-Markov machine. The proposed algorithm has been validated on pressure measurements from a laboratory-scale swirl-stabilized combustor and the results are compared with those of existing techniques. From the perspectives of active control of combustion instabilities (e.g., see Ref. [20]), the proposed method may serve as a statistical filter to predict the system state with a high level of confidence and thus potentially improve the performance of active controllers by reducing the (potentially destabilizing) delay in the feedback loop.

This paper is organized into four sections including the present section. Section 2 describes the technical approach for data-driven

modeling of pressure time series and for detection of thermoacoustic instabilities. Section 3 validates the theoretical results on the data collected from a laboratory-scale combustion apparatus. Section 4 summarizes this paper along with conclusions and suggested topics of future research.

2 Technical Approach

This section describes the details of the technical approach that is adopted for data-driven modeling of pressure time series and for detection of thermoacoustic instabilities. The proposed Bayesian nonparametric methodology constructs a D-Markov machine [16,17] from real-valued time series, which are then used for online sequential testing.

2.1 Symbolization.
To construct a D-Markov machine, the (finite-length) data of pressure time series are converted into symbol strings through symbolization, also known as quantization, that requires partitioning of the signal space into a finite number of mutually exclusive and exhaustive regions, each corresponding to a unique symbol. The (finite) set of symbols is called the alphabet Σ and its cardinality is called the alphabet size, denoted as $|\Sigma|$.

Now the notion of maximum entropy partitioning (MEP) [21] is introduced as a partitioning tool for the ensemble of time-series data. The entropy of a generated symbol string is maximized by MEP, which implies that the information-rich cells of a data set are partitioned finer and those with sparse information are partitioned coarser; in other words, each cell contains (approximately) equal number of data points under MEP. Considerations for the choice of an alphabet size $|\Sigma|$ include the maximum discrimination capability of a symbol sequence and the associated computational complexity [17].

DEFINITION 2.1. *(D-Markov machine) A symbol sequence is called a D-Markov machine (i.e., a Markov chain of depth D) if, conditioned on the most recently generated D symbols, the distribution of the current symbol is independent of its more distant past. That is,*

$$p(y_t|y_{t-1}, ..., y_{t-D}, ...) = p(y_t|y_{t-1}, ..., y_{t-D}) \quad (1)$$

where y_t denotes the symbol at the time instant t. It is noted that the actual lags (i.e., delayed measurements) to identify the distribution of y_t could be a subset of $(y_{t-1}, ..., y_{t-D})$ and that the maximal order, which is less than or equal to D, is the least positive integer beyond which the lags are not significant.

The assumption of finite-length memory in the definition of a D-Markov machine is reasonable for many (statistically stationary) engineering systems with fading memory, especially combustion systems [12], which tend to forget their initial conditions and distant past rather quickly. Apparently, the mechanism of thermoacoustic instabilities can be better understood through the distinct structure of the D-Markov machine constructed from stable and unstable pressure measurements. However, a significant challenge is how to determine the maximal order and respective lags, which are critical for identification of the probability distribution of the current symbol. As possible combinations of lags and the number of parameters to be estimated increase rapidly with depth D and alphabet size $|\Sigma|$, it is desirable to have a flexible and interpretable procedure to infer the D-Markov machine from limited data. This issue is addressed in Secs. 2.2–2.4.

2.2 Conditional Tensor Factorization.
The conditional probability density $p(y_t|y_{t-1}, ..., y_{t-D})$ is now treated as a $(D+1)$-order tensor in the $|\Sigma|$-dimensional space, which is hereafter called the conditional probability tensor. It was first reported by Yang and Dunson [22] that every conditional probability tensor has the following higher order singular value decomposition:

$$p(y_t|y_{t-1}, ..., y_{t-D}) = \sum_{s_1=1}^{k_1} \cdots \sum_{s_D=1}^{k_D} \lambda_{s_1,...,s_D}(y_t) \prod_{j=1}^{D} \omega_{s_j}^{(j)}(y_{t-j}) \quad (2)$$

where $1 \le k_j \le |\Sigma|$ for $j = 1, ..., D$ and the parameters $\lambda_{s_1...s_D}(y_t)$ and $\omega_{s_j}^{(j)}(y_{t-j})$ are all non-negative real and satisfy the following constraints:

$$\sum_{y_t=1}^{|\Sigma|} \lambda_{s_1,...,s_D}(y_t) = 1 \quad \text{for each } (s_1, ..., s_D) \quad (3)$$

$$\sum_{s_j=1}^{k_j} \omega_{s_j}^{(j)}(y_{t-j}) = 1 \quad \text{for each } (j, y_{t-j}) \quad (4)$$

Because such a factorization exists for every conditional probability tensor, the previously mentioned constraints are not restrictive but they ensure that $\sum_{y_t=1}^{|\Sigma|} p(y_t|y_{t-1} \cdots y_{t-D}) = 1$.

2.3 Bayesian Nonparametric Modeling.
For development of a statistically interpretable and parsimonious model, the tensor factorization in Sec. 2.2 is converted into a Bayes network by introducing latent allocation variables and assigning sparsity-inducing priors. More formally, let $\{y_1, y_2, ..., y_T\}$ be a time-indexed symbol string (of length T) with finite memory of D; the elements y_k of the symbol string are symbols from the alphabet Σ.

