Classification and Computation of Extreme Events in Turbulent Combustion

Malik Hassanaly[a,b], Venkat Raman[a,*]

[a] *Department of Aerospace Engineering, University of Michigan, Ann Arbor, MI 48109, USA*
[b] *Computational Science Center, National Renewable Energy Laboratory, Golden, CO 80401, USA*

Abstract

In the design of practical combustion systems, ensuring safety and reliability is an important requirement. For instance, reliably avoiding lean blowout, flame flashback or inlet unstart is critical for ensuring safe operation. Currently, the science of predicting such events is based on prior experience, limited modeling or diagnostic tools and purely statistical approaches. Even though computational and experimental tools for studying combustion devices have vastly advanced in the last three decades, the analysis of such failure events has not been pursued widely. While the use of data for model development and calibration is being widely accepted, the extension to failure events introduces numerous challenges. In particular, the focus here is on so-called data-poor problems, where the cost of generating data is extremely high and is not easily amenable to existing computational and experimental approaches. Data-poor problems are particularly relevant when related to extreme events (also called anomalous events) that can lead to catastrophic failure of the system. It is argued that transient events that describe such failure can have different causal mechanisms. To develop the scientific inference process, a classification of such problems is used to determine specific modeling paths as well as computational tools needed. Research opportunities in the emerging field of extreme event prediction are highlighted in order to identify critical and immediate needs.

Keywords: Data-poor problems, extreme events, rare events

Contents

*Corresponding author
Email addresses: malik.hassanaly@gmail.com (Malik Hassanaly), ramanvr@umich.edu (Venkat Raman)

1. Introduction: Scientific inference in the age of data

The deluge of data, broadly in society and specifically in engineered systems, has prompted new research directions which is altering our perspective on engineering and related sciences. In the field of turbulent combustion, such data can come from machines in operation, as well as experiments and high-performance computations designed to study basic processes. The use of such disparate data to create and validate models or obtain new theories is broadly classified as data sciences. While data sciences, which include the use of machine learning (ML) and artificial intelligence (AI), have both positive and negative effects on society, they originate from fields that are not based on physics. This includes language translation, online advertising, and image recognition, where arguably more sophisticated AI-based approaches have replaced statistical inference. The popularity of these tools has heralded a new push towards using similar ML techniques in physical sciences. For instance, there has been a rapid increase in the use of neural networks to model specific fluid mechanics-related physics (discussed in [1, 2]). However, to see the long-term impact of data sciences, it is essential to have a more in-depth look at what new information can be obtained from the data-driven process compared to the conventional scientific process. More specifically, it is important to recognize where data can be useful, and what types of problems are appropriate for such data-based modeling.

In combustion science, the use of data for inferring models started with assimilating experimental measurements to calibrate kinetics models [3–6]. It has since been used for model calibration [7] and turbulent combustion model inference [8]. More broadly, such data assimilation has also been used for detecting sensor faults in gas turbines [9], devising control loops for internal combustion engines [10], and for building digital twins of physical devices [11]. In all these cases, data is assumed to exist and its use is limited to the calibration of models that can predict some physical process or device performance over a range of conditions. In other words, much of this development is in the context of interpolation. The novel aspect of data-driven models is that because large amounts of data are available, models can be directly inferred from data rather than merely calibrating model parameters. Of course, the true value of prediction is in extrapolation, at conditions for which prior data does not exist. Currently, such an extrapolation is fundamentally not feasible.

In this context, Raman and Hassanaly [2] defined three categories of problems in combustion science based on the availability of data: A) data-rich problems where the deluge of data requires special processing tools to identify critical features, B) data-sufficient problems which follow the traditional scientific inference process and produce predictive models that may be used with high-performance computational tools, and C) data-poor problems that are yet to be tackled by conventional or data-driven inference processes.

Category A describes the problems that use machine-learning tools currently, focusing more on identifying features within vast amounts of data that are being generated by operational devices such as gas turbines or internal combustion engines. A review of such tools is provided in [2]. For this class of problems, the primary use is in predictive analytics. In other words, repeated decisions that need to be made for a large number of systems (fleets of aircraft, for instance) are readily handled by machine learning tools. Alternatively, such tools are useful for the control of complex systems, where predictive tools can be used for actuation purposes. In both cases, detailed physics-based models are not tractable due to the time available for decision making, and fast-executable tools based on classification or regression are needed.

Category B problems focus on the design of new combustion devices, for which a hierarchical validation process is necessary to establish the credibility of models. For instance, the flamelet concept has been used in a variety of configurations and was found to yield remarkably accurate predictions [12]. The modeling activity organized around target flames [13–17] for which models have been gradually refined and extended to more and more extreme computational scales (by increasing the number of degrees of freedom). These two trends (model refinement and increase in computational scale) are illustrated in Fig. 1. It shows the evolution of prediction error for a target partially premixed turbulent jet flame, the Sandia D flame [13], which has been commonly used to test new combustion models and techniques. All the results shown were obtained using large eddy simulation (LES) data published between 2005 and 2018. The values of the quantity of interest (QoI) chosen to compare different simulations are shown close to the axial location where partial flame extinction occurs, at $x/D = 15$, where x is the axial location, and D is the jet diameter. The error between the predicted CO mass fraction and the experiments are plotted at different radial locations against the number of degrees of freedom used in the simulation.

Such validation tests are central to the category B problems, which allows new designs to be simulated by building trust in the models. The validated models are now integral to the design process and are used alongside experiments to advance new propulsion concepts. However, the data shown in Fig. 1 also paints a different picture. Although computational models have become increasingly complex, and the computational expense has gone up, there is not a significant convergence in the accuracy of the methods. This saturation in the modeling accuracy has led to the use of more simple models (RANS or simple combustion models) instrumented with tools for uncertainty quantification that aids in the decision-making process [18]. If the complexity of high-fidelity computational models over simpler models is deemed not worthy for practical applications, it raises doubt about the current direction of modeling regarding category B problems. While this question requires exploration by itself, this topic is beyond the scope of the current discussion.

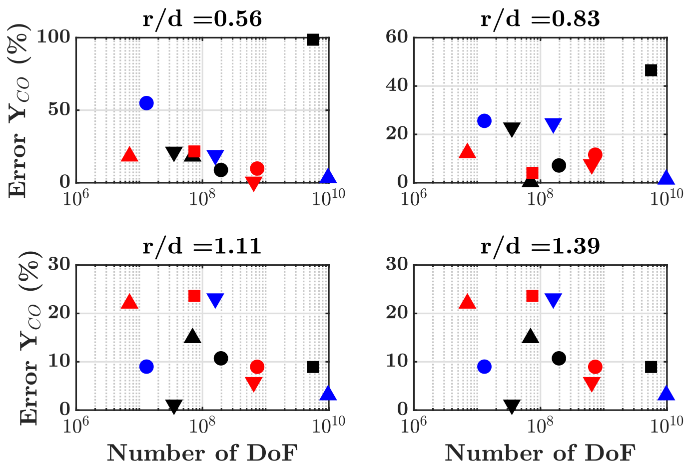

Figure 1: Experimental and numerical mismatch in percentage of CO mass fraction at $x/D = 15$ for different radial locations. The mismatch is plotted against the number of degrees of freedom (DoF) used in the following computations: ● [19]; ▲ [20, 35K particles]; ▼ [21]; ■ [22]; ● [23]; ▲ [24]; ▼ [25]; ● [26]; ▲ [27]; ▼ [28]; ■ [20, 175K particles].

Instead, the focus here is on category C problems. In the broader world of modeling and computational sciences, there has not been equal attention to category C problems, those with limited availability of data. Herein lies an opportunity for HPC, since the limited access to data makes any detailed computation extremely valuable. As the treatment of category B problems mature, there is an increasing need to focus on this newer class of problems, which are often not amenable to the same modeling, numerical and experimental treatment. The focus of this paper is to identify in what contexts data-poor problems are relevant, and to review existing computational tools than can form the basis of a new inference cycle for the combustion community.

This article is based on a preceding work [29]. While the topics discussed are similar, there are some fundamental differences between these two works. The current work provides a more comprehensive definition of data-poor problems, distinguishing them from extreme and rare events (which form a subset). Further, a discussion on the need for computational tools for such problems has been added. The central aspect of this work is the classification of events, which allows different computational approaches to be used. This classification of extreme events has been reformulated to represent the span of feasible causal mechanisms. Within each class of events, the use of advances from machine learning has been added as appropriate in order to reflect recent trends.

2. Data-poor problems, extreme events and rare events

The lack of data in any engineering problem is driven by two considerations: the perception of risk of failure of the system and the cost of generating the data relative to its value associated. For instance, if measuring the temperature at the exit of the combustor can improve performance but the cost of sensors in a high-temperature high-pressure environment is large, then such data collection may not be pursued. While this choice may only affect fuel efficiency, there are other design choices that can lead to calamitous outcomes, as evidenced by the 2019 Boeing 737 Max events [30]. In most instances, conservative choices are made such that the possible occurrence of such events is minimized. As discussed below, such choices may only decrease the perceived and not the actual odds of such events. Hence, the ability to unbiasedly estimate failure risk but at acceptable cost will be of immense value. As a quantitative example, consider that in 2015, 61 airline companies reported spending on average \$513M on fuel and oil [31, p.7]. Decreasing by 5% the fuel expenditure would save on average \$25M per airline per year. The goal of category B research is to advance such a fuel-efficient engine concept. Meanwhile, replacing a single new Airbus A320 that experienced a failure costs about \$100M [32]. The overall cost of a failure can actually far exceed the price of the lost device when one takes into account indirect costs (increase of insurance prices, loss of reputation). Fortunately, such failures are rare due to the conservative design choices. However, when these failures do happen, the system is complex enough that no two failures may have a common root cause. As a result, learning only from experience may not suitably reduce future risk of failure. It is necessary to have tools and techniques that can robustly quantify risks in any design. At a more fundamental level, simulation and design tools used to develop new vehicle concepts should be able to estimate the possibility of failure, preferably from a physically-descriptive modeling approach.

It is important to note that the notion of failure need not be limited only to catastrophic outcomes. More broadly, the interest is in understanding non-deterministic outcomes given a set of operating parameters and controllable inputs to the system. The discussion below casts category C as the study of such risk estimation.

2.1. Illustrative examples of extreme events

Extreme events (also referred to as anomalous in the rest of the manuscript) in engineering are marked by large excursions of a device from its design point. Although combustors are usually designed to be resilient to extreme conditions, turbulent combustion is not free of extreme events which can expose the device to catastrophic failures and expensive/unacceptable human and/or financial losses. As a starting point, consider these examples (shown in Fig. 2):

- One of the critical issues in scramjets is the stabilization of the shock structures in the pre-combustor region, termed as the isolator section. Since scramjets lack turbomachinery for compression, shocks are essential for maintaining flame stability in the combustor section. Under certain conditions, the shock structures inside the combustor can be ejected, leading to a total loss of compression, which is called engine unstart [33]. This event leads to catastrophic failure of the device.

- At lean operation of both aircraft and stationary gas turbines, there is an increased chance of flame blow-out [34]. If the fuel flow rate is reduced faster than the compressor response time, it can lead to conditions that cannot sustain a flame. Since flame blow-out can have disastrous consequences, reliable relight procedure is required for aircraft [35]. The relight procedure introduces a pocket of high energy fluid (a spark) into the engine while injecting liquid fuel [36, Chap 5.11]. Depending on the flow properties next to the spark (the level of turbulence and the fuel-air mixing), the engine can either reignite or remain extinguished. As a result, the ignition process is described probabilistically given a set of operating conditions for the engine.

- In order to limit carbon emissions in stationary gas turbines, hydrogen-enriched combustion could be considered. For example, syngas is used in integrated gasification combined cycle (IGCC). While it decreases pollutant emissions and facilitates carbon capture and storage [37], hydrogen use has a negative impact on the stability properties of the flame [38]. In particular, the fuel-air mixture can be so reactive that the flame may not remain in the combustion chamber but instead, propagate upstream [39]. In swirl combustors, for example, the flame can take advantage of the low near-wall velocity to creep upstream into the fuel-air mixing chamber. The premixing chamber is typically not designed to support the presence of high-temperature gases and can be damaged because of the flashback.

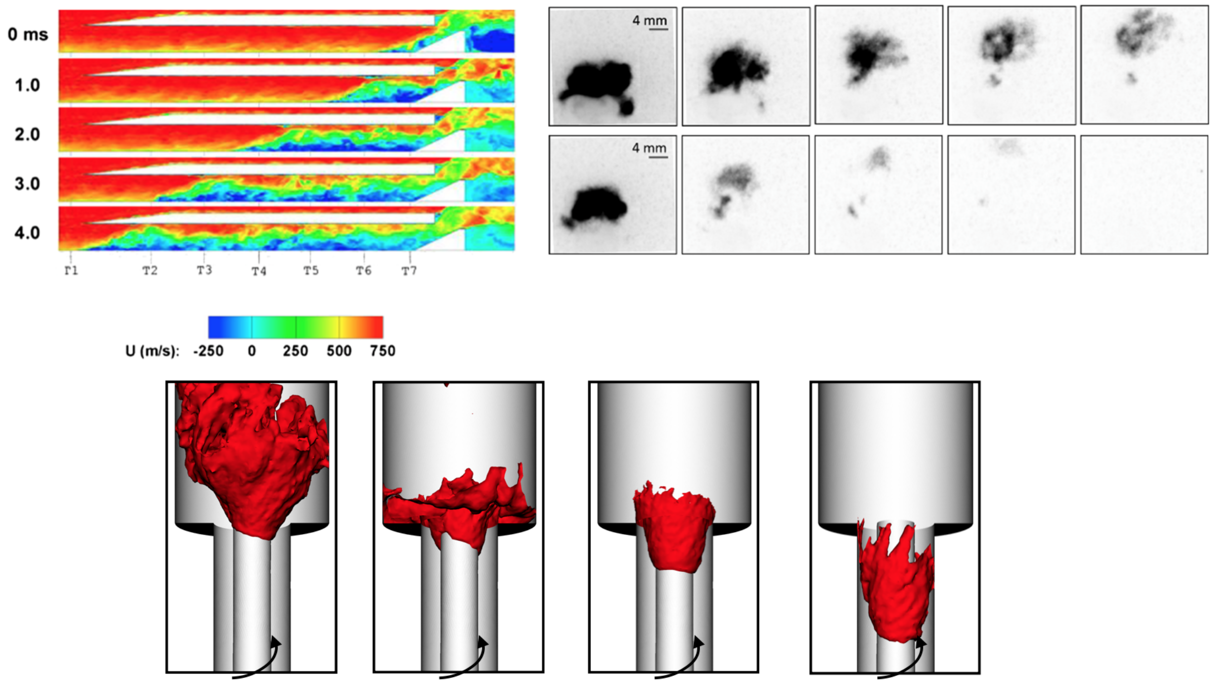

Figure 2: Top left: Streamwise velocity contours from LES of an isolator during unstart. Time advances from top to bottom. Reprinted from [40] with permission of American Institute of Aeronautics and Astronautics (AIAA). Top right: Sequence of line of sight measurement of OH^* during successful ignition (top) and ignition failure (bottom). Time advances from left to right. Reprinted from [41] with permission of American Society of Mechanical Engineers (ASME). Bottom: LES of boundary layer flashback in a premixed swirl combustor. The red iso-surface indicates the flame location. The arrows indicate the flow direction. Time advances from left to right.

