

Reliable Calculation of Thermoacoustic Instability Risk Using an Imperfect Surrogate Model

Shuai Guo¹

Mem. ASME

Professur für Thermofluidynamik,
Technische Universität München,
Boltzmannstr. 15,
Garching D-85748, Germany
e-mail: guo@fd.mw.tum.de

Camilo F. Silva

Professur für Thermofluidynamik,
Technische Universität München,
Boltzmannstr. 15,
Garching D-85748, Germany
e-mail: silva@fd.mw.tum.de

Wolfgang Polifke

Professur für Thermofluidynamik,
Technische Universität München,
Boltzmannstr. 15,
Garching D-85748, Germany
e-mail: polifke@tum.de

One of the fundamental tasks in performing robust thermoacoustic design of gas turbine combustors is calculating the modal instability risk, i.e., the probability that a thermoacoustic mode is unstable, given various sources of uncertainty (e.g., operation or boundary conditions). To alleviate the high computational cost associated with conventional Monte Carlo simulation, surrogate modeling techniques are usually employed. Unfortunately, in practice, it is not uncommon that only a small number of training samples can be afforded for surrogate model training. As a result, epistemic uncertainty may be introduced by such an “inaccurate” model, provoking a variation of modal instability risk calculation. In the current study, using Gaussian process (GP) as the surrogate model, we address the following two questions: First, how to quantify the variation of modal instability risk induced by the epistemic surrogate model uncertainty? Second, how to reduce the variation of risk calculation given a limited computational budget for the surrogate model training? For the first question, we leverage on the Bayesian characteristic of the GP model and perform correlated sampling of the GP predictions at different inputs to quantify the uncertainty of risk calculation. We show how this uncertainty shrinks when more training samples are available. For the second question, we adopt an active learning strategy to intelligently allocate training samples such that the trained GP model is highly accurate particularly in the vicinity of the zero growth rate contour. As a result, a more accurate and robust modal instability risk calculation is obtained without increasing the computational cost of surrogate model training. [DOI: 10.1115/1.4049314]

1 Introduction

The thermoacoustic behavior of a gas turbine combustor usually exhibits a high level of sensitivity to uncertain parameters such as operation conditions and acoustic boundary conditions [1]. As a result, conventional thermoacoustic instability predictions, which employ only the nominal values of the system parameters when calculating the modal frequencies and growth rates, may yield unreliable results, e.g., a combustor that is predicted as stable with nominal input parameters may become unstable when inputs deviate only slightly from their nominal values. Consequently, uncertainty quantification (UQ) analysis, which focuses on quantifying uncertainties on outputs given uncertain inputs, is essential to achieve reliable thermoacoustic instability predictions and constitutes a fundamental step toward robust thermoacoustic design.

One of the main tasks in thermoacoustic UQ analysis is calculating the modal instability risk (P_f), i.e., the probability that a thermoacoustic mode is unstable, given various uncertain parameters and their corresponding probability density functions (PDF). To calculate P_f induced by those uncertainties, a straight-forward method is to perform Monte Carlo simulation: first a large number of samples (at the order of 10^4) are drawn according to the PDFs of uncertain inputs. Then for each sample, its corresponding modal growth rate values are calculated via thermoacoustic solvers (e.g., acoustic network model [2], Helmholtz solver [3]). Finally, P_f can be simply estimated as the number of samples with positive modal growth rate values divided by the total sample number.

It is obvious that a direct application of Monte Carlo simulation using expensive thermoacoustic solvers would induce prohibitive computational cost, thus rendering it infeasible for realistic UQ

analysis. In order to solve the efficiency problem, various surrogate modeling techniques have been explored in previous studies. In this framework, first a small number of training samples (around 1% of full Monte Carlo sample size) are carefully selected according to the input PDFs. Their corresponding outputs (growth rate values) are calculated via thermoacoustic solver. Then, machine learning techniques are usually employed to train a cheap surrogate model, based on the training samples and their responses, to approximate the thermoacoustic solver. Subsequently, Monte Carlo simulation can be applied on the surrogate model to derive P_f with negligible cost. Typical examples of surrogate modeling techniques include polynomial regression [4], adjoint method [3,5,6], polynomial chaos expansion [7–9], as well as Gaussian process approach [10].

Despite the remarkable progress made in terms of accelerating P_f calculation, the uncertainty contained in the surrogate model itself has been largely ignored in previous studies. This issue becomes particularly prominent as the parameter variation ranges become larger and the fidelity of the thermoacoustic solver improves: a larger number of training samples is then required to build a surrogate model that is accurate over a larger parameter space. Meanwhile, every calculation of a training sample induces significant computational cost. Consequently, in many practical cases, only a small training sample size can be afforded, which leads to a potentially “inaccurate” surrogate model. As a result, epistemic uncertainty, which is defined as uncertainty induced by lack of knowledge or simplifications and can be reduced by collecting more information [11], is introduced by such a model. In this situation, P_f is not a deterministic value anymore, but rather an uncertain variable itself. To summarize, in the framework of surrogate model based uncertainty quantification, P_f exists due to the presence of system parameter uncertainty; P_f varies due to the presence of surrogate model uncertainty.

Among various surrogate modeling techniques, the Gaussian process (GP) approach stands out. This is due to the fact that GP

¹Corresponding author.

Manuscript received August 26, 2020; final manuscript received August 28, 2020; published online January 4, 2021. Editor: Jerzy T. Sawicki.

not only provides predicted values at unsampled inputs but also estimates the prediction variances, thanks to its Bayesian nature [12]. This unique feature of GP allows us to quantify the associated epistemic uncertainty when using GP to calculate P_f .

Therefore, in this study, we employ GP as the surrogate model. We aim to address the following two questions that are closely related to a realistic application of GP approach in thermoacoustic UQ analysis: (Q1) how to quantify the variation of modal instability risk P_f induced by the epistemic GP model uncertainty? Although surrogate modeling techniques have become popular in the thermoacoustic community to perform risk analysis, the extra uncertainties induced by imperfectly trained surrogate models have been largely ignored until now. To our best knowledge, our work is the first to quantify such uncertainties as well as the impact of these uncertainties on the variation of thermoacoustic risk calculation. (Q2) How to reduce the variation of P_f given a limited computational budget for the GP model training? Q2 directly builds on Q1: after understanding how to quantify the uncertainty of risk calculation, we now focus on how to reduce this uncertainty. This point is of particular interest to industrial applications. For the first question, we leverage on the prediction variance naturally offered by the GP model and perform correlated sampling of the GP predictions at different inputs [13] to quantify the uncertainty of P_f calculation. For the second question, we apply active learning methods [12] from the machine learning community, where the training samples are sequentially allocated such that the trained GP model is gradually refined particularly in the vicinity of the zero growth rate contour. As a result, we are able to obtain a more accurate and robust P_f calculation without increasing the computational cost for surrogate model training.

This work is closely related to our previous work [10], where our focus was to employ the GP approach to efficiently address various robust thermoacoustic design tasks. In that work, P_f calculations were carried out with the assumption that GP models have negligible epistemic uncertainty. In this work, we will drop this assumption and investigate how GP model uncertainty impacts the P_f calculation, as well as how to reduce this impact given a limited training sample budget. It is worth mentioning that the procedure of a reliable calculation of P_f established in this work can be seamlessly integrated into the robust design analysis investigated in our previous work [10], without compromising the associated workflow developed there.