Let s_j be the realization of a latent allocation variable $x_{j,t}$, whose support is $\{1, 2, ..., k_j\}$ all time t. Then, the transition probability $p(y_t|y_{t-1}, ..., y_{t-D})$, factorized as in Eq. (2), can be rewritten in the following form:

$$p(y_t|y_{t-1} \cdots y_{t-D}) = \int_{x_{1,t}} \cdots \int_{x_{D,t}} p(y_t|x_{1,t}, ..., x_{D,t}) \prod_{j=1}^{D} p(x_{j,t}|y_{t-j})$$
$$(5)$$

where $x_{j,t}$, for $j = 1, ..., D$ and $t = D+1, ..., T$, are latent allocation variables and follow the distribution:

$$(x_{j,t}|y_{t-j}) \sim \mathbf{Mult}\left(\boldsymbol{\omega}_{k_j}^{(j)}(y_{t-j})\right) \quad (6)$$

$$(y_t|x_{1,t}, ..., x_{D,t}) \sim \mathbf{Mult}(\boldsymbol{\lambda}_{s_1,...,s_D}) \quad (7)$$

where $\mathbf{Mult}$ stands for multinomial distribution [23], $\boldsymbol{\omega}_{k_j}^{(j)}(y_{t-j}) = \left\{(\omega_i^{(j)}(y_{t-j})\right\}_{i=1}^{k_j}$, and $\boldsymbol{\lambda}_{s_1,...,s_D} = \{\lambda_{s_1,...,s_D}(i)\}_{i=1}^{|\Sigma|}$.

The previously mentioned hierarchical reformulation of higher order singular value decomposition illustrates the features of the model in two ways: first, Eq. (6) shows that soft clustering is implemented for each lag y_{t-j} across the alphabet Σ to borrow statistical strength among different symbols. Then, by Eq. (7), the clustering assignments $x_{j,t}$ are used to capture the interactions among the lags in an implicit and parsimonious manner by allowing the latent populations indexed by $(s_1, ..., s_D)$ to be shared among the various state combinations of the lags.

It is noted that the number of latent classes for jth lag (always less than or equal to k_j) determines the inclusion of lag y_{t-j} in the model, because $p(y_t|y_{t-1} \cdots y_{t-D})$ does not vary with y_{t-j} when there is only one latent class. Therefore, only one latent class is retained by eliminating the unnecessary lags, and the maximal order is determined by identifying the most distant lag with multiple latent classes. However, in real-life applications, the tensor $\lambda_{s_1...s_D}(y_t)$ still may have more components than required, because $\prod_{j=1}^{D} k_j$ could be large for moderate values of D and $|\Sigma|$. To alleviate this difficulty, $\lambda_{s_1,...,s_D}(y_t)$ is clustered among different combinations of $(s_1, ..., s_D)$ in a nonparametric way by imposing the Dirichlet process prior [24] on it. Thus, employing the stick-breaking representation of Dirichlet Process [25], it follows that:

$$\pi | \gamma \sim \mathbf{GEM}(\gamma) \tag{8}$$

$$\theta_{s_1,\ldots,s_D} | \pi \sim \mathbf{Mult}(\pi) \tag{9}$$

$$\lambda_l = \{\lambda_l(1), \ldots, \lambda_l(|\Sigma|)\} | \alpha \sim \mathbf{Dir}(\alpha) \tag{10}$$

$$y | \theta, \lambda, (x_{j,t}, j = 1, \ldots, D) \sim \mathbf{Mult}(\lambda_{\theta_{x_{1,t},\ldots,x_{D,t}}}) \tag{11}$$

where **GEM** stands for Griffiths, Engen, and McCloskey process [23], **Dir** represents Dirichlet distribution, $\theta = \{\theta_{s_1,\ldots,s_D}\}_{s_1,\ldots,s_D}$, and $\lambda = \{\lambda_l\}_{l=1}^{\infty}$. Finally, the priors on mixture probability vectors $\omega_{k_j}^j(y_{t-j})$ and k_j are set as follows:

$$\omega_{k_j}^j(z_{j,t}) | k_j, z_{j,t}, \beta_j \sim \mathbf{Dir}(\beta_j) \tag{12}$$

$$p(k_j = k) | \mu \propto \exp(-\mu j k) \tag{13}$$

where $\mu > 0$ and $z_{j,t} = y_{t-j}$. The prior assigns increasing probabilities to smaller values of k_j as the lag j becomes more distant, reflecting the natural belief that increasing lags have diminishing impact on the distribution.

A Bayes network representation of the model is obtained by combining Eqs. (6)–(13) and its dependency structure is summarized in Fig. 1. Although the (nonparametric) posterior distribution has no analytic form, the inference of this Bayes network can be accomplished by Markov chain Monte Carlo (MCMC) simulation, as described in Algorithm 1.

To execute Algorithm 1, several hyperparameters, i.e., μ, α, and β_j (see Fig. 1), need to be chosen. The implication and determination of μ have been addressed earlier and those of other hyperparameters are discussed here. It is noted that α and β_j are hyperparameters of Dirichlet distribution and serve as pseudo-counts. Their determination is dependent on the users' prior belief and often they are chosen to be small values when no additional information is available. In this paper, these hyperparameters are chosen to be: $\alpha = 1$ and $\beta_j = 1/3$.