Each of the above examples can result in a complete failure of the combustor. The most direct solution is to consider design choices that remove such events. However, this is not a fail-proof approach for two main reasons. First, turbulent systems are chaotic, which implies that certain events cannot be reproduced easily to determine the causal mechanism. Understanding an extreme event requires identifying what infinitesimal perturbations could lead to an extreme event, and be able to track its evolution. Second, combustor physics contain a large range of scales (typically from molecular scale to meters), and the devices themselves are coupled with a multitude of other components in the flow path. Therefore, a system can fail or reach an unwanted state in a myriad of ways. All such paths cannot be exhaustively accounted for during design. More critically, it may not be possible to thoroughly search all such possibilities, since certain events may be catalyzed by conditions that are present only during operation. For instance, the effect of operational cycles on the combustor or particular changes to fuel composition may not have been anticipated during the design phase. As a result, even a conservative design based on known failure modes cannot completely guarantee that all paths to failure have been identified.

2.2. Rare events and extreme events in relation to data-poor problems

To define the scope of the paper in relation to extreme and rare events, the definition of such events is provided in below. Since labeling problems as rare or extreme is a subjective choice, the boundaries between different classes of problems cannot be precisely defined. However, it is useful to identify what attributes of a problem make it suited for a data-poor type of analysis. Here, these attributes are the probability, the severity and the observational cost of the events.

Figure 3 schematically illustrates the type of problems that enter the scope of the paper. These problems are tied to extreme events and rare events, which are defined as follows. An event is characterized as extreme only based on its severity (Figure 3, left). However, not all extreme events are relevant in a design. In particular, an extreme event is relevant only if it occurs often enough to be a cause of concern in terms of an estimated impact, monetary or otherwise. For example, if an extreme event that causes cabin turbulence with a probability lower than once every ten flights, it may be considered less important for design purposes. However, if an extreme event leads to engine failure, it is relevant even if it occurs less than once every 10^3 flights. The relevance of an extreme event is a function of its probability and severity of impact.

In turn, rare events refer to low-probability events. There is no probability level that universally defines a rare event. Instead one can consider that an event is rare when the overall observational cost is so large that a specific procedure is required to induce this outcome. For example, if an event has low probability but can be analyzed with random observations, then it is not considered to be a rare event. On the other hand, if an event has a higher probability but the cost of each observation of the system is so high that random observations are not sufficient for the analysis, then these types of events are rare events. Here, these events are labelled as data-poor problems (Figure 3, right).

Since data-poor problems have a high observational cost, they are valuable if studied when they are associated with high severity, i.e., extreme events. However, the opposite is not true. While extreme events are rare through design choices, there is no necessity that such events always occur with low probability. For example, in the early stages of design, it is not uncommon to expect a high-probability extreme event. In fact, non-rare extreme events are useful to gain a fundamental understanding of the mechanism that leads to an extreme event (see Sec. 4.2.2). In practical applications, the relevant problems are at the intersection of the class of data-poor problems and extreme relevant problems.

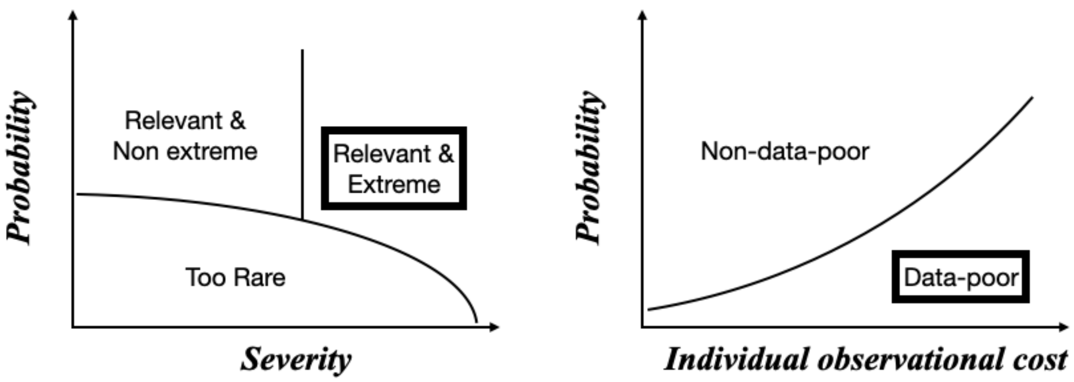

Figure 3: Illustration of the relation between rare and extreme events: (Left) extreme events depends on their severity. If the extreme event has too low probability, it may not be relevant. (Right) rare event are events that require a special treatment to be sampled either because they occur with low probability or because the observing the system is costly. Problems that are relevant, extreme, but also data-poor are the focus of this work.

2.3. The importance of computational tools

Traditionally, computational tools have been used to facilitate the observation of certain quantities of interest more rapidly than what experiments would allow, and also at a lower cost. In the context of extreme event analysis, computational tools may actually have no substitute. For example, if one is interested in the stability properties of a system, it is not possible to characterize the effect of perturbations using experiments only, since the state of the system cannot be controlled. Furthermore, if extreme events drive the system towards an irreversible state (by damaging the system), experimental investigations can only allow a few observations. As a result, very limited experiments exist for the problems mentioned above [33, 39, 42]. It is important to note that even when such experiments exist, only certain types of events can be probed. As will be discussed below, there are many types of failure events, and these can arise from different sources.

High-performance computing can play a key role in the analysis of data-poor problems. As opposed to data-sufficient problems where a lot of information is already available and low-order models have had time to mature, the physical mechanisms that drive extreme events are not well understood. Hence, accurate and detailed datasets can be invaluable for gaining a deeper understanding. Moreover, the questions posed by data-poor problems may not be suited for conventional low fidelity tools. For instance, consider the case of boundary layer flashback. Currently, available computational fluid dynamics (CFD) tools such as large eddy simulation (LES) and Reynolds-averaged Navier Stokes equations (RANS) can predict the average velocity of the flame front during a flashback. However, these methods will fail if the following question is posed: what is the fastest flashback velocity possible, or more formally, what is the distribution of flashback velocities possible? Even LES, which considers unsteady evolution of the turbulent flow, only represents a phase-averaged trajectory of the system [43, 44]. The validity of existing models depends on the predictive questions that pertain to the given extreme event.

Such events pose a formidable and relevant challenge not only in turbulent combustion but also in many other scientific communities, such as geophysics [45]. Over the past decade, significant advances have been made in probing these events, both from a mathematical perspective [46–52] and from applications perspective [53–58]. However, a comprehensive understanding of the nature of extreme events in complex nonlinear systems is still unavailable. With this background, the objectives of the paper are as follows: a) establish a classification of extreme events in turbulent combustion-related applications based on their causality (Sec. 3); b) survey the tools available to answer questions related to extreme events, and identify the limits of these tools for their application in the context of turbulent combustion (Sec. 4); c) highlight research opportunities in the emerging field of extreme events (Sec. 5).

3. Classification of extreme events

So far, extreme events have been related only to severity, which is not sufficient in dealing with the causal agents that drive such behavior. In order to understand the modeling issues, it is necessary to identify the different types of extreme events. The classification discussed below seeks to address this issue.

3.1. The need for a classification

Before embarking on the classification of extreme events, it is useful to recognize the need for this categorization. There are three broad benefits:

- Understanding the causality of extreme events helps determine their impact on practical devices. For instance, if the lack of precise knowledge of operating conditions leads to certain extreme events, one can then increase precision or design a more conservative device that can withstand the anomalies created by this uncertainty. Such classification or taxonomies are widely recognized as an approach to decompose complex problems. An interesting example can be found in Ref. [59] which recognized that to improve worker's error rates in factories, mistakes should not be only labeled as "human errors" but analyzed further to establish causality. This was achieved by introducing a classification of errors that precisely pointed out what needed to be improved.

- Causal classification allows identification of open questions and appropriate channeling of scientific efforts. As will be discussed in the following sections, there exist anomalies of various kinds, and their triggering mechanism can be grouped into a handful of categories where the analysis of different anomalies of the same group could be approached similarly. From a practical standpoint, this would encourage modelers and experimenters to gather data about certain anomalies.

- Provide a starting mechanism for the inference cycle [60]. In other words, extracting the events that have similar causal mechanisms can be used as the start of the abduction phase. This aspect is particularly interesting when machines can comb through data from practical devices and other disparate sources to identify patterns or features for such events. This autonomous search for causal mechanisms is an important component of the discussion below. In social sciences, the abundance of information of various kinds that cannot be treated efficiently by humans has encouraged the creation of algorithmic tools for decision making that are based on such classifications [61, 62]. In fact, many machine learning algorithms can be recast as classification algorithms that partition information based on the similarity of features.

3.2. Dynamical systems-based classification of extreme events

Prior reviews of extreme event mechanisms [45, 63] were motivated by geophysics applications. In these cases, extreme events (oceanic waves, heavy rains, landslides, etc.) occur spontaneously, and their existence is known. As will be explained below, some cases in turbulent combustion bear resemblance to these events, which allows capitalizing on advances achieved in these fields. However, other types of extreme events caused by different mechanisms also play a significant role in turbulent combustion. Here, extreme events are classified based on their causality with the goal of highlighting where methodologies applied in other scientific fields can be put to use, and where new methods need to be developed.

The classification is formulated by utilizing a dynamical systems approach for describing combustion systems. Mathematically, the state of the system is considered to be an infinite-dimensional vector (also called *state-vector*) $\boldsymbol{\xi} \in E$ (E being an infinite dimensional space called *phase space*) that evolves according to the evolution equation $\mathcal{F}$ given in its simplest form by:

$$\frac{d\boldsymbol{\xi}}{dt} = \mathcal{F}(\boldsymbol{\xi}, \boldsymbol{\zeta});\ \boldsymbol{\xi}(t=0) = \boldsymbol{\xi}^0, \tag{1}$$

where $\boldsymbol{\xi}^0$ is the set of initial conditions for the state-vector, and $\mathcal{F}$ is the operator that describes the time-evolution of the system. Typically, the operator $\mathcal{F}$ is obtained from physics principles. In the present work, the focus is on fluid systems, and the functional form of $\mathcal{F}$ may be derived from the conservation of momentum. Note that the operator $\mathcal{F}$ can depend on variables $\boldsymbol{\zeta}$ that are independent of $\boldsymbol{\xi}$. For example, $\boldsymbol{\zeta}$ can refer to boundary conditions or an external body force. From a practical point of view, $\boldsymbol{\zeta}$ also contains the operating conditions denoted by $\boldsymbol{I}$. In general, $\boldsymbol{I}$ is a set of macroscopic inputs, such as pressure, mass flow rate, or fuel-split in multi-injection combustors. With this notation, the time evolution of the dynamical system can be described by the *trajectory* of $\boldsymbol{\xi}$ in E.

In practice, a finite-dimensional discretized version of the governing equations is used using spatial discretization. The dimension of a dynamical system is the number of degrees of freedom. In turbulent combustion applications, the dimension can easily reach $O(10^9)$. For any given dynamical system, a quantity of interest, q, that characterizes the presence of an anomaly may be defined. While q is considered a scalar, a vector list of observables may also be used. q is defined as

$$\begin{aligned} Q\colon E &\to \mathbb{R} \\ \boldsymbol{\xi} &\mapsto q = Q(\boldsymbol{\xi}) \end{aligned}$$

There are two distinct assumptions made in this characterization. First, it is assumed that the dynamical system is deterministic, which implies that the future state of the system is determined only by $\boldsymbol{\xi}^0$ and $\mathcal{F}$. An important corollary is that trajectories in phase space never cross. For two trajectories that are distinct at time t_0 (i.e. $\boldsymbol{\xi}_1(t_0) \neq \boldsymbol{\xi}_2(t_0)$), if one can find two subsequent times t_1 and t_2 with $t_2 > t_1 > t_0$ such that $\boldsymbol{\xi}_1(t_1) = \boldsymbol{\xi}_2(t_2)$, then the trajectories of $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$ are the same, and are only delayed from one another. Second, it is assumed that the governing equations themselves are invariant with time. In Sec. 3.3, both these assumptions will be discussed and relaxed.

The dynamical behavior of the system is central to this work. Indeed, the deviations from "normal behavior" that the system experiences are the events of interest. The behavior of the system depends on the part of the phase space that it traverses. As a consequence, the system may exhibit different q based on the initial conditions. While turbulent systems never reach a steady-state, the dynamics of the system may be bounded to a subspace of the phase space [64–66]. The unsteadiness of the long-time behavior is mostly governed by the inertial forces while the boundedness of the subspace is due to the dissipative nature of the equations, along with the thermodynamical constraints. The system never spontaneously escapes this subspace which is called an *attracting set* [67]. The attracting set can be composed of different disjoint *attractors*. This means that for the same governing equations, multiple stable conditions can be observed. For instance, a scramjet can be stable in a "started" configuration or an "unstarted" configuration if the right perturbation is applied to it [40]. More trivially, both the ignited and extinguished states are thermodynamically reachable depending on whether an ignition source of sufficient strength is present. If several disjoint attractors are present, the path from a given initial point to a specific attractor depends on its basin of attraction, which contains all initial conditions that produce a trajectory ending on the same attractor.

Extreme events can then occur in different parts of the phase space. These events are triggered by many different causal mechanisms, including the presence of disjoint attractors. The classification below is based on these triggering mechanisms.

3.3. Causality classification of extreme events

The description of extreme events provided in Sec. 3.2 is general but does not describe how extreme events may manifest in practice. To better approach the problem of extreme events, a classification based on their causality is proposed. Three main types of extreme events are isolated and illustrated using turbulent combustion problems.

3.3.1. Type I: Extreme events associated with a controllable state

The Type I events represent the reliable behavior of complex devices. Here, for any given macroscopic operating condition $\mathcal{I}$, the output is precisely known. In other words, there is a direct connection between the input state and the output state. When such behavior is observed, it is possible to develop a map that relates input to output variables, either through experimental or computational procedures. Once this map is known, the device can be operated within boundaries such that unwanted output states are not observed. For most combustion systems, such an operating map is devised in order to approximate the stability limits.

Here, operating conditions $\mathcal{I}$ themselves could lead to unwanted regimes, where potentially catastrophic behavior is possible. Most studies of combustion instabilities and transient phenomena in combustion devices have focused on Type I events due to their easy reproducibility. In particular, Type I events can be readily studied using experiments, since the outcomes are directly dictated by the operating or boundary conditions at a macroscopic scale (such as pressure, inflow velocity, boundary layer thickness, etc.), and can be precisely controlled and/or measured. Type I events are also leveraged to collect data about a particular phenomenon with a limited number of observations. For example, the experimental work related to the problem of boundary layer flashback that was mentioned in Sec. 2.1 falls in the category of the Type I problems [39]. There, the flashback (the extreme event) is triggered with a probability equal to 1 by increasing the global equivalence ratio (one of the macroscopic inputs of the system). Other experimental studies of scramjet unstart fall in the same category [33]. There, an unstart (the extreme event) is triggered with a probability 1 by a sudden change in the flow outlet condition.