This paper is organized as follows: The section “Thermoacoustic Framework” outlines the thermoacoustic problem under investigation. The section “Gaussian Process Modeling” presents the technical details of the employed GP approach and gives the formula for estimating the prediction variance. The section “Uncertainty of P_f Calculation” focuses on quantifying the P_f calculation uncertainty induced by the epistemic GP model uncertainty. Procedures are introduced first, followed by a case study to show how the uncertainty of P_f calculation shrinks when more GP training samples are available. The section “Reliable P_f Calculation” introduces the active learning scheme for adaptive GP model training. The case study in Sec. 4 is revisited and the effectiveness of the adopted active learning scheme is demonstrated. All the code and data to produce the results presented in this paper can be found at²

2 Thermoacoustic Framework

The thermoacoustic problem investigated in the current study is taken from our previous work [10]. Here, we briefly review the key aspects to ensure the self-containedness of the present paper.

For the combustor configuration, we investigate a turbulent pre-mixed swirl combustor test rig [14,15]. An equivalence ratio of 0.77 of perfectly premixed methane–air mixture and a thermal power of 30 kW are considered as the operation conditions.

We employ a low-order acoustic network model (shown in Fig. 1) to calculate the eigenmode of the burner test rig. Geometry, thermodynamic parameters, as well as modeling details can be found in Refs. [2] and [10]. The reflection coefficient at the combustor exit is modeled as $|R_{out}|e^{i\pi}$.

A flame impulse response (FIR) model describing the swirling flame dynamics is adopted in the current work. This model was proposed by Komarek et al. [14] and validated by Oberleithner et al. [16]. This model decomposes the shape of the flame impulse response into three Gaussian distributions: one positive (+) Gaussian to account for the flame response to axial velocity fluctuations, and one (+) combined with another (–) Gaussian to describe the response to swirl number fluctuations. The parametrization of the model is shown in Fig. 2. Here, τ_c represents the time needed for an inertial wave to propagate from the swirler to the flame base [17]. τ_1 and σ_1 represent the mean and standard deviation of the flame heat release response under an impulse axial velocity perturbation, respectively. τ_{s1} and τ_{s2} represent the characteristic time lags for the flame response to the swirl number fluctuations. The standard deviations of the associated positive and negative Gaussian functions are assumed to be one-third of τ_{s1} and τ_{s2} , respectively. Therefore, we can express the value of the FIR coefficient h_k as [14]

$$h_k = \frac{\Delta t}{\sigma_1 \sqrt{2\pi}} e^{-\frac{(k\Delta t - \tau_1)^2}{2\sigma_1^2}} + \frac{\Delta t}{\sigma_2 \sqrt{2\pi}} e^{-\frac{(k\Delta t - \tau_2)^2}{2\sigma_2^2}} - \frac{\Delta t}{\sigma_3 \sqrt{2\pi}} e^{-\frac{(k\Delta t - \tau_3)^2}{2\sigma_3^2}}, \quad k \in [1, 2, \dots, N] \quad (1)$$

where N is the number of FIR coefficients, and

$$\begin{aligned} \tau_2 &= \tau_c + \tau_{s1}, & \tau_3 &= \tau_c + \tau_{s2} \\ \sigma_2 &= \tau_{s1}/3, & \sigma_3 &= \tau_{s2}/3 \end{aligned} \quad (2)$$

In this study, we consider six uncertain parameters: five flame model parameters $\tau_1, \sigma_1, \tau_c, \tau_{s1}, \tau_{s2}$, and one acoustic boundary condition, i.e., the magnitude of the reflection coefficient at the combustor outlet $|R_{out}|$. Their nominal values and their respective

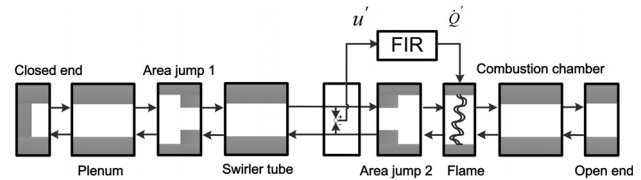


Fig. 1 Sketch of acoustic network model, flow from left to right

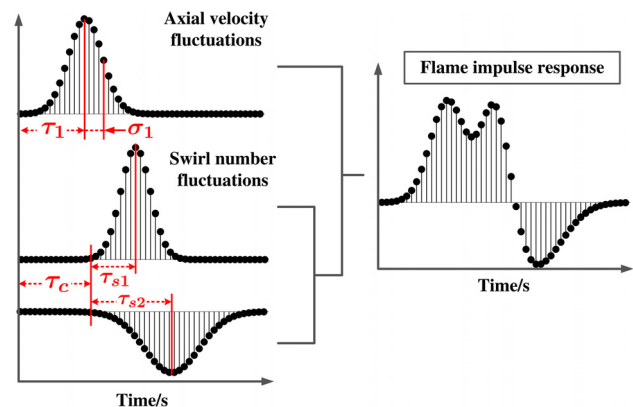


Fig. 2 Sketch of the employed FIR model for the flame dynamics

²<https://github.com/ShuaiGuo16/ASME20>

Table 1 Uncertainty information of the investigated parameters

| Parameters | | Nominal | Range |
|-------------------|-------------|---------------------|-----------------------------------|
| Flame (units: ms) | τ_1 | $\tau_1^0 = 2.85$ | $0.9\tau_1^0 - 1.1\tau_1^0$ |
| | σ_1 | $\sigma_1^0 = 0.7$ | $0.9\sigma_1^0 - 1.1\sigma_1^0$ |
| | τ_c | $\tau_c^0 = 3$ | 2–4.8 |
| | τ_{s1} | $\tau_{s1}^0 = 1.8$ | $0.9\tau_{s1}^0 - 1.1\tau_{s1}^0$ |
| | τ_{s2} | $\tau_{s2}^0 = 3.3$ | $0.9\tau_{s2}^0 - 1.1\tau_{s2}^0$ |
| Acoustic BC | $ R_{out} $ | $ R_{out} ^0 = 0.9$ | 0.6–1 |

uncertain ranges are taken from Ref. [10] and summarized in Table 1. In practice, it is preferable to construct a GP model that is valid over a large parameter space so that it can be reused for different tasks in the analysis process (e.g., various robust design tasks detailed in Ref. [10]). Therefore, we will build the GP model on the full parameter space as described in Table 1. In the section “Uncertainty of P_f Calculation,” we will reinvestigate “Q1: Risk Analysis” in Ref. [10] to assess the impact of epistemic surrogate model uncertainty on the calculation of P_f . Since only τ_1 , σ_1 , τ_{s1} , and τ_{s2} are considered as the uncertain parameters in that case study, we can effectively test the flexibility and robustness of the GP modeling approach.

3 Gaussian Process Modeling

In this section, we summarize the key features of GP surrogate modeling that are relevant for our current study. For a comprehensive introduction of GP approach, readers are referred to Ref. [18]. For the application of GP approach in the context of thermoacoustic instability analysis, readers are referred to Refs. [10] and [19–21].

GP approach models the output $f(\mathbf{x})$ of the high fidelity model at \mathbf{x} as the realization of a Gaussian process

$$f(\mathbf{x}) = \beta + Z(\mathbf{x}) \quad (3)$$

where β is a constant value and $Z(\mathbf{x})$ corresponds to a Gaussian process with zero mean and covariance defined as

$$\text{Cov}[Z(\mathbf{x}^i, \mathbf{x}^j)] = \sigma^2 R(\mathbf{x}^i, \mathbf{x}^j) \quad (4)$$

where σ^2 is the process variance and $R(\mathbf{x}^i, \mathbf{x}^j)$ is the correlation function between any two locations \mathbf{x}^i and \mathbf{x}^j in the input parameter space. In this study, we adopt a Gaussian correlation function to describe $R(\mathbf{x}^i, \mathbf{x}^j)$

$$R(\mathbf{x}^i, \mathbf{x}^j) = \exp \left[- \sum_{k=1}^M \theta_k (x_k^i - x_k^j)^2 \right] \quad (5)$$

where M denotes the dimension of the input (i.e., number of input parameters), and θ_k is a hyperparameter that controls the correlation strength between the points within dimension k .