Algorithm 1 MCMC algorithm for inference

Require: Depth D, Alphabet size $|\Sigma|$, truncating components L, number of samples N, symbol sequence $\{y_t\}_{t=1}^T$ and initial $\theta, \pi, \lambda, \omega = \{\omega_{kj}^{(j)}(z)\}_{z=1,\ldots,|\Sigma|}^{j=1,\ldots,D}, x = \{x_{j,t}\}_{t=D+1,\ldots,T}^{j=1,\ldots,D}, k = \{k_j\}_{j=1,\ldots,D}$

Ensure: Posterior samples $\{_{(n)}\theta, _{(n)}\pi, _{(n)}\lambda, _{(n)}\omega, _{(n)}x, _{(n)}k\}_{n=1}^N$

1: **for** $n = 1$ to N **do**
2:   For each $(s_1, \ldots, s_D)$, sample $\theta_{s_1,\ldots,s_D}$ from its multinomial full conditionals: $p(\theta_{s_1,\ldots,s_D} = l | \xi) \propto \pi_l \prod_{c=1}^{|\Sigma|} \{\lambda_l(c)\}^{n_{s_1,\ldots,s_D}(c)}$ where $n_{s_1,\ldots,s_D}(c) = \sum_{D+1}^T 1\{x_{1,t} = s_1, \ldots, x_{D,t} = s_D, y_t = c\}$.
3:   For $l = 1, \ldots, L$, update π_l as follows: $p(V_l | \xi) = \mathbf{Beta}(1 + n_l, \gamma + \sum_{k>l} n_k)$; $\pi_l = V_l \prod_{m=1}^{l-1}(1 - V_m)$, where $n_l = \sum_{(s_1,\ldots,s_D)} 1\{\theta_{s_1,\ldots,s_D} = l\}$.
4:   For $l = 1, \ldots, L$, sample λ_l from their Dirichlet full conditionals $\lambda_l | \xi \sim \mathbf{Dir}\{\alpha + n_l(1), \ldots, \alpha + n_l(|\Sigma|)\}$, where $n_l(c) = \sum_{(s_1,\ldots,s_D)} 1\{\theta_{s_1,\ldots,s_D} = l\} n_{s_1,\ldots,s_D}(c)$.
5:   For $j = 1, \ldots, D$ and $c = 1, \ldots, |\Sigma|$, sample $\{\omega_1^{(j)}(c), \ldots, \omega_{k_j}^{(j)}(c)\} | \xi \sim \mathbf{Dir}\{\beta_j + n_{j,c}(1), \ldots, \beta_j + n_{j,c}(k_j)\}$, where $n_{j,c}(s_j) = \sum_{t=D+1}^T 1\{x_{j,t} = s_j, z_{j,t} = c\}$.
6:   For $j = 1, \ldots, q$ and for $t = D+1, \ldots, T$, sample the $x_{j,t}$ from their multinomial full conditionals: $p(x_{j,t} = s | \xi, x_{l,t} = s_l, l \neq j) \propto \omega_s^{(j)}(z_{j,t}) \lambda_{\theta_{s_1,\ldots,s,\ldots,s_D}}(y_t)$: where ξ is the collection of variables that are not explicitly mentioned.
7:   For $j = 1, \ldots, D$, sample k_j using their multinomial full conditionals: $p(k_j = k | \xi) \propto \exp(-\mu j k) \prod_{c=1}^{|\Sigma|} n_{j,c}^{-k\beta_j}$, $k_j = \max_t\{x_{j,t}\}, \ldots, |\Sigma|$, where $n_{j,c} = \sum_{t=D+1}^T 1\{z_{j,t} = c\}$.
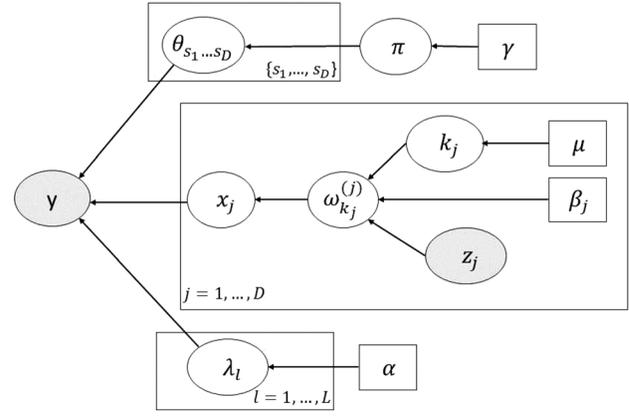8: **end for**

Fig. 1 Graphical representation of the Bayes network: $x_j = \{x_{j,t}\}_{t=D+1}^T$; $y = \{y_t\}_{t=D+1}^T$; and $z_j = \{z_{j,t}\}_{t=D+1}^T$. **Rectangle $\equiv$ deterministic hyperparameter; shaded ellipse $\equiv$ observed random variable; and transparent ellipse $\equiv$ unobserved random variable.**

Remark 2.1. By assigning sparsity-inducing priors on the decomposed algebraic structure, the proposed method yields a parsimonious model that inherits its statistical strength from different categories and predictors to improve the estimation accuracy especially if the data are limited. This allows the proposed method to infer a high-order Markov model from limited data compared to frequency counting methods. Experimental validation is presented in Sec. 3.