The dimension of $\mathcal{I}$ is denoted by $d_{\mathcal{I}}$, and is typically much smaller than that of the dynamical system. The fact that these inputs are sufficient to guarantee the output state shows that either a) there occurs a drastic reduction in the true dimensionality of the system, or b) the output variables are insensitive to much of the state-space of the dynamical system.

3.3.2. Type II: Extreme events associated with a non-controllable state

Type II events are related to the imprecise control of the state of the system using the limited set of input parameters $\mathcal{I}$ available. In other terms, for a fixed set of input parameters, the system can adopt many different states among which, some can lead to extreme events.

First, an anomalous behavior can stem from imprecise knowledge of the initial conditions of the system. Turbulence is a typical context in which this uncertainty would arise. Only macroscopic features of the flow are precisely controlled, while small scale turbulent fluctuations of the flow field are not. In initial conditions-driven anomalous events, a finite time horizon is considered. In other words, the short-time evolution of the system matters. More formally, let Q be some threshold for the observable, and T be some time threshold. During normal operations, $q > Q$ for $t = T$. During an anomaly, $q \leq Q$ for $t = T$. Such behavior is illustrated in Fig. 4. Most initial conditions lead to normal operating conditions, but at times, an extreme event can be encountered with a low probability (highlighted in blue in the figure).

A practical turbulent combustion example is the relight problem mentioned in Sec. 2.1. Here, the flame blows out at some operating condition, typically at high-altitude, and a relight procedure is initiated. Fuel is pumped into the combustor, and an igniter is used to send in high-enthalpy gases that can ignite and stabilize a flame. Both experimental [42] and simulation studies [68, 69] show that uncertainty in igniter output, turbulence state, and fuel-air mixing can lead to failed ignition events. Figure 2 (top right) shows the ignition outcome in a lab-scale experiment at fixed operating conditions. Variations in initial conditions lead to ignition success (top) or ignition failure (bottom). In other terms, the output cannot be solely controlled with the controllable input parameters.

Second, an anomalous behavior can stem from imprecise knowledge of the boundary conditions of the system. This aspect is particularly important for open systems with turbulent boundary conditions. Typically, the mass flow rate through a burner would be known, but the time-dependent turbulent structures entering the domain cannot be prescribed. This imprecision leaves room for extreme events to occur.

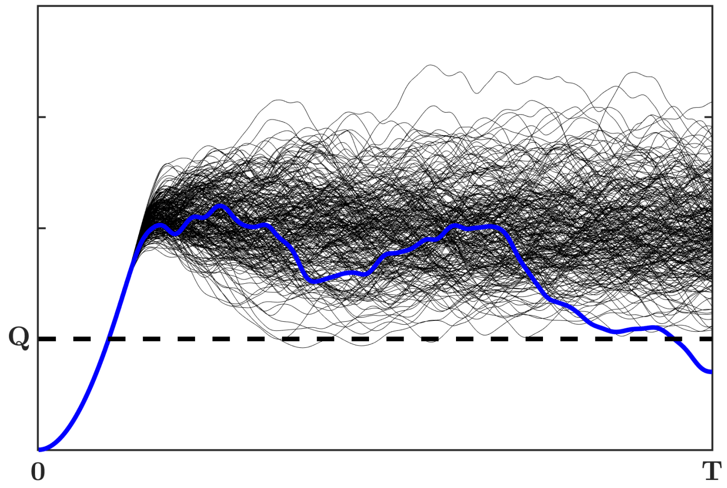

Figure 4: Illustration of the an anomaly driven by initial conditions. Most of the initial conditions lead to a quantity of interest in the normal operating range $q > Q$ for $t = T$ (—). Some initial conditions with low probability can lead to a rare and extreme event (—).

For instance, in swirl premixed burners, it was observed that at lean equivalence ratios, a flame could oscillate between two states: one attached to the nozzle and another detached from the nozzle. This process is illustrated in Fig. 5. There are several possible explanations, including aperiodic high strain rates that cause flame extinction near the base [70], or variations introduced near the boundary due to variations in turbulent inflow [71]. While the first cause would be classified below (spontaneous burst, event type III-A), the latter cause will lead to a Type II event. As a result, even for nominally identical operating conditions, different steady states as well as transitions between these states are feasible.

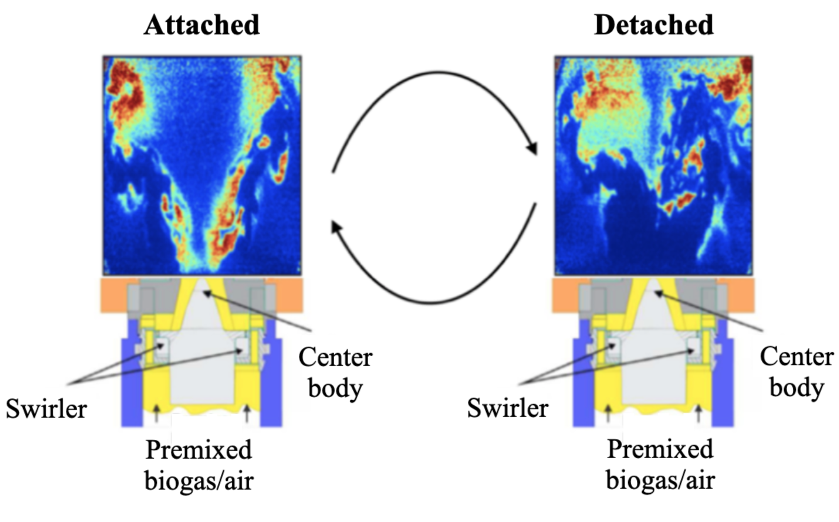

Figure 5: Illustration of the flame transition in a swirl premixed burner with OH PLIF contours taken at the combustor mid-plane. The OH contour denotes the flame location. Left: flame attached to the combustor nozzle. Right: flame detached from the combustor nozzle. Adapted from [72] with permission of Taylor & Francis Ltd (www.tandfonline.com).

3.3.3. Type III: Extreme events associated with the nature of the system dynamics

Type III events are related to the nature of the attracting set in which the system evolves. Pathological subspaces may exist as part of the attracting set, thereby allowing extreme events to occur. Compared to Type I events, while varying macroscopic conditions may change the shape of the attracting set and suppress potential extreme events, the design constraints do not allow such variations. Compared to Type II events, even in the case of a fully determined system (in terms of boundary conditions and initial conditions), extreme events may still occur. Different pathologies for the system dynamics can be identified and are discussed below.

Type III-A: Spontaneous bursts

Here, the QoI exhibits intermittent bursts that are periodically encountered, with a non-fixed period, and characterize the long-term behavior of the system [65]. In other terms, these extreme events are naturally encountered by the system as they are part of the underlying dynamics. This behavior is illustrated in Fig. 6 (left). Note that the frequency of such events can be extremely low, i.e., they can be low probability or rare events.

In the turbulent combustion parlance, such events are termed intermittent behavior. For instance, soot formation in gas turbine combustors can be highly intermittent. An illustration of this behavior is provided in Fig. 6 (right). The instantaneous contour of soot volume fraction in a swirl combustor at different times is shown and highlights this intermittent behavior (here the spatial and temporal localization of soot concentration). The production of soot can be considered an extreme event as the system is required to locally encounter specific conditions over extended periods of time [73–75]. Experiments of practical combustors showed that such behavior is central to the production and transport of soot [76].

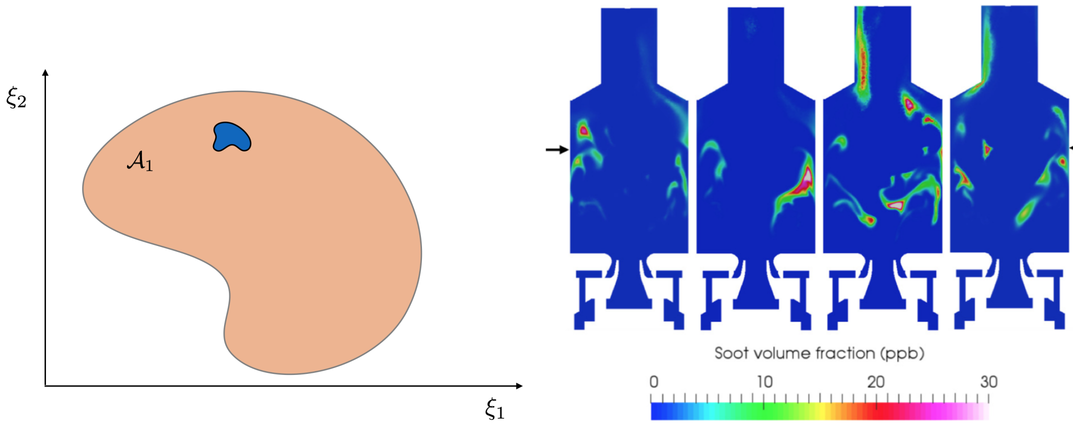

Figure 6: Left: Illustration of the attractor $\mathcal{A}$ of a dynamical system that exposes the system to spontaneous bursts. When a trajectory encounters the blue region, an extreme burst occurs. Because this region is part of the attractor, this burst occurs periodically during the life time of the system. Right: Instantaneous soot volume fraction snapshots at the center plane of a swirl combustor every at different times. Reprinted from [74] with permission of American Society of Mechanical Engineers (ASME).

Similarly, premixed combustors can exhibit macroscopic transitions that can be categorized as spontaneous bursts. At fuel-lean conditions, the flame front can, for example, oscillate between "V"- and "M"-shapes [77, 78] (Fig. 7). Such transitions can be detrimental to the efficiency and the durability of the combustor [79]. Since the flame topology controls the spatial distribution of heat flux to the walls, these transitions may expose walls to large heat loads. Numerical simulations conducted in non-swirling flames by Huang and Yang [80] suggested that this behavior is a spontaneous process that occurs due to the interaction between the flame and the walls of the combustor.

Outside the field of turbulent combustion, this type of events has received considerable attention. In oceanic engineering, the formation of rogue waves has been studied [54] and was shown to be a spontaneous event using an analogy with the non-linear Schrödinger equation (NLSE). Other examples of such spontaneous transition in turbulent flows are extensively discussed in Ref. [81].

Type III-B: Sensitivity to external shocks

Some configurations can reach multiple steady states at identical operating conditions, but the transition between these states cannot occur without an external force (shock) of appropriate magnitude and orientation in the phase space of the dynamical system. In a more general setting, external shocks can also refer to the modification of input parameters $\mathcal{I}$ of the system, which can be interpreted as a variation of initial conditions as illustrated in Fig. 8. One

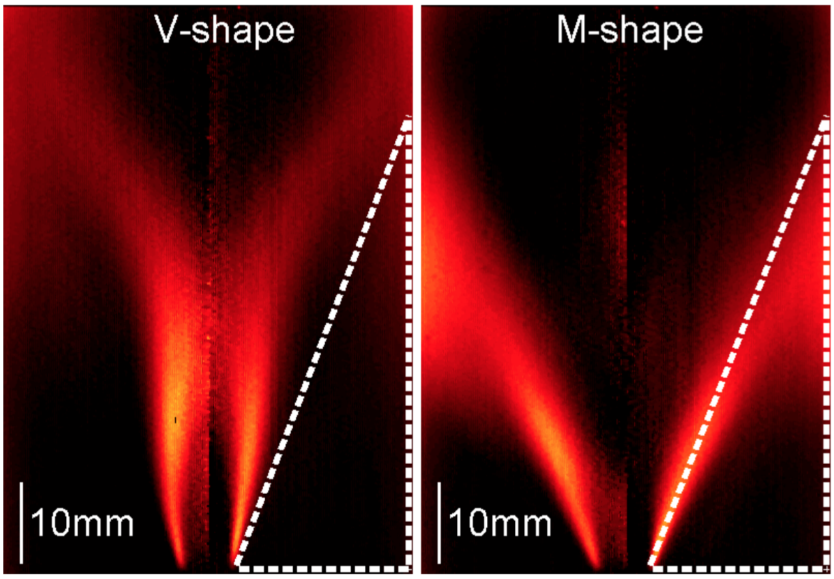

Figure 7: Variation of flame topology that can be encountered in a swirl combustor. The flow goes from the bottom to top. The flame can alternatively stabilize in a V-shape (left) or M-shape (right). Images are captured using flame chemiluminescence. Reprinted from [77] with permission of American Society of Mechanical Engineers (ASME).

way through which the impact of initial conditions can manifest is hysteresis. A common occurrence of this type in combustion is in the thermoacoustic response of combustors [82]. When pressure oscillations and heat release are in phase, an amplification of the fluctuations can occur, leading to catastrophic failure of the device [83, 84]. However, it has been shown that when the operating point of the system is slowly varied, the device can reach the same set of operating conditions without exhibiting thermoacoustic instabilities [85]. Figure 9 shows the RMS pressure fluctuations for the same combustor where the operating equivalence ratio is varied. It is seen that for identical global ϕ, two states are possible based on the trajectory taken.

This hysteresis phenomenon can be explained by the fact that for specific equivalence ratios, there exist two irreversible stable conditions. Depending on the initial conditions (the direction along which equivalence ratio is varied), the system stabilizes in either one of these states. In combustion problems, hysteresis is not only limited to swirl combustors but has also been observed numerically for droplet vaporization. Figure 9 (right) illustrates this effect with the map of the evaporation rate of a droplet as a function of the Reynolds number of a crossflow. It was found that two different vaporization rates can be observed depending on the direction along which the Reynolds number is varied [87]. The two steady states are obtained based on the history of the device traversing a basin of attraction in Fig. 8. Since the attractors are disjoint, the system will reach one of the attractors based on the initial conditions chosen.

Type III-C: Sensitivity to continuous perturbations

In describing any flow configuration, the extent of the domain and physics that is modeled is limited to ensure computational tractability. In this regard, even if the governing equations are fully known, the system is not completely described due to the impact of external factors. In such systems, the propagation functional $\mathcal{F}$ does not completely describe the state of the system. This implies that the state space of the system and the structure of the attracting regions are not precisely known or obtainable. Type III-C events occur due to the imprecise definition of the functional form of the governing equations $\mathcal{F}$. Such systems can be considered as stochastic processes (as opposed to deterministic processes), where a continuous stochastic (or random) forcing term represents the effects of the unmodeled physics.