Based on a set of training samples $\mathbf{X}_D = [\mathbf{x}^1, \dots, \mathbf{x}^N]^T$ and their corresponding responses $\mathbf{Y}_D = [f(\mathbf{x}^1), \dots, f(\mathbf{x}^N)]^T$, we can estimate β and σ^2 according to Ref. [18]

$$\hat{\beta} = (\mathbf{1}^T \mathbf{R}_D^{-1} \mathbf{1})^{-1} \mathbf{1}^T \mathbf{R}_D^{-1} \mathbf{Y}_D \quad (6)$$

$$\hat{\sigma}^2 = \frac{1}{N} (\mathbf{Y}_D - \hat{\beta})^T \mathbf{R}_D^{-1} (\mathbf{Y}_D - \hat{\beta}) \quad (7)$$

where $\mathbf{1}$ is a vector of ones of dimension M . \mathbf{R}_D is the N-by-N correlation matrix between training samples in \mathbf{X}_D . However, since Eq. (6) and (7) depend on θ_k through correlation matrix \mathbf{R}_D , it is first required to estimate θ_k using maximum likelihood estimation

$$\hat{\theta} = \underset{\theta}{\text{argmax}} \left[-\frac{N}{2} \ln(\hat{\sigma}^2) - \frac{1}{2} \ln(|\mathbf{R}_D|) \right] \quad (8)$$

Finally, the GP model predictions \mathbf{Y}_p at unknown locations $\mathbf{X}_p = [\mathbf{x}^1, \dots, \mathbf{x}^L]^T$ follow a multivariate normal distribution, with the mean predictions $\mu(\mathbf{Y}_p)$ and prediction covariance $\text{cov}(\mathbf{Y}_p)$ given as

$$\mu(\mathbf{Y}_p) = \hat{\beta} + \mathbf{R}_{PD}^T \mathbf{R}_D^{-1} (\mathbf{Y}_D - \hat{\beta}) \quad (9)$$

$$\text{cov}(\mathbf{Y}_p) = \hat{\sigma}^2 (\mathbf{R}_p - \mathbf{R}_{PD}^T \mathbf{R}_D^{-1} \mathbf{R}_{PD}) \quad (10)$$

where \mathbf{R}_{PD} represents the N-by-L correlation matrix between the prediction inputs \mathbf{X}_p and the training inputs \mathbf{X}_D . \mathbf{R}_p represents the L-by-L correlation matrix between the prediction inputs \mathbf{X}_p .

Specifically, when GP prediction y_p at only a single point \mathbf{x} is required, the corresponding prediction mean $\mu(y_p)$ and variance $\sigma^2(y_p)$ can be directly derived from Eqs. (9) and (10) as

$$\mu(y_p) = \hat{\beta} + \mathbf{r}(\mathbf{x})^T \mathbf{R}_D^{-1} (\mathbf{Y}_D - \hat{\beta}) \quad (11)$$

$$\sigma^2(y_p) = \hat{\sigma}^2 (1 - \mathbf{r}(\mathbf{x})^T \mathbf{R}_D^{-1} \mathbf{r}(\mathbf{x})) \quad (12)$$

In the section “Uncertainty of P_f Calculation,” Eqs. (9) and (10) are used to generate realizations of GP predictions to facilitate the quantification of P_f calculation uncertainty. In the section “Reliable P_f Calculation,” Eqs. (11) and (12) are used to derive the learning function to facilitate adaptive GP model training.

4 Uncertainty of P_f Calculation

In this section, we investigate the procedure to quantify the P_f calculation uncertainty and demonstrate the procedure through a case study. We start with the case setup, followed by detailing the uncertainty quantification procedure. Finally, we examine how the P_f calculation uncertainty evolves once more training samples are made available for GP model training.

4.1 Case Set-Up. In this work, we focus on one of the eigenmodes investigated in Ref. [10]: the quarter wave mode of the combustor, which is labeled as a *cavity mode* [22]. Given nominal values of the flame model parameters and $|R_{out}|$, the nominal values of the modal frequencies (ω) and growth rates (α) of the cavity mode are calculated via the acoustic network model (Fig. 1), i.e., ($\omega = 287.5$ Hz, $\alpha = -27.7$ rad/s) for cavity mode.

When uncertainties are present in flame model parameters and $|R_{out}|$, the cavity mode may exhibit a certain level of risk to be unstable. Here, we reinvestigate the problem outlined in “Q1: Risk Analysis” in Ref. [10]. For that problem, we need to calculate the P_f of cavity mode when τ_c and $|R_{out}|$ are fixed at their nominal values, while the rest of the parameters follow independent uniform distributions with the parameter range indicated in Table 1.

In GP-based Monte Carlo simulation, P_f of a thermoacoustic mode can be calculated as [13]

$$P_f = \frac{1}{n_{MC}} \sum_{i=1}^{n_{MC}} I(y_p(\mathbf{x}^i)), \quad y_p(\mathbf{x}^i) \in \mathbf{Y}_p \quad (13)$$

where n_{MC} denotes the number of Monte Carlo samples, \mathbf{x}^i is the i th Monte Carlo sample, $y_p(\mathbf{x}^i)$ represents the GP prediction of the modal growth rate value given the i th Monte Carlo sample. $I(x)$ constitutes an indicator function, which equals 1 if $x > 0$ and equals 0 if $x \leq 0$. Note that there are other more sophisticated approaches (e.g., importance sampling [23], subset sampling [24]) to compute P_f . However, we choose direct Monte Carlo simulation due to the fact that it is extremely easy to implement and versatile, therefore making it especially favorable for industrial

applications. In addition, by adopting straight-forward Monte Carlo method, we can better focus on quantifying the impact of GP model uncertainty on the variation of risk calculation, which is the main purpose of this work.

Here, due to the stochastic nature of the GP model, each $y_p(\mathbf{x}^i)$, $i = 1, \dots, n$ is a random variable, following a normal distribution with the mean and variance given by Eqs. (11) and (12). This is the marginal distribution of the ensemble of all the $y_p(\mathbf{x}^i)$'s, which follows a multivariate normal distribution with the mean and covariance given by Eqs. (9) and (10). In our previous work [10], we assume that the GP model uncertainty is negligible, i.e., we directly substitute $y_p(\mathbf{x}^i)$, $i = 1, \dots, n$ with $\mu(y_p(\mathbf{x}^i))$, $i = 1, \dots, n$ in Eq. (13), thus leading to a deterministic value of P_f . However, when $\text{cov}(Y_p)$ (determined by Eq. (10)) is sufficiently large, the training sample size is insufficient and GP model predictions exhibit nonnegligible uncertainties. In that case, $y_p(\mathbf{x}^i)$'s have to be treated as a random vector. As a result, P_f would not be a deterministic value, but rather a random variable with an associated PDF. In Sec. 4.2, we adopt an efficient procedure proposed by Ref. [13] to account for the randomness of $y_p(\mathbf{x}^i)$'s and using Monte Carlo technique to derive the PDF of P_f for the cavity mode.

4.2 Quantification of Gaussian Process Model Uncertainty. To determine the uncertainty of P_f , we employ a direct Monte Carlo approach: First, we generate realizations of GP predictions Y_p as Monte Carlo samples. Then, for each sample of Y_p , we calculate its corresponding P_f value via Eq. (13). Finally, based on an ensemble of P_f values, we can obtain its histogram, thus quantifying the uncertainty of P_f induced by the epistemic uncertainty of the GP model.

However, the Monte Carlo sample size n_{MC} in Eq. (13) is usually very large (at the order of 10^4). As a result, the covariance matrix of Y_p would be at the order of 10^4 -by- 10^4 . Since drawing samples from a multivariate normal distribution usually involves performing eigenvalue decomposition on the covariance matrix [25], the resulting computational cost of drawing samples of Y_p would be prohibitively expensive.