2.4 Online Sequential Testing. The proposed method for detection of thermoacoustic instabilities is comprised of an offline training phase and an online testing phase. All the training and testing data of pressure time-series are required to be symbolized with the same alphabet size using the MEP [21], and the classification process is conducted based on the generated symbol strings. In the training phase, posterior samples $\{_{(n)}^{(i)}\theta, _{(n)}^{(i)}\lambda, _{(n)}^{(i)}\omega\}_{n=1}^N$ are obtained from a training symbol string $^{(i)}y$ for each class i, where zero represents the stable class and one represents the unstable class. In the testing phase, the conditional probability $p(y|^{(i)}y)$ for an observed symbol string y with length L is calculated using posterior samples as follows:

$$p(y|^{(i)}y) = \prod_{t=D+1}^L p(y_t | y_{t-1}, \ldots, y_{t-D}, ^{(i)}y) \tag{14}$$

$$p\left(y_t | y_{t-1}, \ldots, y_{t-D}, ^{(i)}y\right)$$
$$\approx \frac{1}{N} \sum_{n=1}^N \left(\sum_{s_1=1}^{k_1} \cdots \sum_{s_D=1}^{k_D} {}_{(n)}^{(i)} \lambda_{{}_{(n)}^{(i)}\theta s_1, \ldots, s_D}(y_t) \prod_{j=1}^D {}_{(n)}^{(i)} \omega_{s_j}^{(j)}(y_{t-j}) \right) \tag{15}$$

Based on the conditional probability $p(y|y^i)$, the likelihood ratio test [26] is constructed as follows:

$$\frac{p\left(y|^1y\right)}{p\left(y|^0y\right)} \underset{0}{\overset{1}{\gtrless}} \tau \tag{16}$$

where τ is the threshold; one criterion to choose the threshold τ is the receiver operating characteristic (ROC) [26]. The ROC curve, which is obtained by varying τ, provides a trade-off between the probability of successful detection $p_D = p(\text{decide } 1 | 1 \text{ is true})$ and the probability of false alarms $p_F = p(\text{decide } 1 | 0 \text{ is true})$. A combination of p_D and the test data length for a given p_F is selected based on ROC curves, which would lead to a choice of the
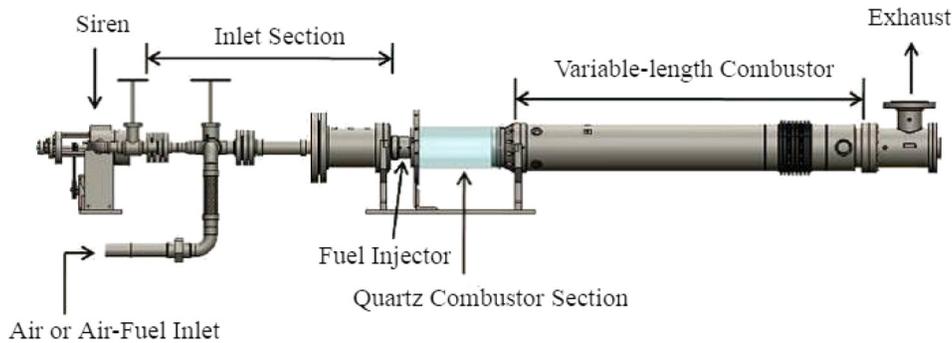
Fig. 2   Schematic diagram of the combustion apparatus

threshold τ. This detection algorithm is sequential, because the conditional probability $p(y|^{(i)}y)$ could be evaluated sequentially as shown in Eq. (14). For fast implementation, the values of $p(y_t|y_{t-1}, \ldots, y_{t-D}, {}^{(i)}y)$ in Eq. (15) can be precomputed and stored for different combinations of $(y_t, y_{t-1}, \ldots, y_{t-D})$.

3   Experimental Data Collection

This section presents the experimental details for data collection from a laboratory-scale apparatus with the objective of analyzing the nonlinear dynamics that occur during the instability phenomena.

3.1   Description of the Test Apparatus. The test apparatus is built upon a swirl-stabilized, lean-premixed, laboratory-scale combustor [27] that has been used to perform the experimental investigation. Figure 2 shows a schematic diagram of the variable-length combustor that consists of an inlet section, an injector, a combustion chamber, and an exhaust section. The combustor chamber consists of an optically accessible quartz section followed by a variable-length steel section. High-pressure air is delivered to the apparatus from a compressor system after passing through filters to remove any liquid or solid particles that might be present in the inlet air. The air supply pressure is set to approximately 1.34 MPa using a dome pressure regulator. The air is preheated to a maximum temperature of 250 °C by an 88 kW electric heater. The fuel for this study is natural gas (approximately 95% methane). It is supplied to the system at a pressure of approximately 1.48 MPa. The flow rates of the air and natural gas are measured by thermal mass flow meters. The desired equivalence ratio and mean inlet velocity are set by adjusting these flow rates with needle valves. Tests are conducted at a nominal combustor pressure of 1 atm over a range of operating conditions, as listed in Table 1. Under each operating condition, 8 s pressure time series are collected at a sampling rate of 8192 Hz.