For instance, Popov et al. [88] studied the onset of thermoacoustic instabilities due to variability in the acceleration of a rocket. Stable and unstable behaviors of this system are shown in Fig. 10, along with the spectrum of the rocket acceleration leading to this behavior. The system used there was not deterministic as the perturbations could be neither

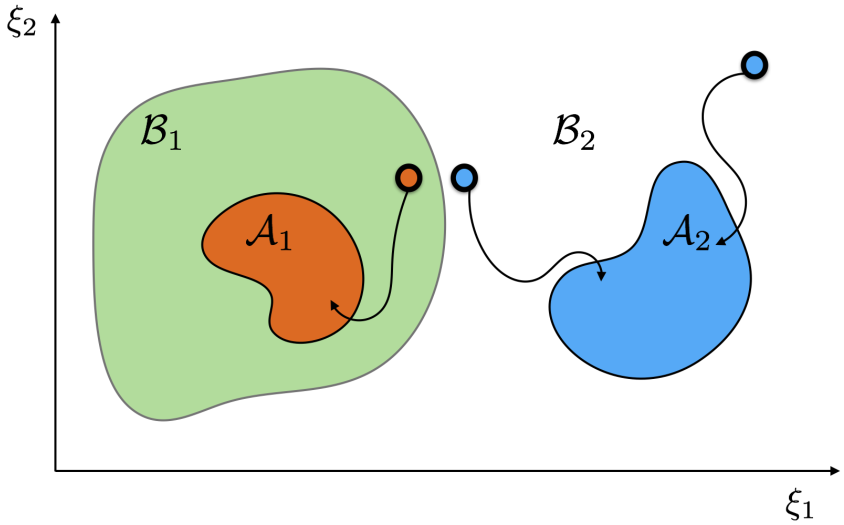

Figure 8: Illustration of the phase space in which a dynamical system evolves. The space is partitioned between two attractors $\mathcal{A}_1$ and $\mathcal{A}_2$ of basin of attraction $\mathcal{B}_1$ and $\mathcal{B}_2$. The arrows denote the time evolution of the dynamical system. Initial conditions are denoted by red dots when they are attracted by $\mathcal{A}_1$ and blue dots when attracted by $\mathcal{A}_2$.

reasonably assumed to be known *a priori*, nor are functions of the solution. In other studies unrelated to combustion applications, the effect of continuous external perturbations was also investigated for the stall of rotorcraft [89]. Similar to the thermoacoustic instability due to the acceleration of a rocket, the rotorcraft is exposed to atmospheric flow, which contains some level of variation. These variations act as external perturbations and can eventually lead to the stall of the rotorcraft (an extreme event). Modeling the physics including the external airflow would be intractable. Instead, the system's dynamics are assumed to be continuously affected by external perturbations.

As a final remark, it is noted that the system dynamics can evolve over time depending on wear or damages incurred during operations. As a result, the function $\mathcal{F}$ changes with continued use of the system, which can introduce dynamics that is not present in the original design. In turn, the long-term behavior of the device would change making it susceptible to Type III events.

Based on the above discussion on the different classes of events, a summary is provided in proposed in Tab. 1. In particular, the three main factors are the trigger mechanism, the structure of the attracting set in phase space, and the sensitivity of the output of interest to either macroscopic variables or the state of the system. In Tab. 1, $\mathcal{I}$ refers to the input macroscopic parameters of the system, Q is the quantity of interest, ξ is the state of the system, ζ are the boundary conditions of the system, δ notation refers to small perturbations, $\mathcal{A} = \{\mathcal{A}_i\}$ is the attracting set where $\mathcal{A}_i$ are possibly disjoint attractors, and $\mathcal{F}$ are the governing equations.

4. Computational tools for extreme events

The classification above showcases the different paths to observing anomalous events in systems. However, these paths alone are not indicative of the type of analysis required. In other words, different classes of anomalous events may be present at different stages of the scientific inference process. The questions answered by the analysis depend on the specific engineering needs. This section aims at clarifying what the word *prediction* entails in the context of extreme events. With this discussion in mind, the rest of the article describes relevant questions and the computational tools that can be used for data-poor problems.

4.1. Predicting anomalous events

Just as the definition of an extreme event is non-unique (Sec. 2.2), the notion of "prediction" also can be considered in different ways. In fact, the question that is posed will determine not only the algorithms used but also the level of expert input needed. Below, a nominal set of prediction targets for engineering applications are first discussed.

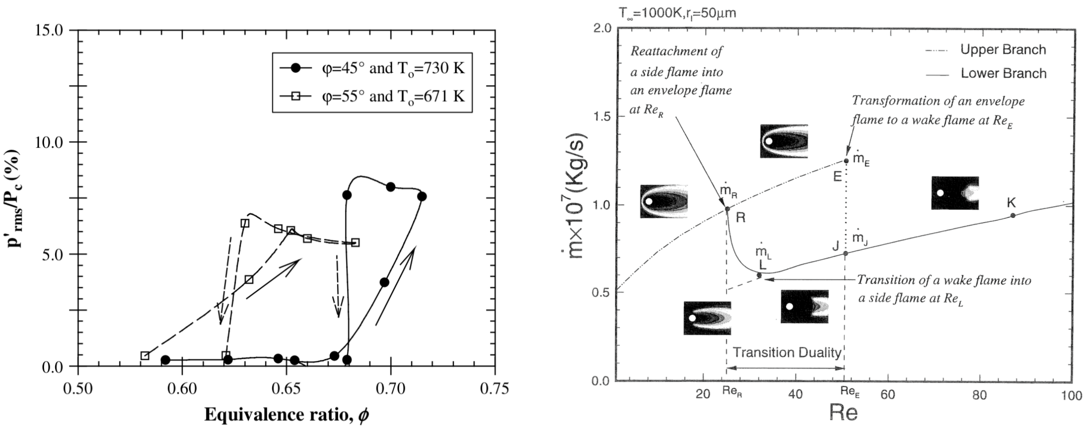

Figure 9: Left: Stability maps of a swirl combustor as a function of the equivalence ratio for two different swirl angles. Arrows indicate the chronological order of measurements. Reprinted from [86] with permission of the author. Right: Illustration of the multistability phenomenon for droplet vaporization. In the transition regime, two different vaporization rates are possible at fixed operating conditions. Reprinted from [87] with permission of Begell House.

Table 1: Classification of extreme events.

	Type I	Type II	Type III-A	Type III-B	Type III-C
Trigger	$\boldsymbol{\mathcal{I}}$	$\delta\boldsymbol{\xi}^0$ or $\delta\boldsymbol{\zeta}$	None	External shock	$\delta\mathcal{F}$
Attracting set	Non-relevant	Non-relevant	Existence of pathology	Disjoint attractors	$\lvert\frac{\delta\mathcal{A}}{\delta\mathcal{F}}\rvert \gg 1$
Output sensitivity	$\boldsymbol{Q}(\boldsymbol{\xi},\boldsymbol{\mathcal{I}}) \approx \boldsymbol{Q}(\boldsymbol{\mathcal{I}})$	$\lvert\frac{\delta Q}{\delta\xi^0}\rvert \gg 1$ or $\lvert\frac{\delta Q}{\delta\zeta}\rvert \gg 1$	$\exists \boldsymbol{\xi}_c \in \mathcal{A}$, $\lvert\frac{\delta Q}{\delta\xi}\rvert_{\xi_c} \gg 1$	$\lvert\frac{\Delta Q}{\Delta\mathcal{A}_i}\rvert \gg 1$	$\lvert\frac{\delta Q}{\delta\mathcal{A}}\rvert \gg 1$

The goal of computations may be defined as follows:

- Establish a stability map: When extreme events are fully controllable via the input parameters (Type I events), the goal is to find the set of parameters that leads to non-extreme conditions, while not hindering the performances of the device. In other terms, the goal is to establish a stability map for the combustor. The stability map is then used by the operator to use the device in safe, yet efficient conditions.
- Predict for real-time control: In practical systems, actuation can be used to push the system away from an anomalous event predicted sufficiently ahead of time. However, for computational tractability, this requires that precursors for such events be identified using a projection of the state-space. This problem is relevant for Type II and Type III events, where one may want to avoid an imminent unstable behavior.
- Estimate the probability of an event: For design purposes, it is useful to determine the probability of encountering a pre-determined anomalous event. It can help determine whether an extreme event is actually relevant for the design, including design considerations to improve the robustness or resilience of the device. For instance, if for safety reasons the ignition time for high altitude relight should not exceed T, then estimating the probability $P(T_{ig} > T)$, where T_{ig} is the time needed to ignite the engine, is a useful design metric. This problem is relevant for Type II and Type III events, where uncertainty about the initial or current state or the dynamics suggests treating QoIs as random variables with given probability density function.

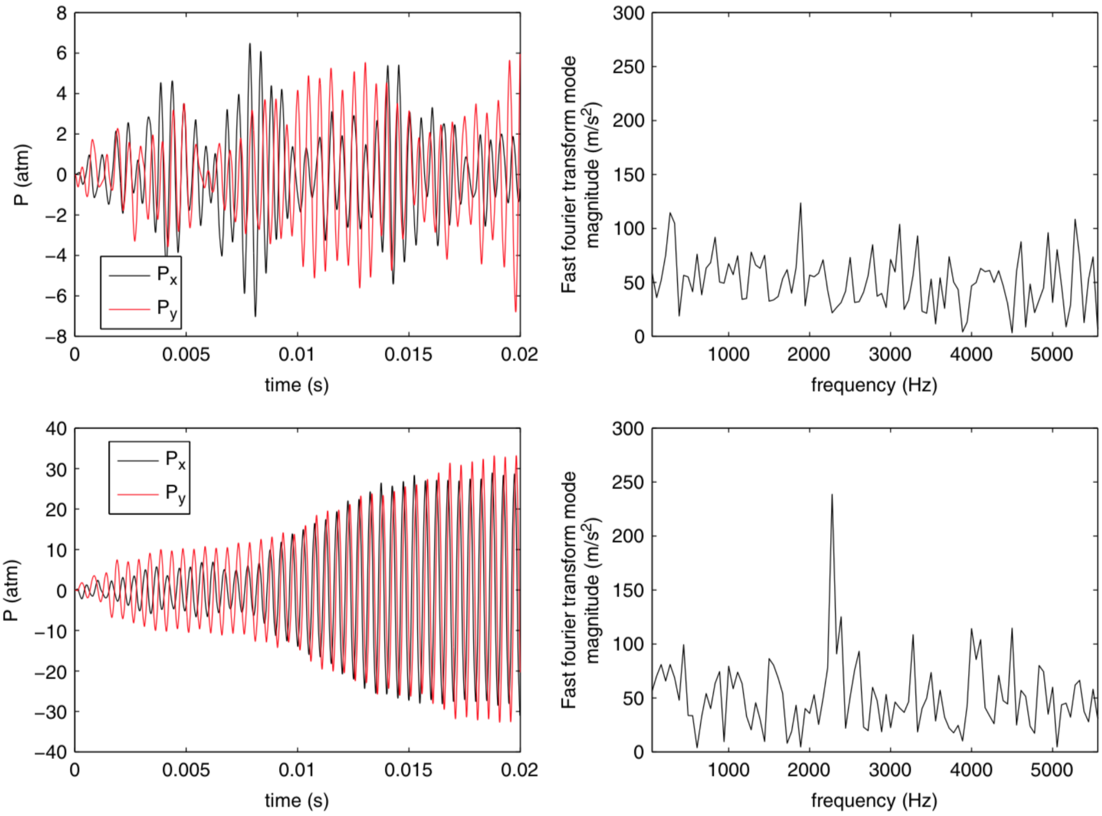

Figure 10: Pressure history of stable (top) and unstable (bottom) configuration in a model rocket engine exposed to different accelerations. The acceleration acts as an external perturbation that can drive the system to an unstable behavior (here, an acoustic instability). Reprinted from [88] with permission of American Institute of Aeronautics and Astonautics (AIAA).

- Predict unobserved events: A truly remarkable use of computational tools will be the ability to predict previously unobserved anomalous events. While the exploration of the state space is feasible in many engineering domains, the large dimensionality of a turbulent flow combined with a low probability of traversing certain regions makes it difficult to obtain such events through Monte Carlo type random searches. Since this problem suggests that the unobserved event is not known yet, it is not adherent to a particular type of event, but is applicable to all extreme events in general. However, the likely events that will be targeted are Type III, where one may want to identify unstable parts of an attracting set.

- Bound quantities of interest:
Safeguarding against failure is one of the primary design constraints. In this sense, being able to predict the worst outcomes possible, or more precisely, providing bounds on quantities of interest would be a valuable tool. Predicting bounds on quantities of interest is valuable for all events but might be particularly relevant for Type II and Type III events, where uncertainty about the initial or current state or the dynamics induces a range of possible values for the QoI and drives the extreme event.

- Revise a design:
If a given design gives rise to an extreme event, it could be advantageous to revise the design in order to make extreme events not occur anymore. For this purpose, it is necessary to understand the cause of the extreme event, i.e., analyze its trajectory in phase space. This prediction target is valuable for all the types of events.

4.2. Survey of computational tools

In this section, the available computational tools that could be used to tackle the target questions listed above are surveyed. Importantly, the relevance of these computational tools for turbulent combustion problems is evaluated.

This point is crucial as many of the available tools are often tested only on low-dimensional canonical problems, and are not tailored for combustion-related physics. Hence, applicability to the set of equations that govern reacting flow, as well as scalability to high-dimensional problems that result from a numerical discretization of these governing equations should be considered. The different approaches are listed in a table at the beginning of each subsection, along with requirements and advantages of the methods, and detailed discussion is provided underneath.

4.2.1. Establish a stability map

Table 2: Overview of computational methods that establish stability maps.

Establish a stability map			
Method	Ref.	Advantages	Requirements
Edge marching	[90]	Efficient in high-dimensions	Observations of instabilities and smoothness of the map
Edge parameterization	[91]	Efficient with low number of samples	
Polynomial annihilation	[92, 93]	Limited geometric specification	
Support vector machine (SVM)	[94]	Arbitrary edge topology, very few points needed	

Stability maps are of interest for Type I events, where the outcome is fully controllable by the set of input parameters $\boldsymbol{I}$. Here, the main goal is to precisely find the edge between safe and unsafe operating conditions in the $d_{\boldsymbol{I}}$ parameter-space, where $d_{\boldsymbol{I}}$ is the dimension of $\boldsymbol{I}$ which can be expected to be $O(10)$. An overview of methods to construct stability maps is provided in Tab. 2.

Not all combinations of input parameters can be tested due to observational cost, and the stability map must be constructed with the smallest possible set of data points. This is still a case of a data-poor problem. Determining the envelope of stable operating conditions can be recast as edge detection, which has been popularized by the field of image recognition [95], and has since been extended to engineering applications. Techniques such as edge marching [90] have demonstrated a linear dependence of the number of points required with the dimension of space. In the same vein, parameterization techniques for the geometry of the edge were demonstrated and showed how a balance between cost and accuracy could be achieved [91]. Often, these methods require an indicator function for the discontinuity. In particular, the polynomial annihilation indicator, which is null everywhere except near a discontinuity, has received significant attention in the past [93] and it has been demonstrated for up to 6 dimensions [92]. Building upon the polynomial annihilation method, an additional technique was recently introduced, and uses support vector machine (SVM) to approximate discontinuities [94]. It has the advantage of invoking a limited set of assumptions about the structure of the edge. In particular, disconnected unsafe regions of the parameter space (see Fig. 11) can be identified, and it requires two orders of magnitude less function evaluations than Ref. [90, 93] to identify a discontinuity. In terms of scaling, the number of training samples grows exponentially with the number of dimensions, but sub-exponential scaling can be achieved at the expense of minimal accuracy loss. Finally, we note that data-driven methods that identify boundaries between stable and unstable modes may face numerical issues in the case where there exist more stable than unstable data points. This issue is known as "class imbalance" and can be addressed via data augmentation or data selection [96–98].