To resolve this efficiency issue, Nannapaneni et al. [13] proposed a procedure to reduce the problem dimensionality, i.e., the number of random variables in Y_p that has to be taken into account. The fundamental idea is illustrated in Fig. 3: since the output of the indicator function $I(x)$ only depends on the sign of its input, not all the variations of $y_p(\mathbf{x}^i)$ will lead to a change of P_f . Figure 3 illustrates two potential scenarios: For the GP prediction of the first Monte Carlo sample $y_p(\mathbf{x}^1)$, its mean and variance values can be calculated via Eqs. (11) and (12). Since more than 99.7% (three sigma) of the associated PDF lies in the unstable side, our GP model is extremely confident to classify this sample

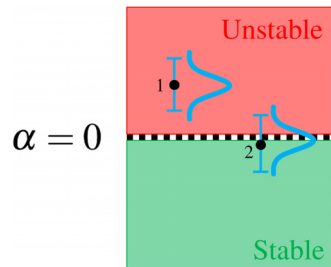


Fig. 3 An illustration of the fundamental idea to partition all the Monte Carlo samples into two groups. Two samples are depicted here, and their associated GP prediction PDFs are also plotted. Sample 1 belongs to the first group since its prediction variations would not change P_f . While for sample 2, it can hop between stable and unstable by chance, thus directly affecting P_f calculation. Therefore, we only need to consider samples in group 2 as random variables, thus significantly improving the UQ efficiency.

to be unstable, i.e., $I(y_p(\mathbf{x}^1)) = 1$, although its exact growth rate value may not be estimated precisely. In another word, there is negligible probability that our GP model makes a mistake in estimating the value of $I(y_p(\mathbf{x}^1))$. On the other hand, for the GP prediction of the second Monte Carlo sample $y_p(\mathbf{x}^2)$, our GP model no longer exhibits a high level of confidence regarding the stability nature of the sample: there are large portions of its PDF that lie in both the unstable side and stable side, respectively. Consequently, there is a high probability that our GP model makes a mistake in estimating the value of $I(y_p(\mathbf{x}^2))$.

Based on the probability that our GP model makes a mistake in estimating the sign of $y_p(\mathbf{x}^i)$, we can partition the n_{MC} Monte Carlo samples into two groups: The first group corresponds to predictions $y_p(\mathbf{x}^j)$, $j = 1, \dots, n_1$, for which the probability of making mistakes in estimating the signs of $I(y_p(\mathbf{x}^j))$, $j = 1, \dots, n_1$ is sufficiently low (i.e., below 0.3%). For this group, $y_p(\mathbf{x}^j)$ are not treated as random variables and the mean predictions $\mu(y_p(\mathbf{x}^j))$, $j = 1, \dots, n_1$ are directly used in Eq. (13). The remaining samples $y_p(\mathbf{x}^j)$, $j = 1, \dots, n_2$ (with $n_1 + n_2 = n_{MC}$) belong to the second group. They are treated as random variables, following a multivariate normal distribution, with mean and covariance determined by Eqs. (9) and (10). Therefore, Eq. (13) can be rewritten as

$$P_f = \frac{1}{n_{MC}} \left(\sum_{j=1}^{n_1} I(\mu(y_p(\mathbf{x}^j))) + \sum_{j=1}^{n_2} I(y_p(\mathbf{x}^j)) \right) \quad (14)$$

To achieve the partition, we adopt the following classification function proposed by Echard et al. [12]:

$$U(\mathbf{x}^i) = \frac{|\mu(y_p(\mathbf{x}^i))|}{\sigma(y_p(\mathbf{x}^i))} \quad (15)$$

Here, $U(\mathbf{x}^i)$ is used to estimate the probability of making a mistake in estimating the sign of $y_p(\mathbf{x}^i)$ (see Appendix A). The first group consists of samples with a U -value larger than 3, and the rest of the samples fall into the second group.

The following procedure summarizes the procedure to derive the uncertainty of P_f calculation. We assume the GP model has already been trained:

- (1) Generate n_{MC} Monte Carlo samples according to the distributions of uncertain flame and acoustic parameters;
- (2) Use GP model to calculate the mean and variance values of each sample (Eqs. (11) and (12));
- (3) Calculate U -value for each sample (Eq. (15)). All the samples with a U -value larger than 3 fall into group 1, the rest of the samples fall into group 2. Define n_1 and n_2 as the total sample number in two groups, respectively. Re-index the samples in the first group as $g_1(1), \dots, g_1(n_1)$. Re-index the samples in the second group as $g_2(1), \dots, g_2(n_2)$.
- (4) Using Eqs. (9) and (10) to calculate the mean and covariance matrix of the random vector $[y_p(\mathbf{x}^{g_2(1)}), \dots, y_p(\mathbf{x}^{g_2(n_2)})]$. Generate L realizations of $[y_p(\mathbf{x}^{g_2(1)}), \dots, y_p(\mathbf{x}^{g_2(n_2)})]$. For each realization, using Eq. (14) to calculate the corresponding value of P_f^k , $k = 1, \dots, L$;
- (5) Construct the histogram of P_f and extract desired statistical indices.

4.3 Case Study. We train our GP model based on the full parameter space described in Table 1. To generate training samples, we adopt the Halton sampling technique [26], which is a low-discrepancy sampling method with excellent space-filling property. In this study, we label this sampling method as *passive sampling*, to indicate that this sampling technique merely tries to distribute the training samples as evenly as possible in the parameter space. This is in direct contrast with the *adaptive sampling*

technique, which we will discuss in Sec. 5, where the sampling method exploits the already learned “landscape” of the underlying function, and adaptively allocate samples to the region where prediction accuracy is not satisfactory.

In line with the problem setting in “Q1: Risk Analysis” in Ref. [10], we consider four uncertain parameters: τ_1 , σ_1 , τ_{s1} , and τ_{s2} , which follow independent uniform distribution with the parameter ranges indicated in Table 1. τ_c and $|R_{out}|$ are fixed at their nominal values. Here, we use 102 training samples (same as in Ref. [10]) to train a GP model and follow the procedure outlined in section “Quantification of GP model Uncertainty” to derive the PDF of P_f . Monte Carlo sample size is 20,000 for both n_{MC} and L . After partitioning the Monte Carlo samples, only 1630 out of 20,000 samples belong to the second group, resulting in significantly lowered computational cost associated with generating realizations.

The calculated PDF of P_f for the cavity mode is shown in Fig. 4. Also shown in Fig. 4 is the reference value for P_f , which is calculated by the network model (Fig. 1) using the same n_{MC} Monte Carlo sample. It can be seen that although the reference value is covered by the PDF, P_f prediction exhibits a high level of variation, ranging from 5.5% to 10%, indicating a nonnegligible epistemic GP model uncertainty induced by a lack of training samples. No GP model uncertainty was considered in our previous work [10]. The predicted P_f has a deterministic value of 7.6%, which approximately corresponds to the mean of the P_f PDF shown in Fig. 4.

It is worth mentioning that the calculation of the reference P_f value via direct Monte Carlo is entirely feasible with an acoustic network model, thus making it possible to benchmark our current approach to quantify the uncertainty of P_f calculation. For other more computational intensive acoustic models, like the ones characterized by the Helmholtz equation or the Linearized Navier–Stokes equation, direct Monte Carlo would no longer be an option, then a Gaussian Process surrogate model with quantified uncertainty becomes really valuable for a robust calculation of P_f .

It would also be interesting to see how the PDF of P_f varies when the number of training samples is gradually increased. Figure 5 demonstrates the PDFs of P_f when the number of training samples are 100, 200, and 400, respectively. It can be seen clearly that with more training samples, the mean value of P_f prediction moves closer to the reference P_f value, thus indicating an improvement of accuracy in P_f calculation; in addition, the PDF becomes more concentrated, indicating that the epistemic uncertainty of GP model reduces with more available information (training samples).