3.2   Algorithm Validation on Experimental Data. This subsection analyzes the experimental data of pressure oscillations collected from the test apparatus. The details of data analysis are presented later.

3.2.1   Construction of D-Markov Machines. Oversampling is often preferred for data collection from engineering perspectives and may mask the true nature of system dynamics when real-valued time-series data are symbolized. To avoid such a problem,

Table 1   Operating conditions

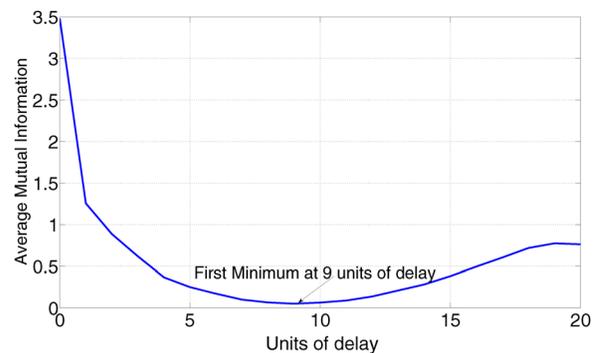| Parameters	Value
Equivalence ratio	0.525, 0.55, 0.60, 0.65
Inlet velocity	25–50 m/s in 5 m/s increments
Combustor length	25–59 (in) in 1 (in) increments

**Fig. 3   Profile of the average mutual information**

the pressure measurements from combustors are first down-sampled by a factor that is representative of a typical lag, which is the first minimum of the average mutual information plot [28], as demonstrated in Fig. 3. Then, the time-series data are symbolized through maximum entropy partitioning [21] with a ternary alphabet $\Sigma = \{1, 2, 3\}$ for both stable and unstable cases, where the ground truth is decided such that the root mean square value of pressure greater than 0.483 kPa (0.07 psi) indicates an unstable situation.

To construct a D-Markov machine, symbolized pressure time-series data are analyzed using the proposed Bayesian nonparametric method. The hyperparameters in Eqs. (8)–(13) are set at $\mu = 0.5$, $\gamma = 1$, $\alpha = 1/3$, and $\beta_j = 1/3$ for each j. For each case, the data length is 500 with depth $D = 10$ and 2000 posterior samples are collected from 10,000 MCMC iterations. Figures 4 and 5 summarize the MCMC results for pressure data under stable and unstable cases, respectively. Figures 4(a), 4(b), 5(a), and 5(b) illustrate the method's ability to identify the maximal order and relevant lags of the D-Markov machine. The proposed method also leads to parsimonious representations, as shown in Figs. 4(c), 4(d), 5(c), and 5(d). A comparison of Figs. 4(a) and 5(a) reveals that the maximal order for the unstable case is significantly higher than that for the stable case, which indicates a more deterministic (i.e., less random) behavior and is in agreement with the previously published works [15,29].