The computational efficiencies of the methods mentioned in this section are all dependent on the number of dimensions of the input parameter space $d_{\boldsymbol{I}}$. In order to decrease the number of experiments required to find the edge between safe and unsafe conditions, it can be beneficial to reduce the number of relevant dimensions. Methods such as active subspace [99–102] and global sensitivity [103–105] analysis can be useful in that regard.

For combustion applications, stability maps could be useful for predictable events that can be triggered with macroscopic quantities (Type I events). For example, if a boundary layer flame flashback [39, 106] is triggered by the global equivalence ratios, mean temperature or pressure, one can establish a stability map to identify "safe" and

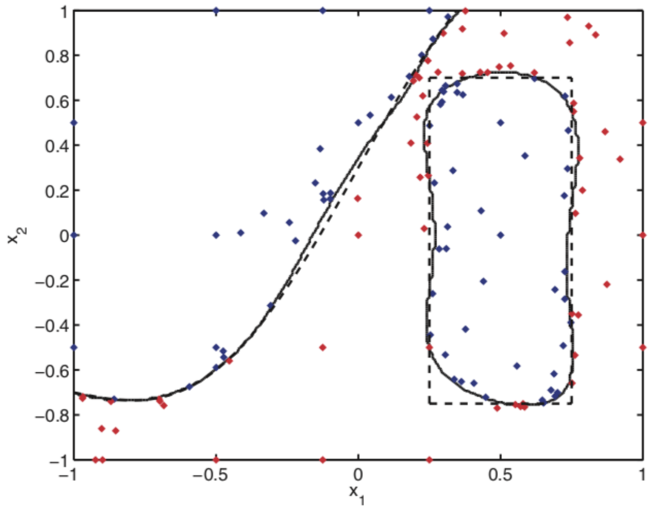

Figure 11: Estimation of the discontinuity in a two dimensional space between two different labelled output (blue dot and red dots). Dashed line is the exact edge. Solid line is the approximated edge after 120 iterations. Adapted from [94] with permission of SIAM. Copyright 2014 Society for Industrial and Applied Mathematics. All rights reserved.

"unsafe" modes. Acquiring the "unsafe" data could be done either via simulations or in representative lab-scale combustors.

4.2.2. Predict the occurrence of an extreme event

Real-time or close to real-time predictions of extreme events are particularly useful when an alternative course of action can be used to avoid catastrophic consequences. In the context of the classification, such predictions are most useful for Type II, Type III-A, and Type III-C events, which are endemic to the system. Since fast predictions are needed, direct solution of the dynamical system (Eq. 1) is not useful, but some reduced form that still captures the mechanism for the generation of anomalous events is preferable. An overview of methods to predict that an extreme event will occur is provided in Tab. 3.

However, long-time predictions in chaotic systems are particularly challenging since even small initial variations can lead to large changes in outcomes (butterfly effect). This problem is central to weather forecasting [114–116], where limited initial observational data are used to predict future precipitations or the trajectory of a hurricane. It is only possible to make predictions over finite time horizons, where the growth in initial uncertainty is still not large, although this time horizon is not straightforward to estimate since it depends on the system considered and the quantity of interest.

At the moment, predictions can be made using *precursors*, which are variables that can be tracked and can indicate an imminent an extreme event:

- Ad-hoc precursors: For systems where the mechanism for the production of extreme events is well-understood, it is possible to derive precursors based on the general knowledge of the system. As an example, Cousins and Sapsis [107] derived precursors for the occurrence of rogue waves by separating each wave group and performing a Gabor transform, which quantifies the rate of change of the energy of waves. For turbulent combustion applications, similar precursors have been defined, for example in the case of the swirling flame detachment already mentioned in Sec. 3.3.2. This premixed swirling flame was shown to detach from the nozzle intermittently. Several experimental investigations [70, 71] found that the flame detachment event was correlated with the existence of a precessing vortex core (PVC). Therefore, as a precursor of flame detachment, it was proposed to track the asymmetry of the flow field. This kind of precursor could only be obtained because prior knowledge about the flame detachment mechanism exists. The main disadvantage of this approach is its lack of generality. Advantages include its robustness and interpretability. Arguably, there is no well-defined cost for the methods shown here. The bulk of the cost is driven by the collection of relevant domain knowledge.

Table 3: Overview of computational methods able to forecast extreme events.

Predict ahead of time			
Method	Ref.	Advantages	Requirements
Gabor precursor	[107]	Interpretable and easily measurable	Domain knowledge
System properties precursor	[70, 71]		
Optimally Time-Dependent (OTD) modes	[52, 108]	No prior knowledge of instability mechanism	Observe full system
Variational probing	[51]		Define realizability, observe full system
Supervised data-based	[109]	Mesurable	Observe many instabilities
Semi-supervised data-based	[110]		Artificially generate instabilities
Projection	[111]	Observe no instabilities	Large differences with normal conditions
Clustering	[112]	Measurable	Observe many instabilities
Trajectory identification	[113]		

- Systematic precursors: For systems where little is known about the onset of the extreme event, the ad-hoc strategy is not suitable and should be replaced by a more systematic procedure. The search for systematic precursors has attracted a lot of attention in the dynamical system community. Using the so-called "optimally time-dependent" (OTD) modes [52], precursors of extreme dissipation events could be predicted for Kolmogorov flows [108]. In order to obtain r OTD modes, $r + 1$ direct numerical simulations (DNS) are necessary. In the case of Kolmogorov flow, $r = 2$ was found sufficient to derive precursors [108]. More recently, Farazmand and Sapsis [51] formulated an optimization problem to find the optimal system state that grows the fastest over a short period of time. In particular, this rate was constrained to trajectories that lie on the attractor, which allows for the prediction of spontaneous bursts. This methodology was applied by Blonigan et al. [117] to predict extreme dissipation in turbulent channel flows. Here, a precursor based on the distance to the optimum was shown to be successful at predicting the occurrence of the extreme event. While these high-dimensional examples are comparable to turbulent flow systems, the methodologies rely on the knowledge of the governing equations and produce precursors that require being able to observe the entire state space. As a result, the practical implementation of such tools may be difficult.

  An alternative is the type of precursors that were obtained by Gotoda et al. [109] to predict the onset of thermoacoustic instabilities. The approach used experimental data to fit a predictive model of the pressure oscillations in a lab-scale combustor. However, the construction of this precursor requires a large amount of data which makes this method well-suited for those cases where low-probability events can still be adequately observed.

  With the emergence of techniques able to accurately classify datasets, it may be tempting to naively use such tools to distinguish anomalous behavior from nominal conditions. However, if anomalies have low-probability, these methods may fail as the anomalous class may be underrepresented. One could use semi-supervised learning, i.e. training a generative model [118, 119] able to enrich the data by generating anomalies. Nevertheless, the anomalies generated will necessarily be similar to the ones in the original dataset and may not be representative of all possible anomalies. Careful assessment of such methods is still sparse [110] at the moment.

Another data-based approach consists of training a model that performs well only when exposed to normal data. When exposed to anomalies, the model will then perform poorly which can be used as an indication that the system entered an anomalous mode. For example, Schlegl et al. [111] construct an encoding of normal data on a low dimensional space. When projecting anomalous data on this same space and reconstructing it, discrepancies appear. The advantage of such a method is that the model only needs to see normal data, which are observed with high-probability. However, since extreme events can originate from small discrepancies, it is unclear whether it will be possible to detect extreme events early enough.

- Predictive reduced-order models: Using available observations of the system, one can also build a reduced-order model that can predict the evolution of the quantity of interest over a certain horizon time. The output of the model is richer than a simple anomaly indicator since a time-varying picture can be extracted. Given arbitrary observations, it may not be straightforward to derive a physics-based equation. Instead, data-based tools have been formulated to approximate the trajectory of the system in phase space. These tools can be used as reduced-order models to predict the future outcome, recognizing that in most systems only a small part of phase space can be measured or simulated. For example, the cluster-based reduced-order modeling approached introduced by Kaiser et al. [112] identifies patterns (or clusters) in the flow field and allows for the generation of a transition matrix that gives the probability with which a pattern transitions to another (see also the seminal work of Ref. [120–122]). Other approaches rely on similar clustering techniques to detect anomalies ahead of time by classifying the type of trajectory that a system follows [113]. If a sufficient amount of data capturing the transitions is gathered, these methods can provide a reasonable prediction of the future state of the system. However, the collection of data points is arguably the main driver for the cost for these techniques. Clustering itself has been successfully demonstrated in large dimensional systems [123]. Along the same lines, other reduced-order modeling techniques have gradually emerged (see [2] and references therein) with the perspective of enabling real-time prediction. In the context of data-poor problems, these methods would work if the extreme-event is artificially made more frequent (by triggering it in a surrogate system that emulates the real process) and well-instrumented. For example, if the extreme-event is studied in a lab-scale combustor [72], data that characterize the anomaly can be used to study the phase space trajectory, and later be used on a realistic case. However, just like for any data-driven model, translating conclusions from a lab to a real case is not straightforward.

For practical systems, predicting the onset of an instability is useful when it is known that an instability exists and cannot be avoided. Most engineering applications involve stable designs that avoid such instabilities. In pursuit of efficiency gains, combustion applications will eventually exploit edges of stability maps where one may transition towards an extreme event. Considering again the example of boundary flame flashback mentioned in Sec. 4.2.1; if one decides to operate the combustor near the estimated stability edge, flame flashback may occur at some future time. The ability to predict such an event will allow operating the combustor near the unstable boundaries.

4.2.3. Compute the probability of an event

In the design phase, the probability of extreme events is a valuable metric for all types of events. Such computations have been extensively studied in statistics and dynamical systems fields (see, for example, the review of Morio et al. [133]). In the body of work cited, while the rare events are not necessarily extreme, the techniques address specifically the problem posed by the event rarity. An overview of methods to estimate the probability of a rare event is provided in Tab. 4.

Formally, the computation of the probability of encountering an extreme event is noted as $\gamma = P(Q(\xi > Q))$, where Q is an extreme threshold for the QoI. Monte-Carlo methods provide the most direct approach to estimating this probability. Typically, if the extreme event is driven by uncertainty in initial conditions that leads to an extreme event after a finite time, the strategy would be as follows: 1) sample N initial conditions from the distribution that defines the uncertainty; 2) propagate the system in time and count the number of samples that lead to an extreme event. From an experimental perspective, this strategy would imply repeating the experiment a large number of times in order to sample the right initial conditions sufficiently in order to estimate event statistics. The probability can be

Table 4: Overview of computational methods able to compute the probability of extreme events.

Compute the probability of an event			
Method	Ref.	Advantages	Requirements
Adaptive importance sampling	[124, 125]	Adaptive biasing distribution	Choose a family of biasing distributions
Cross-entropy importance sampling	[126, 127]	Direct approximation of optimal biasing distribution	
Multi-fidelity importance sampling	[128, 129]	Cheap construction of biasing distribution	Low-fidelity model must capture rare event
Adaptive importance splitting	[48, 130]	No need for a biasing distribution	Sample from high-dimensional ensemble
Unstable space sampling	[131]	No need for a biasing distribution and full runs	Identify unstable space
Sequential sampling	[132]	Few forward runs needed	Quantity of interest depends on few parameters

estimated as

$$\widehat{\gamma}^{MC} = \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{I}, \tag{2}$$

where N is the number of observations, and $\boldsymbol{I}$ is an indicator function set to 1 if an extreme event is observed and 0 otherwise. This estimator is unbiased, i.e. its expected value converges to the extreme event probability as $N \to \infty$. While this estimator is sufficient for extreme events of high probability, with a fixed number of observations N, the relative error $RE_{MC} = \frac{\sqrt{Var(\widehat{\gamma}^{MC})}}{\gamma}$ of this estimator can be shown to increase as the probability of the event decreases, i.e. when the extreme event is also a rare event:

$$RE_{MC} = \frac{\sqrt{1-\gamma}}{\sqrt{N\gamma}}. \tag{3}$$

In order to obtain a reliable estimate of probabilities, a large number of realizations will be necessary. Computing a probability of order 10^{-6} with 10% accuracy therefore requires around 10^8 realizations. Alleviating this problem has been the focus of intense research [134, 135] and has led to variance reduction methods for the probability estimator used, while ensuring it remains unbiased. The following methods are the most commonly used.

- Importance sampling: One way to construct an unbiased estimator with reduced variance is by only observing realizations that have a high chance of leading to a low-probability event. This technique is called importance sampling [136]. In practice, this implies that samples for the simulation conditions are not drawn from the nominal probability distribution function (PDF) of the parameters, but a biased distribution that has a higher chance of creating an extreme event. Using the same example as above where the extreme event is driven by uncertainty in initial conditions, instead of sampling from the PDF $\rho(\boldsymbol{\xi}^0)$ of initial conditions, one would sample from a biased distribution $\rho_B(\boldsymbol{\xi}^0)$.

  There exists an optimal biasing distribution that cancels the relative error of the estimator (see [130] for example) which is given by

$$\rho_B(\boldsymbol{\xi}^0) = \begin{cases} \frac{\rho(\boldsymbol{\xi}^0)}{\gamma}, & \text{if } \boldsymbol{\xi}^0 \text{ leads to an extreme event} \\ 0, & \text{otherwise} \end{cases}$$

However, sampling from this optimal distribution is not feasible as it requires to know the probability that one wants to estimate, which defeats the purpose of the computation [133]. An alternative approach is to assume a functional form for the biasing distribution for which the exact expression can be adjusted with parameters that are chosen to minimize biasing criteria [124, 125, 137] (see Ref. [138] for an extensive review). Alternatively, in the cross-entropy importance sampling method, parameters of the biasing distribution are adjusted to minimize the Kullback-Leibler divergence between the biasing distribution and the optimal biasing distribution (the one that leads to zero-variance) [126, 127]. Recently, machine-learning techniques have been used for the approximation of the biasing distribution [139]. In particular, normalizing flows [140] have proven useful as density estimators and are increasingly being used for the generation of events that span the phase-space [141, 142]. In the case of very high-dimensional systems such as a CFD problem, it remains unclear how the biasing distribution can be reasonably defined. Not only does the PDF at every point need to be adjusted, but their joint PDF should also reflect all the physical constraints. Additionally, if the biasing distribution fitting requires many samples that are expensive to collect, the cost importance sampling may remain too high. One path to improving the feasibility of this technique was developed by Peherstorfer et al. [128, 129], where the biasing distribution is obtained with a multifidelity approach. This methodology was applied to predict the probability of large plate deformation due to an external load. As can be seen in Fig. 12, using the biased samples, it was possible to observe extreme deformations more often than by using the nominal distribution of load. Crucially, Ref. [128, 129] demonstrate a cost reduction between one and three orders of magnitude compared to single fidelity methods.