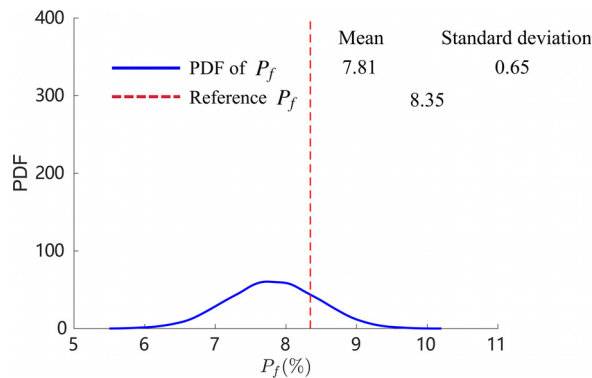


Fig. 4 The PDF of P_f is calculated following the procedure outlined in section “Quantification of GP Model Uncertainty.” Reference value of P_f is calculated by applying Monte Carlo directly on the acoustic network model (Fig. 1). Although reference value is covered by the PDF, P_f prediction exhibits a certain level of variation, which is induced by a lack of samples for GP model training.

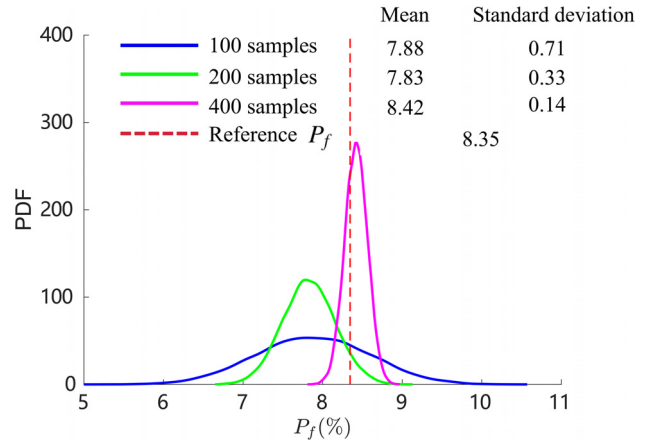


Fig. 5 More training samples lead to a more accurate and robust calculation of P_f

In Sec. 5, we will investigate how to intelligently train the GP model so that the improvement of both accuracy and robustness of P_f calculation can be maximized.

5 Reliable P_f Calculation

In this section, we aim to employ an adaptive sampling scheme to enrich the training samples X_D sequentially, thus enabling us to iteratively increase the accuracy and robustness of the GP model for P_f calculation, while keeping the number of runs of the acoustic solver as low as possible. We start with the motivation for adaptive sampling, followed by outlining the detailed procedures. Subsequently, we revisit the case study investigated in Sec. 4.3 and demonstrate the superior performance of the adaptive sampling scheme over the passive sampling scheme. It is worth mentioning that we have also applied our proposed workflow to quantify the uncertainty of P_f based on a Helmholtz solver. The results are presented in Appendix B.

5.1 Motivations. We can see that the convergence of the PDF of P_f is rather slow in Fig. 5. This is a direct consequence of the employed passive sampling strategy, which evenly distributes the training samples in the parameter space (shown in Fig. 6, left), with the goal of improving the accuracy and robustness of the GP predictions throughout the entire domain. This practice may induce significant wasted expense since as illustrated in Fig. 3, only the predictions of the Monte Carlo samples close to the zero growth rate contour are crucial to the accurate calculation of P_f . Those Monte Carlo samples are associated with low values of $U(x)$ (since their prediction means are small), indicating higher

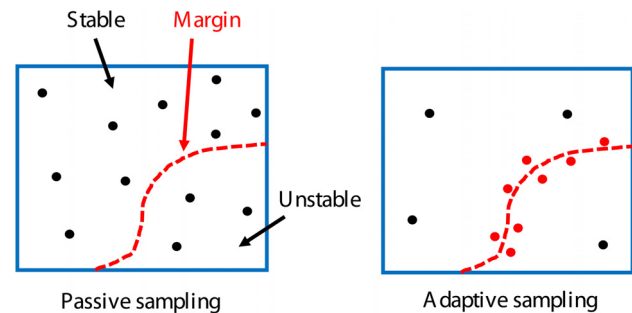


Fig. 6 Passive sampling scheme distributed training samples evenly across the parameter space while adaptive sampling scheme allocated samples in the vicinity of the zero growth rate contour. In this way, the calculation of P_f would be more accurate with less uncertainty.

probabilities of misclassification and contributing to the variations of P_f calculation. Therefore, instead of making the GP model accurate everywhere in the parameter space, it would be more beneficial to make the GP model particularly accurate in the vicinity of the stability border. To achieve that, more training samples should be allocated near the zero growth rate contour, as shown in Fig. 6.

5.2 Adaptive Sampling Scheme. In practice, we can adopt adaptive sampling strategy to intelligently allocate training samples: we start with an initial sampling plan and train an initial GP model. Then, a learning function is applied to determine the location of the next sample, which will be added to the current sampling plan. Afterward, an updated GP model is trained based on the enriched training samples. This iteration is terminated when a stopping criterion is met.

In order to put training samples in the vicinity of the zero growth rate contour, $U(\mathbf{x})$ in Eq. (15) is employed as the learning function, as proposed by Echard et al. [12]. Algorithm 1 summarizes the complete procedure for adaptive GP model training. For the stopping criterion, we require that $\min(U(\mathbf{x})) \geq 1.65$, which corresponds to a probability of misclassification of $\Phi(-1.65) = 0.05$.

Algorithm 1 GP model training via adaptive sampling [12]

-
- 1: Generate L samples X_U as a pool of candidate samples for sample enrichment, set iteration number $q = 0$
 - 2: Generate m initial samples $X_D = [\mathbf{x}^1, \dots, \mathbf{x}^m]^T$
 - 3: Call network model to evaluate their growth rate responses $Y_D = [\alpha^1, \dots, \alpha^m]^T$
 - 4: Construct the initial GP model $\hat{f} \leftarrow (X_D, Y_D)$
 - 5: Identify $\mathbf{x} \leftarrow \min_{\mathbf{x} \in X_U} U(\mathbf{x})$ based on \hat{f}
 - 6: **while** $U(\mathbf{x}) \leq 1.65$ **do**
 - 7: $q = q + 1$.
 - 8: Call network model to calculate the growth rate response σ of \mathbf{x}
 - 9: Enrich samples: $X_D = X_D \cup \mathbf{x}$, $Y_D = Y_D \cup \sigma$
 - 10: Retrain GP model $\hat{f} \leftarrow (X_D, Y_D)$
 - 11: Identify $\mathbf{x} \leftarrow \min_{\mathbf{x} \in X_U} U(\mathbf{x})$ based on \hat{f}
 - 12: **end while**
-

5.3 Case Study: A Revisit. We employ the Algorithm 1 to retrain the GP model. The initial sample size m is set to be 60, which is 10 times the number of input parameters, as suggested in Ref. [27]. The candidate sample size $L = 20,000$. The Latin-hypercube method is employed to generate samples.

Figure 7 displays the convergence history of the GP model training process. Here, the minimum U value among the candidate samples in each iteration is recorded and plotted. After allocating 193 training samples, i.e., 133 iterations, the minimum U value reaches 1.65; therefore, the training is deemed to be converged.

The case study investigated in Sec. 4.3 is revisited here, using the GP model trained via the adaptive sampling strategy. The calculated PDF of P_f for the cavity mode is demonstrated in Fig. 8. Also shown in the figure is the PDF of P_f calculated by the GP model, which is trained by using the passive sampling scheme with 193 training samples. We can clearly see that the adaptive training strategy significantly improved the accuracy and robustness of the P_f calculation, i.e., the PDF centers closer to the reference value with a smaller level of scattering.