3.3   Sequential Testing. One hundred samples of pressure measurements are selected from each of stable phase (class 0) and unstable phase (class 1) to serve as test data. Figure 6 shows the posterior probability of each class as a function of the length of the observed data. It is seen in Fig. 6 that the observed time series is correctly be classified as unstable because the posterior probability of class 1 approaches one, while that of class 0 approaches zero very fast. For all the test data, the posterior probability correctly converged to either zero or one. Figure 7 exhibits a family of ROC curves for the proposed detection algorithm with varying length of test data. It is observed that the ROC curve improves
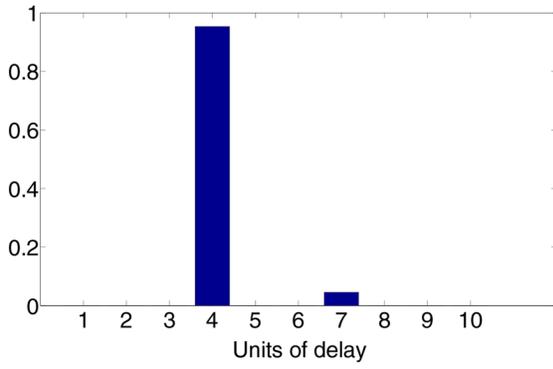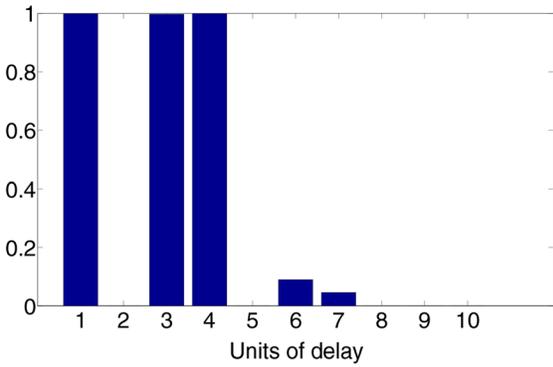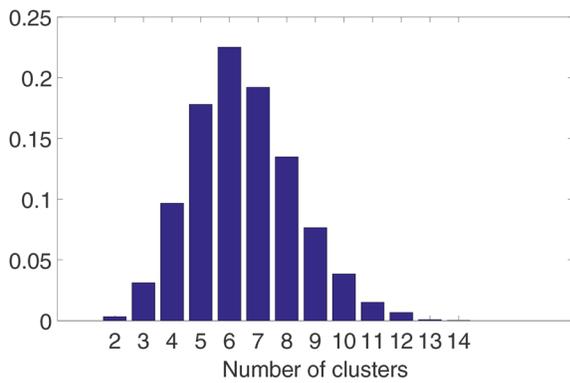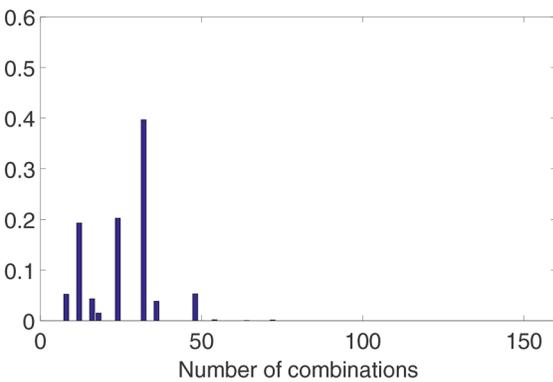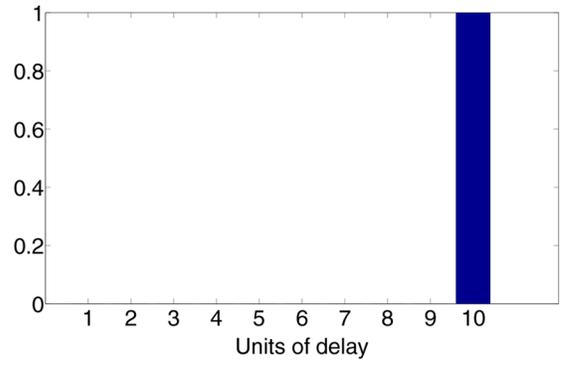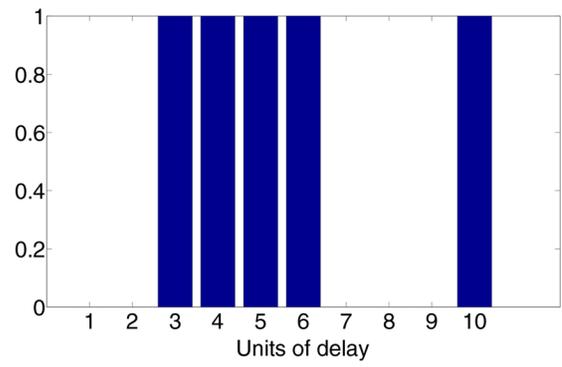
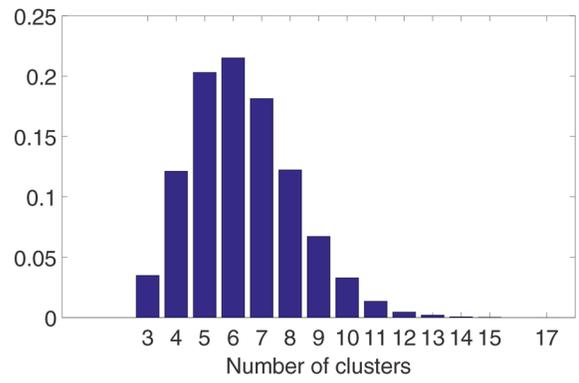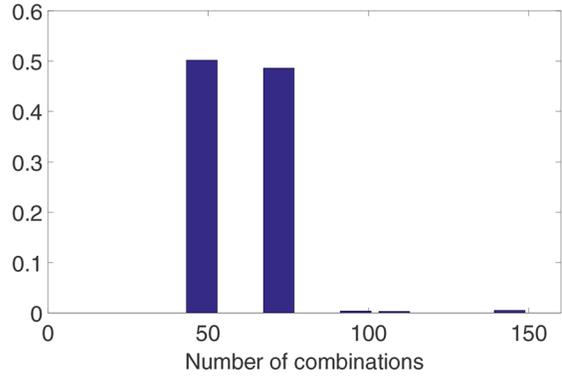Fig. 4 MCMC results for histograms of pressure measurements under stable operation: (*a*) histogram of the maximal order (i.e., depth *D*), (*b*) inclusion proportions of different lags, (*c*) histogram of the number of clusters of the tensor $\lambda_{s_1,\ldots,s_D}$, and (*d*) histogram of combinations of realizations $(s_1,\ldots,s_D)$

Fig. 5 MCMC results for histograms of pressure measurements under unstable operation: (*a*) histogram of the maximal order (i.e., depth *D*), (*b*) inclusion proportions of different lags, (*c*) histogram of the number of clusters of the tensor $\lambda_{s_1,\ldots,s_D}$, and (*d*) histogram of combinations of realizations $(s_1,\ldots,s_D)$
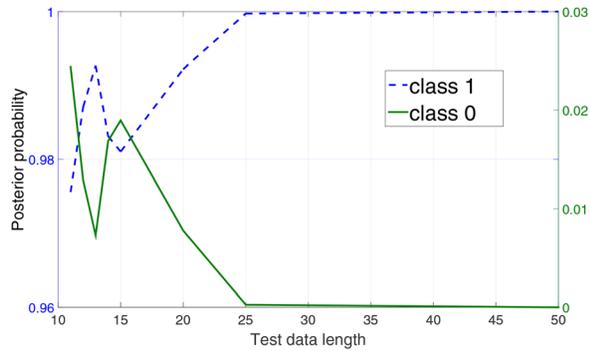
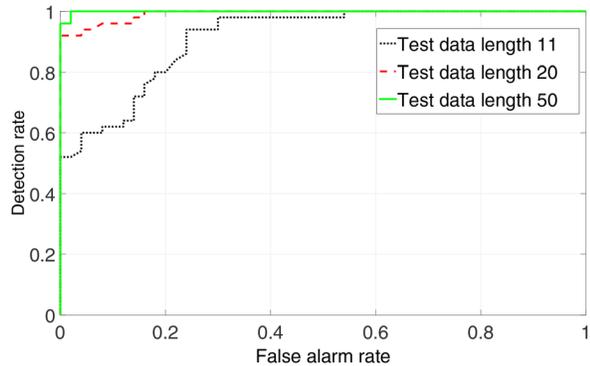**Fig. 6 Posterior probabilities for an unstable test sequence**