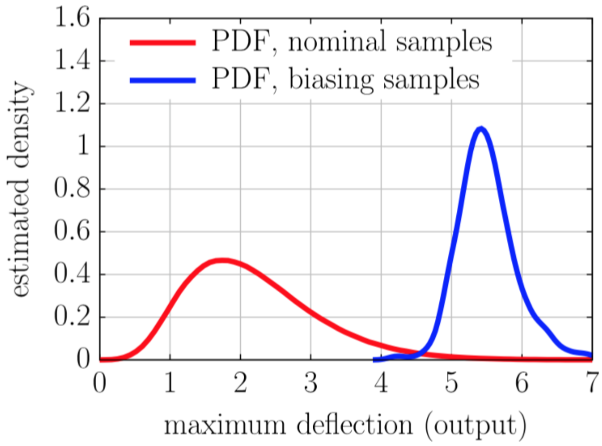

Figure 12: Distribution of observed deformation of a plate exposed to an external load when the external load follows the nominal distribution (red line) and the biased distribution (blue line). Reprinted from [128] with permission of Elsevier

.

Although promising, this approach relies on the fact that the extreme event can be observed with a low-fidelity model. Designing a low-fidelity model with this capability requires to appropriately capture the dynamics of the true system using a limited number of degrees of freedom. At the moment, there exists no robust method that allows designing such models.

- Importance splitting: A different approach to bias the sampling distribution to increase the occurrence of extreme events is importance splitting [133, 143–145]. Compared to importance sampling, importance splitting does not require the definition of a biasing distribution, which offers several numerical advantages [146]. Here, the goal is to find the neighborhood of initial conditions that produce extreme events. The space of initial conditions is sampled based on the nominal distribution, but once a single extreme event is observed, the samples going forward are constrained by the trajectories of observed extreme events. In general, this search is executed in terms of levels defined with respect to the QoI. For instance, a sequence of ordered thresholds for the QoI $Q_1 < Q_2 < ... < Q_n$ and $A_i = \{\boldsymbol{\xi} \mid Q(\boldsymbol{\xi}) > Q_i\}$, then $A_1 \supset A_2 \supset ... \supset A_n$. In importance splitting techniques, the simulations are sampled from A_i to find realizations spanning A_{i+1} and so on. Since its introduction, one key challenge in such methods was the definition of the sequence of thresholds Q_i. However, recent advances allowed adaptive definition of these thresholds [48, 130]. In such cases, the realizations closest to the extreme events are cloned while the ones that are the farthest away are pruned. This method was successfully applied to atmospheric flows and canonical turbulent flows [47, 147]. While these methods are promising, there are also some limitations. First, in the case of time-constrained events, there is no implied guarantee that realizations close to the extreme event at intermediate times will be close to the extreme events at later times. In fact, this assumption fails for oscillating quantities of interest. Rather, the path leading to a low-probability event should be somehow inferred before deciding whether a realization has a high chance of encountering an extreme event. Consider the configuration shown in Fig. 13. In this configuration, a premixed CH_4-air flame (equivalence ratio $\phi = 1$ and unburnt temperature $T_u = 750$ K) is placed in a periodic box and is exposed to a turbulent flow ($Re_\lambda = 30$). The goal is to find the probability of reaching large or low turbulent flame speeds at $t = 15$ ms. Because this problem is deterministic, the QoI depends on the initial flow field distribution, which is a high-dimensional quantity. A typical time-history of the turbulent flame speed is shown in Fig. 13 (top right) and can be compared to an ensemble average obtained using 50 independent realizations (bottom right). It can be observed that the QoI of this problem is highly oscillatory, which implies that an instantaneous large (or low) flame speed does not necessarily lead to a large (or low) final flame speed. Therefore, a naive importance splitting could lead to erroneous results in this case. Instead, the trajectory of a QoI that leads to the low-probability event should be estimated first, before performing any selection. In the context of deterministic systems, a method to estimate the rare trajectory was introduced recently [148]. Second, even in the case where low-probability trajectories might be estimated, it is still necessary to sample turbulent flow realizations, which is itself not obvious, since the PDF of the flow field is high-dimensional and unknown in general. Such questions should be examined to make these methods applicable to turbulent combustion systems. Importance splitting methods require many concurrent DNS runs to explore the space reached after each threshold Q_i. Compared to importance sampling techniques, not much effort has been dedicated to multifidelity or other approaches, and provides new opportunities.

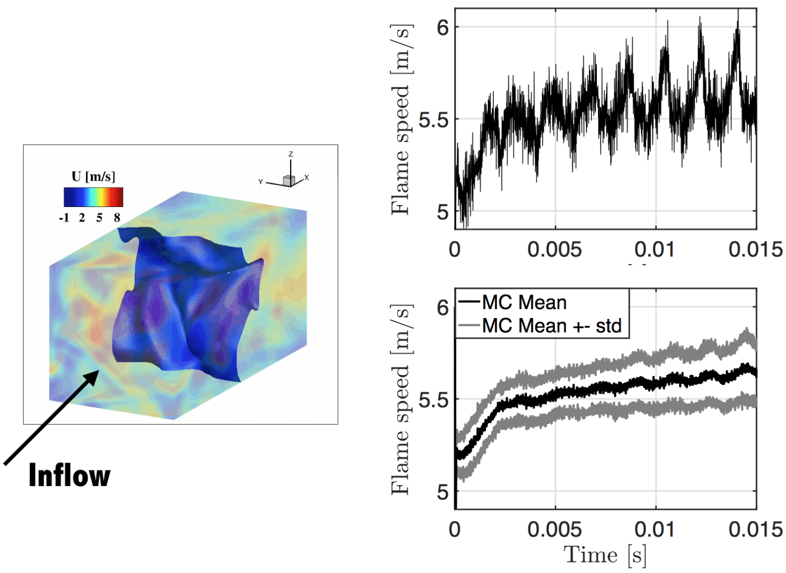

Figure 13: Left: illustration of the flame in a box configuration. A contour of velocity is overlayed with the flame. Top right: time history of the turbulent flame speed of one realization. Bottom right: time history mean (—) and standard deviation (—) of the turbulent flame speed of 50 independent realizations.

- Identify the unstable space: Instead of assuming a distribution for the input set, one can attempt to explore the input space efficiently. In the context of spontaneous bursts (Type III-A events), one could find a reduced-order basis that spans instabilities to efficiently evaluate the volume of an unstable region in an attractor, which can then be used to quantify the probability of a rare event [131]. While the framework is appealing it requires prior understanding of the phenomenon that leads to a rare event. Furthermore, for turbulent flows, there is evidence that the unstable tangent space of turbulent flows becomes more and more aligned with the stable tangent space as the Reynolds number increases [149, 150]. Therefore, it may not be easy to identify a reduced-order description of the unstable subspace in practical applications. One other promising technique was introduced by Mohamad and Sapsis [132] for extreme events that occur in high-dimensional systems, and for which the QoI q is a deterministic function of a small set of parameters α. These parameters are typically not controllable and can be treated as random variables, thereby making the QoI a random variable. For a high-dimensional system, the computational cost associated with evaluating the function $q(\alpha)$ can be large. If one wishes to compute the tails of the PDF of the QoI efficiently, it is possible to find an optimal sequence of samples of α at which the QoI function should be evaluated. This approach has been used to quantify extreme forces on an offshore platform [132]. Since this approach is designed for high-dimensional systems, it is therefore attractive in the context of turbulent combustion. Nevertheless, it comes with important caveats. The function $q(\alpha)$ should be smooth enough to find the optimal sequence of function evaluations. In other words, the QoI should not be too sensitive to the values of the parameters. It is not clear at the moment whether this requirement can be satisfied in turbulent reacting systems. Further, this approach is tractable only if the number of parameters controlling the probability of the QoI is small (of the order $O(10)$).

Knowing the probability of a rare event will help manage risks and weigh the relative cost of control mechanisms. For practical applications, it should be kept in mind that the uncertainty in the probability estimate is not only due to the variance that is being introduced by the different techniques, but also the assumptions made when computing the probability. For example, there are parameters that need to be calibrated based on the application in importance splitting techniques [151]. This is a modeling error that needs to be quantified and is a topic for future exploration.

4.2.4. Predict unobserved events

Table 5: Overview of computational methods able to predict unobserved extreme events.

Predict unobserved events			
Method	Ref.	Advantages	Requirements
Lyapunov weighted dynamics	[50]	No prior knowledge of instability mechanism	Many forward runs
Variational probing	[51, 117]		Define realizability

Simulations can have a dual role in design. A direct use is in the down selection of promising design choices. In this sense, they provide interpolative information, whereby several small changes to the design can be studied rapidly. Another purpose of simulations is to provide information that cannot be obtained using experiments or other approaches. In the current context, predicting events that have not been observed, but could be endemic to the design, will be highly informative for all types of events discussed above. In fact, the ability to predict such events will vastly alter the conservative approach to design. Since the nature of such events is *a priori* not known, it is not feasible to construct probabilistic approaches that map input uncertainty to extreme outputs of interest. Instead, the goal here is to determine causal mechanisms that can lead to the system traversing regions of extreme values. Hence, the exploration of phase space discussed above (Sec. 4.2.4) is not restricted to the space of input or controllable parameters, but the entire phase space accessed by the system. Since for practical systems, the full state space can be very high dimensional, a method to optimally explore this volume in search of extreme events is necessary. Two different exploration methods are considered: first by directly probing the attractor, second by probing different operating conditions. An overview of the methods is provided in Tab. 5.

- Probing the phase space: In practical problems, the phase space of the dynamical system is very high-dimensional. Consequently, a brute force search without a strategy for reducing the volume searched will not be tractable. Hence, the algorithms use some metric to minimize the search volume. For instance, Tailleur and Kurchan [50] use the local measure of chaoticity to determine the region of interest. The chaoticity is measured using a short-time Lyapunov exponent [152], which is obtained by running pairs of simulations that are initially slightly perturbed from one another. The rate of divergence of the fields provides the Lyapunov exponent. A positive and large value shows a high degree of chaoticity, and the possibility of local burst regions. Each pair is termed a phase space walker or crawler. By using several walkers/crawlers in phase space, it is possible to isolate outliers. This method is commonly called *Lyapunov weighted dynamics* and is illustrated for a two-dimensional system in Fig. 14. Although this technique could be used for arbitrarily high-dimensional systems, it has so far been used only in smaller systems. Furthermore, the method required on the order of thousands of concurrent simulations to explore a two-dimensional space. For turbulent combustion problems with much larger dimension ($O(10^9)$), the computational cost may be much larger.

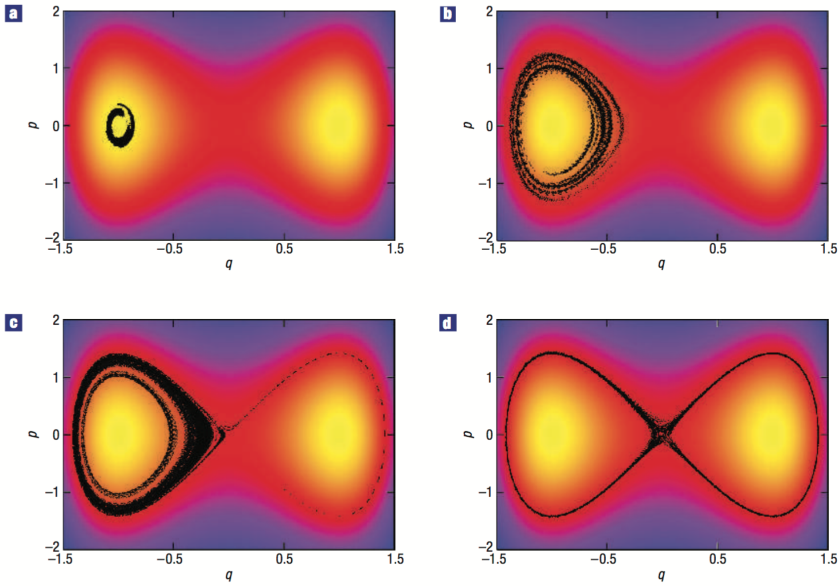

Figure 14: Contour denote a potential surface used to construct a dynamical system. The dots are walkers that explore the phase space in search of the most chaotic trajectories. Such trajectories are obtained by searching for large values of short-time Lyapunov exponents. Reprinted from [50] with permission of Springer Nature

.

In Ref. [51] which was mentioned in Sec. 4.2.2 as a method to derive precursors of extreme events, the authors proposed to solve an optimization problem to discover instabilities in a fluid system. The technique was successfully demonstrated in the case of a Kolmogorov flow and incompressible turbulent channel flow [117]. As presented, the method requires the definition of a realizability condition, which was tailored for incompressible flows. The applicability of this condition to combustion problems needs to be evaluated. Besides the realizability requirement, since most combustion codes are not set up to compute adjoints [153], the cost of optimization may increase significantly if brute force forward simulations are used. The success of such methods relies on the specification of an approach for efficiently probing phase space. This attractor exploration policy would require quantifying the probability of finding an unseen event, given the system states previously explored. The precise definition of the probability could follow the one used in other areas of science [154]. For example, one could define the probability for QoI to reach an extreme event if one advanced the system for n timesteps. The exploration techniques could also be aided by prior knowledge about the structure of the phase space [155–157].

- Operating condition exploration: The attractor of the flow in a combustor does not only depend on the geometry of said combustor but also on the way it is operated. For example, by increasing the Reynolds number starting from a laminar configuration, the system's attracting set will evolve from a being a single point (steady flow) to a multidimensional object [152]. If the operating conditions are variable and lead to anomalous events, it is useful to explore the space of operating conditions. Since the number of such variables is small, such exploration can be computationally tractable. Several alternative approaches to brute force search, including the identification of critical states based on growth rates on the attractors have been developed [158]. There, it was shown that an instability could be anticipated before observing it. However, using this method requires a particular scaling behavior for the system.

Overall, there exist some algorithms in the literature that are capable of phase space and operating condition exploration, especially for large dimensional systems. These methods can be envisioned as computational stress testing that could replace expensive and destructive tests of combustors. However, their cost-effectiveness for turbulent combustion problems has not been established. Efficient exploration techniques cannot only rely on high-fidelity simulations and should be adapted to also use lower fidelity models. Since model development on the combustion side has not focused on extreme events, such a route is not feasible currently.

4.2.5. Obtain bounds on quantities of interest

Table 6: Overview of computational methods able to derive theoretical bounds of quantities of interest.