To provide a more convincing visual demonstration of the effectiveness of the active learning approach, Fig. 9 displays the evolutions of the P_f mean prediction as well as the associated 95% confidence interval against the number of training samples. We can observe that the adaptive sampling scheme leads to a faster convergence of the mean prediction and a more dramatic reduction of the prediction uncertainty. As a matter of fact, the

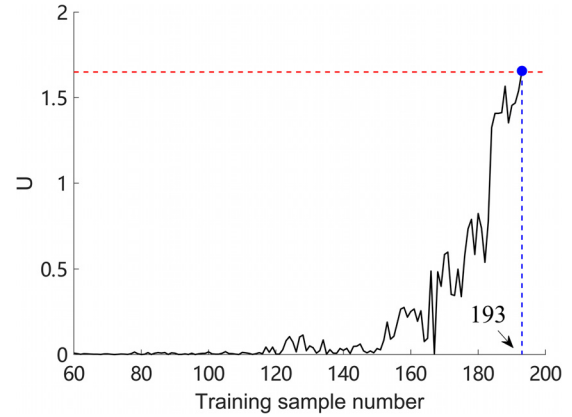


Fig. 7 The minimum value of U in each iteration is recorded and plotted here. After 193 training samples are employed, the minimum U value reaches 1.65, indicating a converged GP training process.

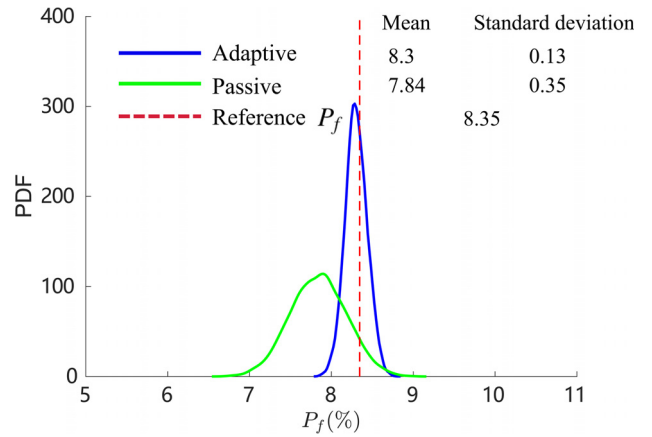


Fig. 8 Given the same number of training samples (i.e., 193), adaptive sampling scheme is able to make better use of the computational resources, yielding a more accurate and robust P_f calculation. The results delivered by the adaptive sampling scheme are comparable to the passive sampling scheme with double the computational cost.

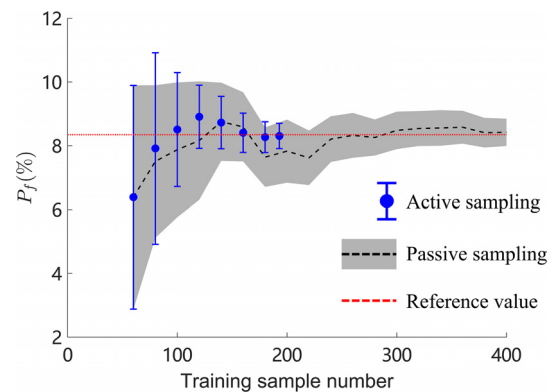


Fig. 9 The variations of the accuracy and robustness of P_f calculation with respect to the number of training samples. For both methods, 95% confidence interval of the P_f prediction are given, which serves as an indicator for prediction robustness.

calculation quality of P_f at the end of the adaptive sampling iteration (193 samples employed) is comparable to the results achieved by the passive sampling scheme with 400 samples. In addition, we emphasize that the GP model trained via the adaptive training

scheme is not only limited to the current case study. Other types of studies illustrated in Ref. [10] can also be investigated with the same GP model without further computational cost.

6 Conclusion

This work investigates the combined effects of aleatory parameter uncertainty and epistemic GP model uncertainty on the calculation of thermoacoustic instability risk (P_f). We addressed two practical questions in the study. For the first question, i.e., “how to quantify the variation of modal instability risk induced by the epistemic surrogate model uncertainty?”, this study leveraged on the prediction uncertainty naturally provided by the GP model and performed correlated sampling of the GP predictions at different inputs to quantify the uncertainty of P_f calculation. We demonstrated that the P_f calculation uncertainty shrinks as more training samples are available. For the second question, i.e., “how to reduce the variation of P_f calculation given a limited computational budget for the surrogate model training?”, this study adopted an adaptive sampling strategy to allocate training samples in the vicinity of the zero growth rate contour, due to the fact that the accuracy of GP predictions in that region plays a major role in achieving a reliable calculation of p_f . Our studies showed that the adaptive sampling strategy can significantly improve the accuracy and robustness of P_f calculation: With half the computational cost, the results delivered by the adaptive training strategy are comparable to the passive training strategy.

Note that in our previous work [10], we have extensively investigated various scenarios of robust design for thermoacoustic stability. However, the foundation of our strategies in Ref. [10], i.e., how to reliably calculate P_f , especially in face of limited training samples, is not addressed directly, but simply assumed feasible. Our current work addresses this shortcoming. Therefore, in combination with our previous work [10], this study is not only of academic interest but also of industrial relevance.

Future studies should focus on the following aspects: (1) when high-dimensional flame/acoustic models are employed, the efficiency of building a GP model may drop significantly due to the “curse of dimensionality.” Therefore, it is important to investigate how the proposed workflow can cope with high-dimensional. (2) When partitioning Monte Carlo samples into two groups to improve the efficiency of GP sampling, a significant number of “group 2” samples may still exist when high-dimensional flame/acoustic models are employed. (3) It is not uncommon that in thermoacoustic control, the effort to stabilize one mode may unexpectedly promote another mode to become unstable [28]. Therefore, it is necessary that our current adaptive strategy can be extended such that it can efficiently track multiple stability margins.

Funding Data

- Chinese Scholarship Council (Grant No. 201606830045; Funder ID: 10.13039/501100004543).

Nomenclature

- A = thermoacoustic modal growth rate
- FIR = flame impulse response
- GP = Gaussian process
- h_i = FIR model coefficient
- N = number of FIR model coefficients
- PDF = probability density function
- P_f = risk factor (%)
- $|R_{\text{out}}|$ = magnitude of reflection coefficients at combustor outlet
- UQ = uncertainty quantification
- Σ_1 = standard deviation of the distributed time lags of flame response for axial velocity perturbation
- τ_c = time lag for swirl fluctuation traveling from swirler to flame base

- τ_{s1} = mean of the distributed time lags of flame response for swirl fluctuation (+)
- τ_{s2} = mean of the distributed time lags of flame response for swirl fluctuation (−)
- τ_1 = mean of the distributed time lags of flame response for axial velocity perturbation

Appendix A: $U(x)$ and Probability of Misclassification

Misclassification at a Monte Carlo sample \mathbf{x} happens when its corresponding GP prediction mean $\mu(y_p) > 0$, whereas the true response $f(\mathbf{x}) < 0$ or vice versa. Since the GP prediction y_p follows a normal distribution with a mean $\mu(y_p)$ (Eq. (11)) and a variance $\sigma^2(y_p)$ (Eq. (12)), we can write out the probability of misclassification at \mathbf{x} as:

$$P_m(\mathbf{x}) = \Phi \left[-\frac{\mu(y_p)}{\sigma(y_p)} \right] \quad (\text{A1})$$

or

$$P_m(\mathbf{x}) = \Phi \left[\frac{\mu(y_p)}{\sigma(y_p)} \right] \quad (\text{A2})$$

where P_m stands for probability of misclassification and Φ represents the cumulative density function of the normal distribution. Equation (A1) corresponds to the case when $\mu(y_p) > 0$ while the true response $f(\mathbf{x}) < 0$ (as shown in Fig. 10(a)); Eq. (A2) corresponds to the case when $\mu(y_p) < 0$ while the true response $f(\mathbf{x}) > 0$ (as shown in Fig. 10(b)).