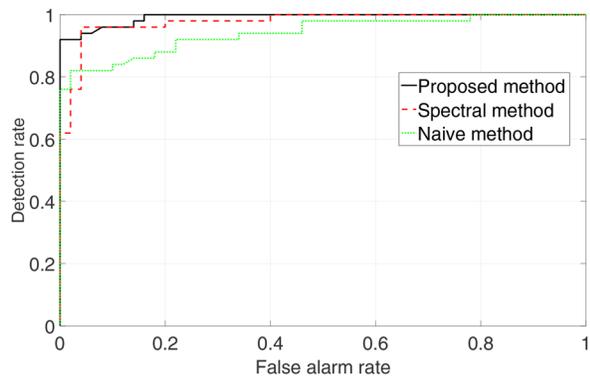**Fig. 7 ROC curves with different test data lengths**

Fig. 8 ROC curves with test data length = 20

(i.e., moves toward the top left corner) considerably as the test data length is increased from 11 to 50. With a test data length of 20, Fig. 8 compares the proposed method with two other reported methods: (i) naive method that selects the maximal order of the D-Markov machine to be 1 and (ii) spectral method that optimizes the maximal order [11]. Both these methods use the frequency counting to estimate the parameters of D-Markov machine. The proposed method yields larger detection rates p_D at different specified false alarm rates p_F, as compared to both naive and spectral methods. The implication is that a larger p_D indicates improved efficacy of the classifier to detect the onset of unstable operating conditions at a given p_F, and this timely information is crucial for adaptation of the combustion system's behavior to a stable operation. Since the proposed method is capable of executing these actions with a very short data length, it is ideally suited for early detection and control of combustion instabilities in real-life combustion systems.

4 Summary, Conclusions, and Future Work

This paper has proposed a dynamic data-driven algorithm for detecting the onset of thermoacoustic instabilities in confined combustors, which requires very short lengths of time-series data and thus provides sufficient lead time for active decision and control. The algorithm is developed in a nonparametric Bayesian setting and is built upon a special class of probabilistic finite state automata, called D-Markov machines [16,17]. These D-Markov machines are constructed from symbol strings, generated by partitioning of (finite length) time series with automated selection of a set of lags for parsimonious representation.

The proposed method has been validated with experimental data of pressure oscillations from a laboratory-scale swirl-stabilized combustor apparatus. Analysis of the experimental data has revealed that the maximal order (i.e., depth D) for pressure time series under unstable operations is relatively higher than those under stable operations, which implies a more deterministic (i.e., less random) behavior of unstable operations compared to that of stable operations.

While there are many areas of theoretical and experimental research to be conducted before this method can be used for real-time monitoring and active control of combustors in industrial applications, the authors suggest the following areas for research topics in the near future:

(1) Treatment of the operating conditions as exogenous labels for pressure time series and extension of the proposed method to spatial-temporal symbol strings.
(2) Integration of the proposed algorithm with those of active combustion control to improve the performance of combustion systems.
(3) Investigation on scalability and feasibility of the algorithm for wider ranges of operation.
(4) Usage of other methods (e.g., confusion matrices), in addition to ROC curves, for performance comparison of different methods.
(5) Theoretical research on trade-off between alphabet size $|\Sigma|$ and depth D for construction of D-Markov machines, supported by experimental validation.

Acknowledgment

References

[1] Lieuwen, T., Torres, H., Johnson, C., and Zinn, B. T., 2000, "A Mechanism of Combustion Instability in Lean Premixed Gas Turbine Combustors," ASME J. Eng. Gas Turbines Power, **123**(1), pp. 182–189.
[2] Dowling, A., and Hubbard, S., 2000, "Instability in Lean Premixed Combustors," Proc. Inst. Mech. Eng. Part A, **214**(4), pp. 317–332.
[3] Huang, Y., and Yang, V., 2009, "Dynamics and Stability of Lean-Premixed Swirl-Stabilized Combustion," Prog. Energy Combust. Sci., **35**(4), pp. 293–364.
[4] Darema, F., 2005, "Dynamic Data Driven Applications Systems: New Capabilities for Application Simulations and Measurements," Fifth International Conference on Computational Science (ICCS), Atlanta, GA, May 22–25, pp. 610–615.
[5] Sarkar, S., Chakravarthy, S. R., Ramanan, V., and Ray, A., 2016, "Dynamic Data-Driven Prediction of Instability in a Swirl-Stabilized Combustor," Int. J. Spray Combust. Dyn., **8**(4), pp. 235–253.
[6] Nair, V., and Sujith, R., 2014, "Multifractality in Combustion Noise: Predicting an Impending Combustion Instability," J. Fluid Mech., **747**, pp. 635–655.
[7] Gianni, G., Mariotti, G., Paganini, E., and Sello, S., 2003, "Characterization and Diagnostics of Combustion Thermoacoustic Instabilities Using Nonlinear Dynamics and Topological Methods," preprint arXiv:Physics/0306199.
[8] Bühlmann, P., and Wyner, A. J., 1999, "Variable Length Markov Chains," Ann. Stat., **27**(2), pp. 480–513.