Obtain bounds on quantities of interest			
Method	Ref.	Advantages	Requirements
Auxiliary function	[159]	Sharp bounds on arbitrary time averages	Ordinary differential equation
Variational method	[160, 161]	Construct optimal (extreme) high-dimensional field	Incompressibility
Background method	[162]	Bounds on the average dissipation rate	Incompressibility
Forward Lyapunov vectors	[163]	Optimum attained in single run	Integration of tangent operator over long time

In some systems, it might be possible to obtain bounds on quantities, without having to explore the phase space directly. For instance, this is particularly relevant for events driven by uncertainty in initial conditions. In the context of high altitude relight (Sec. 3.3), the ignition time is a quantity of interest. The goal is to minimize the time to re-ignite the combustor. However, due to the uncertainty in operating conditions, there might be significant variations in the re-ignition time. One of the computational targets could be to predict the distribution of ignition times. These problems can be recast as determining the bounds of the QoI, which makes them amenable to optimization-based approaches. Two such techniques are as follows (summarized in Tab. 6):

- Bounding QoI: Before direct numerical simulations became routine in fluid mechanics, there had been many efforts dedicated to obtaining analytical bounds on certain flow field quantities such as friction factor or energy dissipation [162, 164] or the drag coefficient of bluff bodies [165]. More recently, optimal bounds for heat transport in Rayleigh-Bénard flows have been developed [160]. It was shown in a subsequent work [159] that even long-time averages of certain QoIs can depend on initial conditions. Similarly, other studies have optimized the best mixing attainable constrained on the amount of energy input to the system [166, 167]. Figure 15 shows results from [166] where a passive tracer is optimally mixed in a two-dimensional box in a divergence-free flow, starting from a sinusoidal distribution in one direction. While these studies are specific to incompressible flows, extensions to variable-density reactive flows need to be pursued.
- Optimal perturbation growth rate: For events driven by initial conditions variations, it can be useful to formulate an optimization problem to find the minimal-perturbation that can cause an extreme event. More formally, the

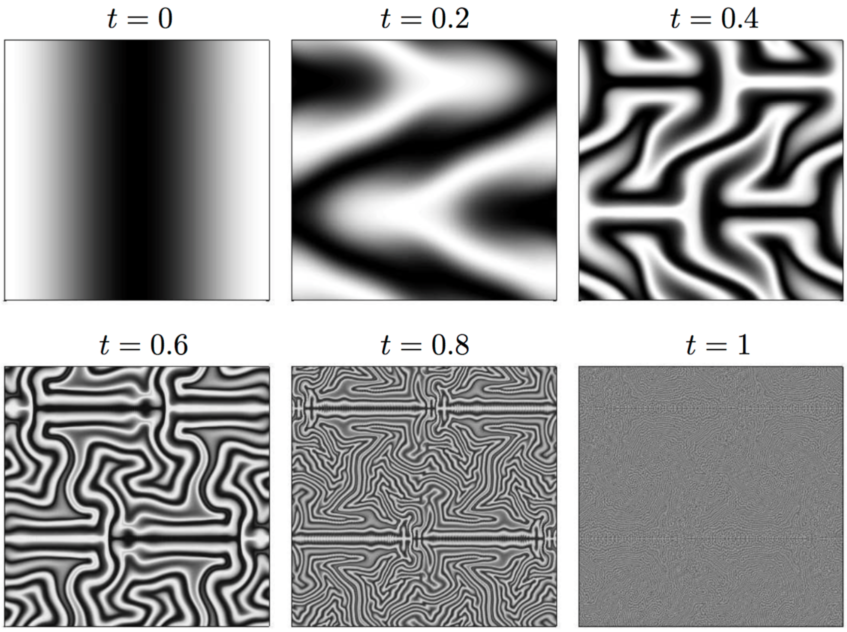

Figure 15: Evolution of a two-dimensional scalar field that mixes from $t = 0$ to $t = 1$ using an optimal mixing strategy starting from a sinusoidal distribution. Reprinted from [166] with permission of Cambridge University Press

.

optimization problem can be formulated as follows: given a fixed energy E for a disturbance $\delta\boldsymbol{\xi}$ that can be added to the flow at time $t = 0$, given a time interval $[0, T]$, find the optimal initial disturbance such that the QoI grows the most over the time interval:

$$\arg\max_{\delta\xi} f(\delta\boldsymbol{\xi}) = \{\frac{Q(\ \boldsymbol{\xi}'(T)\ )}{Q(\ \boldsymbol{\xi}'(0)\ )} \mid \boldsymbol{\xi}'(0) = \boldsymbol{\xi}(0) + \delta\boldsymbol{\xi};\ \ E(\delta\boldsymbol{\xi}) = E_0\}. \tag{4}$$

Such an optimization procedure can be conducted for 3D flows and has been used to study the transition to turbulence (see Ref. [161] and references therein). In the above formulation, the energy E refers to the kinetic energy which is also the QoI. However, such constrained optimization can be applied to any QoI. This approach is central to weather prediction, where bounds in the forecast are routinely obtained by estimating the fastest-growing perturbation. For instance, this approach is implemented within the European weather forecasting model [163]. Here, a coarse representation of initial conditions is obtained from satellite observations and other sensor data. The optimal growth rate is used to estimate the impact of the uncertainty in initial conditions on weather prediction. Starting from a baseline initial condition (the best guess), the tangent linear operator of the dynamical system $\frac{\partial \mathcal{F}}{\partial \xi}\Big|_{\xi}$ is computed over time to find the optimal initial perturbation whose amplitude will grow the most.

These methods have been successfully demonstrated in the context of high-dimensional systems and are therefore good candidates for rare events driven by initial conditions in turbulent combustion problems. Note however that the implementation of such tools can be intrusive in existing combustion codes. In particular, the tangent operator cannot always be obtained easily.

4.2.6. Revise the design of a device

This question requires understanding the cause of an extreme event, for which it is necessary that observational data already exists. Hence, these questions can be dealt with in a data-sufficient environment, provided that methods exist to isolate extreme events and their causal mechanisms from large amounts of data, much of which will contain only nominal events. Here, both data-driven tools and dynamical systems approaches are feasible. An overview of methods to identify the cause of an extreme event is provided in Tab. 7.

Table 7: Overview of computational methods that can be used to infer the cause of an extreme event.

Revise the design of a device			
Method	Ref.	Advantages	Requirements
Conditional averaging of extreme events	[168]	Interpretable path to extreme event	Many observations of extreme event
Reduced order model of extreme path	[72, 112]		Large differences between paths
Sparse sensing	[169]	Done with only forward runs	Interpretation of sensors and many observations
Visualize unstable manifold	[152, 156, 170]	Efficient methods for computation of Lyapunov vectors	Interpretation of Lyapunov vectors

- Trajectory analysis: The purpose of these tools is to use data to classify trajectories that lead to extreme events. In other terms, one can isolate the trajectory that a system follows when it reaches an extreme state. In this vein, the method of instanton filtering [168] has been used for Burgers equations with a stochastic source term. There, the goal was to find the initial conditions leading to the solution with the largest velocity gradient. This is achieved by averaging the initial conditions obtained with several successful runs leading to the desired velocity gradients. More formally, let the extreme event be $q < Q$ at time $t = T$. For $t < T$, a conditional average of the paths that lead to this extreme event is obtained as

$$\overline{\boldsymbol{\xi}(t)} = \{E(\boldsymbol{\xi}(t)) \mid Q(\boldsymbol{\xi}(T)) < Q\}. \tag{5}$$

  Other data analysis tools that are used to approximate the phase space trajectory of the system can also be used to understand how an extreme event occurs. The cluster-based reduced-order modeling approach [112] or symbolic dynamics approach [113] mentioned in Section. 4.2.2 effectively provides a map describing how transitions occur in a flow. This has been recently used to analyze the transition mechanism of flames in swirl combustors [72]. This process is shown in Fig. 16. The trajectory extracted can be used with domain knowledge to infer the cause of an extreme event. However, like most data-based approaches, it requires many observations of the event.

- Discriminant analysis: Given many observations of normal and anomalous conditions, one can pinpoint the main differences by using a sparse sensing approach along with a discriminant analysis [171]. This procedure was applied to identify the cause of ignition failure in a lab-scale combustor [169]. Constructing a relevant discriminant requires many high-fidelity function evaluations. For instance, on the order of $\mathcal{O}(100)$ were used in Ref. [169]. It was found that interpretation of the discriminant analysis may be difficult when the differences between normal and anomalous conditions are small. Nevertheless, when supplemented with additional analysis to confirm or infirm interpretation, the discriminant analysis was successful at guiding causality inference in a high-dimensional turbulent combustion problem.

- Extract unstable regions of the manifold: The last strategy to understand how extreme events occur is to perform a perturbation analysis of the flow field by using the Lyapunov theory (this was briefly discussed in the context of phase space exploration in Sec. 4.2.4). The Lyapunov theory characterizes the tangent operator of the system, i.e., the propagation of perturbations applied to a system. The deterministic dynamical system written in Eq. 1 can be rewritten for infinitesimal perturbations $\delta\boldsymbol{\xi}$ as:

$$\left.\frac{d\delta\boldsymbol{\xi}}{dt}\right|_{\xi} = \left.\frac{\partial\mathcal{F}}{\partial\boldsymbol{\xi}}\right|_{\xi}\delta\boldsymbol{\xi}, \tag{6}$$

  which can now be solved given an initial condition, $\delta\boldsymbol{\xi}^0$. This solution can then be decomposed into a series expansion in terms of Lyapunov vectors (LVs) and Lyapunov exponents (LEs) [172, 173]. The LEs characterize

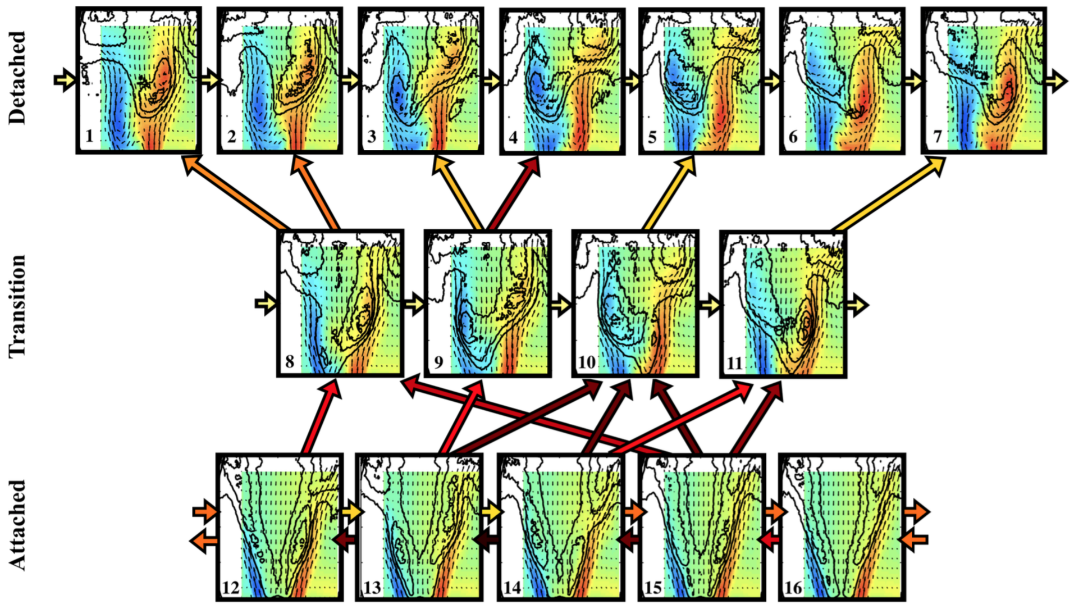

Figure 16: Mechanism leading to flame detachment in a swirl combustor obtained using a clustering strategy on two-dimensional experimental data. The flow goes from bottom to top. Contours indicate the axial velocity and are overlapped with an isoline of OH PLIF data. Arrows indicate paths between each pattern. Light-colored arrows indicate high probability paths. Dark-colored arrows indicate small probability paths. Reprinted from [72] with permission of Taylor & Francis Ltd (www.tandfonline.com)

.

the rate at which the LVs grow and rank the perturbations with regard to their contribution to the chaotic nature of the flow field. Different types of Lyapunov vectors can be used and either describe perturbations to which a system is sensitive (forward LV) or the response of the system to perturbations (backward LV). To compute these quantities, several algorithms have been devised [172, 174, 175]. The Lyapunov vectors can then be correlated to physical phenomena to explain the cause of instabilities. This approach was used by Hassanaly and Raman [156] in the context of turbulent combustion to identify flame extinction and reignition as a large contributor of chaos. Such information could be crucial for understanding how spontaneous and forced extreme events are triggered in turbulent combustion systems.

The LEs and LVs are defined in the long-time limit only. Nevertheless, useful information can also be extracted from the short-time counterpart of the Lyapunov spectrum. Vastano and Moser [152] found from the short-time LV that Kelvin-Helmholtz instability led to the onset of chaotic behavior in the Taylor-Couette flow. Similarly, the generation of instabilities in Couette turbulence [170] and Rayleigh-Bénard convection [176] was explained in great detail using covariant LVs. Note that these approaches require the convergence of LV and are more suited for extreme events related to the long-time behavior of the system (Type III), than the ones that are short-lived (Type II). While one may want to compute many Lyapunov vectors for an extensive characterization of the dynamical system [156, 173], computing the first vector which characterizes the fastest growing instabilities can be obtained at the cost of two concurrent DNS evaluations [172].

Lyapunov analysis can be a powerful tool for chaotic flows. However, it requires detailed simulations that capture the dynamics of interest in the flow field. As a result, this approach can become expensive for turbulent combustion problems. Additionally, some numerical issues specific to variable-density flows need to be handled appropriately [155]. Nevertheless, the breadth of literature on this topic makes it an interesting candidate for the immediate analysis of extreme events in turbulent combustion systems.

5. Future challenges and recommendations

The above discussion shows that extreme events analysis can be a powerful approach to changing the design mindset in the industry. By understanding combustor failure from a computational standpoint, it would be possible to develop less conservative designs. Other fields of engineering and science, most notably weather forecasting, have embraced extreme events modeling as a prime analysis approach. This has helped focus modeling and research priorities, and has provided valuable insights that are not obtainable from analyzing just the average behavior of systems. More importantly, there appear to be many commonalities in how such extreme events arise, through what can be broadly classified as emergent behavior due to the synchronization of different elements of the system. For instance, [177] have studied extreme events that are triggered when multiple coupled neurons can synchronize, or in [178] where energy grid blackouts can occur because of the same synchronization mechanism. Such behavior is reminiscent of the emergence of thermoacoustic instabilities [83, 179].

Given that studying extreme events provides a new approach to gaining insight into complex systems, there is a need for a research strategy moving forward. Below, some essential pathways are identified in an effort to start a discussion on this important topic.

5.1. Develop low-fidelity tools that capture extreme events

At the moment, most numerical tools mentioned in Sec. 4 require to capture with some model the dynamics of the real system that exhibits an extreme event. For example, the computation of the probability of the extreme event requires to be able to observe some realization of this event. Any attractor exploration technique requires that the structure of the attractor is correctly captured. In turbulent combustion problems, the only way to do so is to resolve all the dynamics using direct numerical simulations (DNS). This approach is not acceptable for realistic systems which can be governed by a large range of length and time scales.