Combining Equations (A1) and (A2), we can derive the probability of misclassification as

$$P_m(\mathbf{x}) = \Phi \left[-\frac{|\mu(y_p)|}{\sigma(y_p)} \right] = \Phi[-U(\mathbf{x})] \quad (\text{A3})$$

Therefore, a larger $U(\mathbf{x})$ indicates a smaller probability of misclassification.

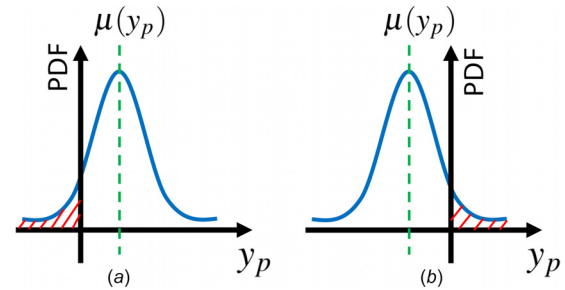


Fig. 10 Two situations for misclassification: (a) represents the situation that the truth is negative while GP predicts positive and (b) represents the situation that the truth is positive while GP predicts negative

Appendix B: Case Study With Helmholtz Solver

To further demonstrate the effectiveness of our proposed workflow to quantify and control GP model uncertainty, here we conduct another case study, where a finite volume-based Helmholtz solver [3] is adopted.

The configuration under investigation is the EM2C turbulent swirler combustor [29]. Here, we consider configuration C4 [30] and the operation condition B of Ref. [29]. An in-house finite volume-based Helmholtz solver [3] is adopted to calculate the first

Table 2 Input uncertain parameters

| Parameters | Range |
|------------|---------------|
| n | 0.4 – 2 |
| τ | 3.5 – 6.5(ms) |
| R_{out} | 0.5 – 1 |

longitudinal thermoacoustic mode of the test rig, and its variation under input uncertainties will be the focus of the current study. The governing equation is given as

$$\nabla(\bar{c}^2 \nabla \hat{p}) - s^2 \hat{p} + \alpha s \hat{p} = s(\gamma - 1) \hat{q}(x) \quad (B1)$$

where \hat{p} and \hat{q} denote the amplitude of pressure and heat release rate distributions, respectively. \bar{c} represents the mean value of the local speed of sound. A damping coefficient α is adopted to globally model the acoustic energy dissipation effect, which rests on the assumption that the acoustic system behaves like a second-order harmonic oscillator [30]. In this study, the damping rate of the combustor was evaluated to be $\alpha = 170s^{-1}$. Robin boundary conditions are employed

$$\nabla \hat{p} + \beta \hat{p} = 0, \quad \text{where} \quad \beta = s \frac{1(1-R)}{\bar{c}(1+R)} \quad (B2)$$

where R is the reflection coefficient at a given boundary; $R_{in} = |R_{in}|$ at the combustor inlet and $R_{out} = |R_{out}|e^{i\pi}$ at the combustor outlet. A $n - \tau$ model is adopted to describe the flame heat release rate response to velocity perturbation at the reference position, which can be written as

$$\frac{\hat{q}}{\bar{q}} = \frac{\hat{u}_{ref}}{\bar{u}} n \exp(i\omega\tau) \quad (B3)$$

In this case study, we consider the flame gain n , flame time delay τ , and the magnitude of reflection coefficient at the combustor outlet $|R_{out}|$ to be uncertain. Their variational ranges are displayed in Table 2. An independent uniform distribution is assigned to all three parameters.

We start with 30 training samples (10 times of the number of uncertain parameters) and employ the active learning scheme to adaptively train the GP model. Figure 11 displays the convergence history of the GP model training process. Here, the minimum U value among the candidate samples in each iteration is recorded and plotted. After allocating 53 samples, the minimum U value is above 1.65, indicating the training process is converged.

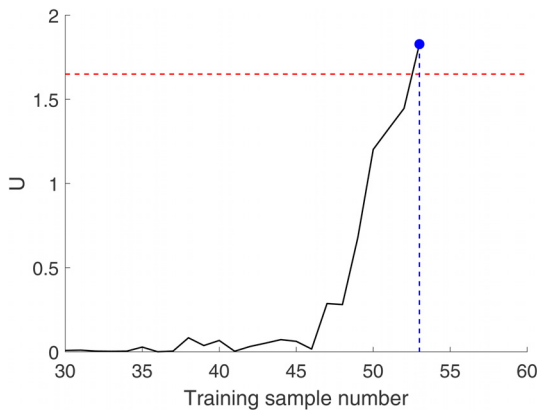


Fig. 11 The minimum value of U in each iteration is recorded and plotted here. After 53 training samples are employed, the minimum U value reaches 1.65, indicating a converged GP training process.

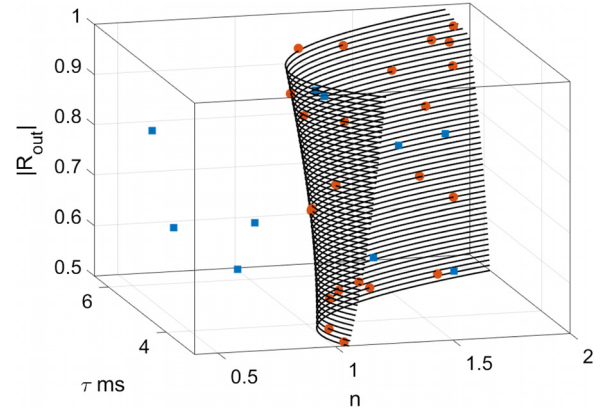


Fig. 12 The location of training samples in the 3D parameter space. The initial training samples are shown in cubics (only 10 of them are shown for clear illustration). The enriched samples are shown in circles.

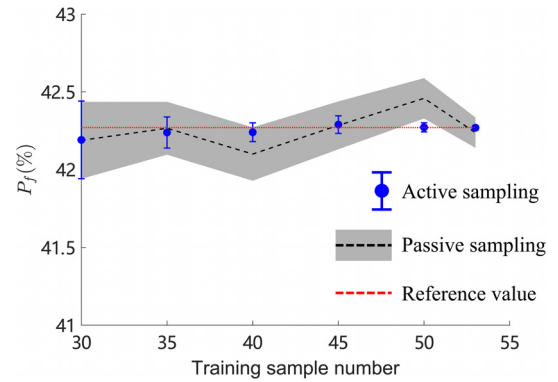


Fig. 13 The variations of the accuracy and robustness of P_f calculation with respect to the number of training samples. For both methods, 95% confidence interval of the P_f prediction are given, which serves as an indicator for prediction robustness.

Figure 12 shows the position of the training samples in the 3D parameter space, where blue cubics represent the initial samples (only 10 of them are shown for clear illustration), and red circles represent enriched samples. It can be seen that all the enriched samples are located in the vicinity of the stability margin, thus indicating our algorithm is behaving as we expected.

Compared with the passive sampling scheme, the active sampling approach significantly improved the accuracy and robustness of the P_f calculation. This point is illustrated in Fig. 13, where the predicted mean value of P_f as well as its uncertainty (95% confidence interval) is plotted against the available sample numbers. For the active sampling scheme, first, the predicted mean value of P_f converges much faster toward the reference value. Second, the prediction uncertainty of P_f reduces monotonically and with a faster speed. Finally, the provided confidence interval always covers the reference P_f value. Therefore, we are able to show that for a more complicated thermoacoustic model (i.e., Helmholtz-equation-based model), active sampling scheme remains to be effective and delivers a more accurate and less uncertain calculation of P_f , compared with the passive sampling scheme with the same computational budget.