[9] Teh, Y. W., 2006, "A Hierarchical Bayesian Language Model Based on Pitman-YOR Processes," 21st International Conference on Computational Linguistics and the 44th Annual Meeting of the Association for Computational Linguistics (ACL), Sydney, Australia, July 17–18, pp. 985–992.

[10] Jääskinen, V., Xiong, J., Corander, J., and Koski, T., 2014, "Sparse Markov Chains for Sequence Data," Scand. J. Stat., **41**(3), pp. 639–655.

[11] Jha, D. K., Srivastav, A., Mukherjee, K., and Ray, A., 2015, "Depth Estimation in Markov Models of Time-Series Data Via Spectral Analysis," American Control Conference (ACC), Chicago, IL, July 1–3, pp. 5812–5817.

[12] Hauser, M., Li, Y., Li, J., and Ray, A., 2016, "Real-Time Combustion State Identification Via Image Processing: A Dynamic Data-Driven Approach," American Control Conference (ACC), Boston, MA, July 6–8, pp. 3316–3321.

[13] Sarkar, S., Lore, K. G., and Sarkar, S., 2015, "Early Detection of Combustion Instability by Neural-Symbolic Analysis on Hi-Speed Video," International Conference on Cognitive Computation: Integrating Neural and Symbolic Approaches (COCO), Montreal, QC, Canada, Dec. 11–12, pp. 93–101.

[14] Sen, S., Sarkar, S., Chaudhari, R. R., Mukhopadhyay, A., and Ray, A., 2017, *Lean Blowout (LBO) Prediction Through Symbolic Time Series Analysis*, Springer, Singapore, pp. 153–167.

[15] Sarkar, S., Ray, A., Mukhopadhyay, A., and Sen, S., 2015, "Dynamic Data-Driven Prediction of Lean Blowout in a Swirl-Stabilized Combustor," Int. J. Spray Combust. Dyn., **7**(3), pp. 209–241.

[16] Ray, A., 2004, "Symbolic Dynamic Analysis of Complex Systems for Anomaly Detection," Signal Process., **84**(7), pp. 1115–1130.

[17] Mukherjee, K., and Ray, A., 2014, "State Splitting and Merging in Probabilistic Finite State Automata for Signal Representation and Analysis," Signal Process., **104**, pp. 105–119.

[18] Xiong, S., Li, J., and Ray, A., 2017, "Bayesian Nonparametric Modeling of Markov Chains for Detection of Thermoacoustic Instabilities," American Control Conference (ACC), Seattle, WA, May 24–26, pp. 3758–3763.

[19] Sarkar, A., and Dunson, D. B., 2015, "Bayesian Nonparametric Modeling of Higher Order Markov Chains," preprint arXiv:1506.06268.

[20] Annaswamy, A. M., and Ghoniem, A. F., 2002, "Active Control of Combustion Instability: Theory and Practice," IEEE Control Syst., **22**(6), pp. 37–54.

[21] Rajagopalan, V., and Ray, A., 2006, "Symbolic Time Series Analysis Via Wavelet-Based Partitioning," Signal Process., **86**(11), pp. 3309–3320.

[22] Yang, Y., and Dunson, D. B., 2015, "Bayesian Conditional Tensor Factorizations for High-Dimensional Classification," J. Am. Stat. Assoc., **111**(514), pp. 656–669.

[23] Pitman, J., 2002, *Combinatorial Stochastic Processes*, Springer, Berlin.

[24] Ferguson, T. S., 1973, "A Bayesian Analysis of Some Nonparametric Problems," Ann. Stat., **1**(2), pp. 209–230.

[25] Ishwaran, H., and James, L. F., 2001, "Gibbs Sampling Methods for Stick-Breaking Priors," J. Am. Stat. Assoc., **96**(453), pp. 161–173.

[26] Poor, H., 2013, *An Introduction to Signal Detection and Estimation*, Springer Science & Business Media, New York.

[27] Kim, K., Lee, J., Quay, B., and Santavicca, D., 2010, "Response of Partially Premixed Flames to Acoustic Velocity and Equivalence Ratio Perturbations," Combust. Flame, **157**(9), pp. 1731–1744.

[28] Abarbanel, H., 1996, *Analysis of Observed Chaotic Data*, Springer, New York.

[29] Nair, V., Thampi, G., Karuppusamy, S., Gopalan, S., and Sujith, R., 2013, "Loss of Chaos in Combustion Noise as a Precursor of Impending Combustion Instability," Int. J. Spray Combust. Dyn., **5**(4), pp. 273–290.