References

- [1] Juniper, M. P., and Sujith, R. I., 2018, "Sensitivity and Nonlinearity of Thermoacoustic Oscillations," *Annu. Rev. Fluid Mech.*, **50**(1), pp. 661–689.
- [2] Guo, S., Silva, C. F., Ghani, A., and Polifke, W., 2019, "Quantification and Propagation of Uncertainties in Identification of Flame Impulse Response for Thermoacoustic Stability Analysis," *ASME J. Eng. Gas Turbines Power*, **141**(2), p. 021032.

- [3] Silva, C., Magri, L., Runte, T., and Polifke, W., 2017, "Uncertainty Quantification of Growth Rates of Thermoacoustic Instability by an Adjoint Helmholtz Solver," *ASME J. Eng. Gas Turbines Power*, **139**(1), p. 011901.
- [4] Ndiaye, A., Bauerheim, M., and Nicoud, F., 2015, "Uncertainty Quantification of Thermoacoustic Instabilities on a Swirled Stabilized Combustor," *ASME Paper No. GT2015-44133*.
- [5] Magri, L., Bauerheim, M., Nicoud, F., and Juniper, M. P., 2016, "Stability Analysis of Thermo-Acoustic Nonlinear Eigenproblems in Annular Combustors. Part II—Uncertainty Quantification," *Comput. Phys.*, **325**, pp. 411–421.
- [6] Mensah, G. A., Magri, L., and Moeck, J. P., 2017, "Methods for the Calculation of Thermoacoustic Stability Margins and Monte Carlo-Free Uncertainty Quantification," *ASME Paper No. GT2017-64829*.
- [7] Avdonin, A., Jaensch, S., Silva, C. F., Česnovar, M., and Polifke, W., 2018, "Uncertainty Quantification and Sensitivity Analysis of Thermoacoustic Stability With Non-Intrusive Polynomial Chaos Expansion," *Combust. Flame*, **189**, pp. 300–310.
- [8] Avdonin, A., and Polifke, W., 2019, "Quantification of the Impact of Uncertainties in Operating Conditions on the Flame Transfer Function With Non-Intrusive Polynomial Chaos Expansion," *ASME J. Eng. Gas Turbines Power*, **141**(1), p. 011020.
- [9] Silva, C. F., Pettersson, P., Iaccarino, G., and Ihme, M., 2018, "Uncertainty Quantification of Combustion Noise by Generalized Polynomial Chaos and State-Space Models," *Combust. Flame*, pp. 113–130.
- [10] Guo, S., Silva, C. F., and Polifke, W., 2019, "Efficient Robust Design for Thermoacoustic Instability Analysis: A Gaussian Process Approach," *ASME J. Eng. Gas Turbines Power*, **142**(3), p. 031026.
- [11] Balesdent, M., Morio, J., and Brevault, L., 2016, "Rare Event Probability Estimation in the Presence of Epistemic Uncertainty on Input Probability Distribution Parameters," *Method. Comput. Appl. Probab.*, **18**(1), pp. 197–216.
- [12] Echard, B., Gayton, N., and Lemaire, M., 2011, "AK-MCS: An Active Learning Reliability Method Combining Kriging and Monte Carlo Simulation," *Struct. Saf.*, **33**(2), pp. 145–154.
- [13] Nannapaneni, S., Hu, Z., and Mahadevan, S., 2016, "Uncertainty Quantification in Reliability Estimation With Limit State Surrogates," *Struct. Multidiscip. Optim.*, **54**(6), pp. 1509–1526.
- [14] Komarek, T., and Polifke, W., 2010, "Impact of Swirl Fluctuations on the Flame Response of a Perfectly Premixed Swirl Burner," *ASME J. Eng. Gas Turbines Power*, **132**(6), p. 061503.
- [15] Tay-Wo-Chong, L., Bomberg, S., Ulhaq, A., Komarek, T., and Polifke, W., 2012, "Comparative Validation Study on Identification of Premixed Flame Transfer Function," *ASME J. Eng. Gas Turbines Power*, **134**(2), p. 021502.
- [16] Oberleithner, K., and Paschereit, C. O., 2016, "Modeling Flame Describing Functions Based on Hydrodynamic Linear Stability Analysis," *ASME Paper No. GT2016-57316*.
- [17] Albayrak, A., Juniper, M. P., and Polifke, W., 2019, "Propagation Speed of Inertial Waves in Cylindrical Swirling Flows," *J. Fluid Mech.*, **879**, pp. 85–120.
- [18] Jones, D. R., Schonlau, M., and Welch, W. J., 1998, "Efficient Global Optimization of Expensive Black-Box Functions," *J. Global Optim.*, **13**(4), pp. 455–492.
- [19] Schneider, E., Staudacher, S., Schuermans, B., Ye, H., and Meeuwissen, T., 2007, "Real-Time Modelling of the Thermoacoustic Dynamics of a Gas Turbine Using a Gaussian Process," *ASME Paper No. GT2007-27468*.
- [20] Chattopadhyay, P., Mondal, S., Bhattacharya, C., Mukhopadhyay, A., and Ray, A., 2017, "Dynamic Data-Driven Design of Lean Premixed Combustors for Thermoacoustically Stable Operations," *Mech. Des.*, **139**(11), p. 111419.
- [21] Chattopadhyay, P., Mondal, S., Ray, A., and Mukhopadhyay, A., 2019, "Dynamic Data-Driven Combustor Design for Mitigation of Thermoacoustic Instabilities," *ASME J. Dyn. Syst., Meas., Control*, **141**(1), p. 014501.
- [22] Albayrak, A., Steinbacher, T., Komarek, T., and Polifke, W., 2018, "Convective Scaling of Intrinsic Thermo-Acoustic Eigenfrequencies of a Premixed Swirl Combustor," *ASME J. Eng. Gas Turbines Power*, **140**(4), p. 041510.
- [23] Echard, B., Gayton, N., Lemaire, M., and Relun, N., 2013, "A Combined Importance Sampling and Kriging Reliability Method for Small Failure Probabilities With Time-Demanding Numerical Models," *Reliab. Eng. Syst. Saf.*, **111**, pp. 232–240.
- [24] Miao, F., and Ghosn, M., 2011, "Modified Subset Simulation Method for Reliability Analysis of Structural Systems," *Struct. Saf.*, **33**(4–5), pp. 251–260.
- [25] Smith, R., 2014, *Uncertainty Quantification: Theory, Implementation, and Applications*, Society for Industrial and Applied Mathematics, Philadelphia, PA.
- [26] Swiler, L., Slepoy, R., and Giunta, A., 2006, "Evaluation of Sampling Methods in Constructing Response Surface Approximations," *AIAA Paper No. 2006-1827*.
- [27] Loepky, J. L., Sacks, J., and Welch, W. J., 2009, "Choosing the Sample Size of a Computer Experiment: A Practical Guide," *Technometrics*, **51**(4), pp. 366–376.
- [28] Lieuwen, T., and Yang, V., eds., 2005, *Combustion Instabilities in Gas Turbine Engines: Operational Experience, Fundamental Mechanisms, and Modeling*, Vol. 210, AIAA, Reston, VA.
- [29] Palies, P., Durox, D., Schuller, T., and Candel, S., 2011, "Nonlinear Combustion Instability Analysis Based on the Flame Describing Function Applied to Turbulent Premixed Swirling Flames," *Combust. Flame*, **158**(10), pp. 1980–1991.
- [30] Silva, C. F., Nicoud, F., Schuller, T., Durox, D., and Candel, S., 2013, "Combining a Helmholtz Solver With the Flame Describing Function to Assess Combustion Instability in a Premixed Swirled Combustor," *Combust. Flame*, **160**(9), pp. 1743–1754.