Therefore, there exists a crucial need to develop low-fidelity models of turbulent combustion that still capture extreme events and the correct response of the system to perturbations. Note that these models do not need to represent all the dynamics of the system, but the most important ones that eventually lead to an extreme event. Thus, the development of low-fidelity models that can capture extreme events goes hand-in-hand with a better understanding of the process through which these events occur. For thermoacoustic instabilities, flame describing functions (FDF) and flame transfer functions (FTF) [84] are useful in this direction. For other types of extreme events, it is not clear how to accurately design low-fidelity models. The most straightforward techniques such as decreasing the mesh cell count were found to affect, to some extent, the dynamics of the system even when the lower-order statistics are correctly captured [156].

In particular, developing methods that can follow individual realizations rather than statistical measures is essential. In this context, approaches based on manifolds may be useful [64, 157, 180]. Other approaches, such as reduced-order modeling (ROM) could be used to construct representations that specifically target extreme events. When developing such models, quantifying the uncertainty introduced by model reduction becomes important [7]. Further, current approaches to model validation may not be viable when dealing with such extreme event statistics. For instance, Hassanaly and Raman [181] recently showed that while well-established forcing techniques for simulating homogeneous isotropic turbulence yield similar low-order statistics, the dimensions of the attractor, and hence the resulting dynamics may vary.

5.2. Develop rare events tools for turbulent combustion

While there exists a large number of tools to probe different aspects of rare and extreme events (Sec. 4), their direct utilization in the field of turbulent combustion is not straightforward. Many of these tools have been developed for low-dimensional problems, and cannot be extended to practical systems at all. Others, such as the optimal transport discussed in Sec. 4.2.5, assume incompressibility or other constraints to enable such analysis. Hence, there exists a wide range of opportunities to develop tools for combustion applications.

The impact of such tools could have effects in a wide variety of fields. Indeed, turbulent combustion continues to serve as a prototype for complex physics problems. For instance, the National Nuclear Security Agency (NNSA) has supported a predictive science research focus for more than two decades, often using combustion as a surrogate problem for the range of physics needed for nuclear stockpile stewardship. As a result, tools initially developed in the context of combustion physics are now used in many different fields. Likewise, the study of rare and extreme events in combustion physics could translate to significant advances for other high-dimensional complex systems.

Part of developing such computational frameworks also involves identifying validation experiments and configurations. Here again, the broad range of diagnostic tools available to probe combustion physics makes this field an ideal candidate for studying extreme events. Nevertheless, currently available tools are often limited to capturing statistical measures. Establishing experimental techniques for low-probability events is essential. Given that a direct approach to recreating such events is practically intractable, innovative approaches to reproducing the causal mechanisms or measurements that expose the phase space characteristics of the flow would be indispensable.

5.3. Model the triggers of extreme events

Among the triggers of extreme events provided in Sec. 3.3, it was shown that some perturbations could make the system strongly deviate from normal conditions. In order to model the response of the system to such perturbations, it is necessary first to model these external variations. While there have been attempts to provide more detailed boundary conditions to high-fidelity combustion tools [182, 183], modeling such triggering events requires more information on the nature of such processes. For instance, in the high-altitude relight problem (Sec. 3.3), the ignitor that provides the necessary high enthalpy ignition source may itself exhibit considerable variability. The level of detail and quantification of related uncertainties is a major modeling challenge.

However, there is a more subtle notion involved here. Currently, combustion models are developed based on the flow physics present inside the combustor. When external variations can introduce large changes to this flow structure, it becomes necessary that models are sensitive to such perturbations. In this context, models that are regime agnostic [184–186] or those that can adapt to different operating conditions [187–190] are more suitable. In some cases, the flow regime itself can be changed. For instance, in deflagration-to-detonation transition, the flow field is initially subsonic but can become supersonic after the transition. Hence, not only is the modeling of external perturbations important but also the ability of the combustion models to respond to these changes.

5.4. Focus research efforts on the phase space structure

The approach to modeling in turbulent combustion has followed the statistical path not only because it is useful but also because it provides an intuitive description of the extreme complexity in flow physics. In particular, the concept of filtering or statistical averaging delineates the terms in the governing equation that require modeling. However, data-driven tools have embraced dynamical systems or phase space driven approach, whereby the treatment of the modeling problem as a description of the structure of the phase space is most straightforward. As machine learning and data-driven approaches are integrated into modeling, there is a need to rethink the notion of statistical modeling.

In particular, many tools for probing extreme events could be greatly improved through *a priori* knowledge about the structure of phase space. For instance, being able to infer the local phase space structure (dimension, volume, and orientation of the stable manifolds, for instance) can vastly accelerate the search for extreme events. It would also be useful in creating biased sampling approaches (see Sec. 4.2.3). Even information on the dimension of the attractor has not been explored for many canonical closed flows. As an example, it was found that the lower bound of this dimension is far less than the theoretical estimates for the well-studied Sandia flame series [156].

Acknowledgements

This work was authored in part by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. This work was financially supported by Air Force Office of Scientific Research (AFOSR) through grant FA9550-15-1-0378 with Chiping Li as program manager, and grant FA9550-16-1-0309 with Fariba Fahroo as program manager. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes. Insightful discussions with Shivam Barwey, Heeseok Koo, Stephen Voelkel and Ral Bielawski are gratefully acknowledged. We also acknowledge the comments from anonymous reviewers that helped improve the manuscript.

References

[1] K. Duraisamy, G. Iaccarino, H. Xiao, Turbulence modeling in the age of data, Annual Review of Fluid Mechanics 51 (2019) 357–377.

[2] V. Raman, M. Hassanaly, Emerging trends in numerical simulations of combustion systems, Proceedings of the Combustion Institute 37 (2) (2019) 2073–2089.

[3] M. Frenklach, Systematic optimization of a detailed kinetic model using a methane ignition example, Combustion and Flame 58 (1) (1984) 69–72.

[4] H. N. Najm, Uncertainty quantification and polynomial chaos techniques in computational fluid dynamics, Annual review of fluid mechanics 41 (2009) 35–52.

[5] K. Braman, T. A. Oliver, V. Raman, Bayesian analysis of syngas chemistry models, Combustion Theory and Modelling 17 (5) (2013) 858–887.

[6] W. Ji, Z. Ren, Y. Marzouk, C. K. Law, Quantifying kinetic uncertainty in turbulent combustion simulations using active subspaces, Proceedings of the Combustion Institute.

[7] M. E. Mueller, V. Raman, Model form uncertainty quantification in turbulent combustion simulations: Peer models, Combustion and Flame 187 (2018) 137–146.

[8] Z. M. Nikolaou, C. Chrysostomou, L. Vervisch, S. Cant, Modelling turbulent premixed flames using convolutional neural networks: application to sub-grid scale variance and filtered reaction rate, arXiv preprint arXiv:1810.07944.

[9] T. Palmé, M. Fast, M. Thern, Gas turbine sensor validation through classification with artificial neural networks, Applied Energy 88 (11) (2011) 3898–3904.

[10] R. Zweigel, F. Thelen, D. Abel, T. Albin, Iterative learning approach for diesel combustion control using injection rate shaping, in: Control Conference (ECC), 2015 European, IEEE, 2015, pp. 3168–3173.

[11] E. Negri, L. Fumagalli, M. Macchi, A Review of the Roles of Digital Twin in CPS-based Production Systems, Procedia Manufacturing 11 (2017) 939–948.

[12] H. Pitsch, H. Steiner, Large-eddy simulation of a turbulent piloted methane/air diffusion flame (Sandia flame D), Physics of Fluids 12 (10) (2000) 2541–2554.

[13] R. Barlow, J. Frank, Effects of turbulence on species mass fractions in methane/air jet flames, in: Symposium (International) on Combustion, Vol. 27, Elsevier, 1998, pp. 1087–1095.

[14] Seventh International Workshop on Measurement and Computation of Turbulent Non-premixed Flames.

[15] International Sooting Flame Workshop, https://www.adelaide.edu.au/cet/isfworkshop (2018).

[16] Fifth Workshop on Measurement and Computation of Turbulent Spray Combustion.

[17] Workshop on measurement and simulation of coal and biomass conversion, http://www.cbc.uni-due.de/?file=workshop (2019).

[18] W. L. Oberkampf, C. J. Roy, Verification and validation in scientific computing, Cambridge University Press, 2010.

[19] J.-Y. Chen, A Eulerian PDF scheme for LES of nonpremixed turbulent combustion with second-order accurate mixture fraction, Combustion Theory and Modelling 11 (5) (2007) 675–695.

[20] M. Cleary, A. Klimenko, J. Janicka, M. Pfitzner, A sparse-Lagrangian multiple mapping conditioning model for turbulent diffusion flames, Proceedings of the Combustion Institute 32 (1) (2009) 1499–1507.

[21] Y. Ge, M. Cleary, A. Klimenko, A comparative study of Sandia flame series (D–F) using sparse-Lagrangian MMC modelling, Proceedings of the Combustion Institute 34 (1) (2013) 1325–1332.

[22] V. Hiremath, S. R. Lantz, H. Wang, S. B. Pope, Large-scale parallel simulations of turbulent combustion using combined dimension reduction and tabulation of chemistry, Proceedings of the Combustion Institute 34 (1) (2013) 205–215.

[23] M. Ihme, H. Pitsch, Prediction of extinction and reignition in nonpremixed turbulent flames using a flamelet/progress variable model: 2. Application in LES of Sandia flames D and E, Combustion and Flame 155 (1-2) (2008) 90–107.

[24] T. Jaravel, E. Riber, B. Cuenot, P. Pepiot, Prediction of flame structure and pollutant formation of Sandia flame D using Large Eddy Simulation with direct integration of chemical kinetics, Combustion and Flame 188 (2018) 180–198.

[25] W. Jones, V. Prasad, Large Eddy Simulation of the Sandia Flame Series (D–F) using the Eulerian stochastic field method, Combustion and Flame 157 (9) (2010) 1621–1636.

[26] V. Raman, H. Pitsch, A consistent LES/filtered-density function formulation for the simulation of turbulent flames with detailed chemistry, Proceedings of the Combustion Institute 31 (2) (2007) 1711–1719.

[27] M. Sheikhi, T. Drozda, P. Givi, F. Jaberi, S. Pope, Large eddy simulation of a turbulent nonpremixed piloted methane jet flame (Sandia Flame D), Proceedings of the Combustion Institute 30 (1) (2005) 549–556.

[28] M. Yaldizli, K. Mehravaran, F. Jaberi, Large-eddy simulations of turbulent methane jet flames with filtered mass density function, International Journal of Heat and Mass Transfer 53 (11-12) (2010) 2551–2562.

[29] M. Hassanaly, V. Raman, Computational tools for data-poor problems in turbulent combustion, in: AIAA Scitech 2019 Forum, 2019, p. 0998.

[30] T. Sgobba, B-737 MAX and the crash of the regulatory system, Journal of Space Safety Engineering 6 (4) (2019) 299–303.

[31] P. Leal de Matos, A. Tautz, P. Andribet, P. Merlo, Standard Inputs for EUROCONTROL Cost-Benefit Analyses, Tech. rep., Eurocontrol, Brussels, Belgium (2018).

[32] Airbus aircraft 2018 average list prices, Tech. rep., Airbus, Blagnac, France (2018).

[33] J. Wagner, K. Yuceil, A. Valdivia, N. Clemens, D. Dolling, Experimental investigation of unstart in an inlet/isolator model in mach 5 flow, AIAA journal 47 (6) (2009) 1528–1542.

[34] B. Zinn, Real-Time Control of Lean Blowout in a Turbine Engine for Minimizing NOx Emissions, Tech. rep., NASA (2004).

[35] F. A. Administration, Turbine engine power-loss and instability ain extreme conditions of rain and hail, advisory Circular 33.78-1 (2000).

[36] A. H. Lefebvre, Gas turbine combustion, CRC press, 1998.

[37] C. Descamps, C. Bouallou, M. Kanniche, Efficiency of an Integrated Gasification Combined Cycle (IGCC) power plant including CO2 removal, Energy 33 (6) (2008) 874–881.

[38] S. Taamallah, K. Vogiatzaki, F. Alzahrani, E. Mokheimer, M. Habib, A. Ghoniem, Fuel flexibility, stability and emissions in premixed hydrogen-rich gas turbine combustion: Technology, fundamentals, and numerical simulations, Applied energy 154 (2015) 1020–1047.
[39] D. Ebi, N. T. Clemens, Experimental investigation of upstream flame propagation during boundary layer flashback of swirl flames, Combustion and Flame 168 (2016) 39–52.
[40] H. Koo, V. Raman, Large-eddy simulation of a supersonic inlet-isolator, AIAA journal 50 (7) (2012) 1596–1613.
[41] S. Wei, B. Sforzo, J. Seitzman, High-Speed Imaging of Forced Ignition Kernels in Nonuniform Jet Fuel/Air Mixtures, Journal of Engineering for Gas Turbines and Power 140 (7) (2018) 071503.
[42] B. Sforzo, J. Kim, J. Jagoda, J. Seitzman, Ignition probability in a stratified turbulent flow with a sunken fire igniter, Journal of Engineering for Gas Turbines and Power 137 (1) (2015) 011502.
[43] P. S. Zandonade, J. A. Langford, R. D. Moser, Finite-volume optimal large-eddy simulation of isotropic turbulence, Physics of fluids 16 (7) (2004) 2255–2271.
[44] R. J. Adrian, Stochastic estimation of sub-grid scale motions, Applied Mechanics Reviews 43 (1990) 214–218.
[45] M. Ghil, P. Yiou, S. Hallegatte, B. Malamud, P. Naveau, A. Soloviev, P. Friederichs, V. Keilis-Borok, D. Kondrashov, V. Kossobokov, et al., Extreme events: dynamics, statistics and prediction, Nonlinear Processes in Geophysics 18 (3) (2011) 295–350.
[46] P. Del Moral, J. Garnier, et al., Genealogical particle analysis of rare events, The Annals of Applied Probability 15 (4) (2005) 2496–2534.
[47] F. Ragone, J. Wouters, F. Bouchet, Computation of extreme heat waves in climate models using a large deviation algorithm, Proceedings of the National Academy of Sciences 115 (1) (2018) 24–29.
[48] D. Cérou, A. Guyader, Adaptive Multilevel Splitting for Rare Event Analysis, Stochastic Analysis and Applications 25 (2007) 417–443.
[49] E. Vanden-Eijnden, et al., Transition-path theory and path-finding algorithms for the study of rare events., Annual review of physical chemistry 61 (2010) 391–420.
[50] J. Tailleur, J. Kurchan, Probing rare physical trajectories with Lyapunov weighted dynamics, Nature Physics 3 (2007) 203–207.
[51] M. Farazmand, T. P. Sapsis, A variational approach to probing extreme events in turbulent dynamical systems, Science Advances 3 (9) (2017) e1701533.
[52] H. Babaee, T. P. Sapsis, A minimization principle for the description of modes associated with finite-time instabilities, Proc. R. Soc. A 472 (2186) (2016) 20150779.
[53] M. Onorato, A. R. Osborne, M. Serio, Extreme wave events in directional, random oceanic sea states, Physics of Fluids 14 (4) (2002) L25–L28.
[54] D. Solli, C. Ropers, P. Koonath, B. Jalali, Optical rogue waves, Nature 450 (7172) (2007) 1054.
[55] S. M. Krause, S. Börries, S. Bornholdt, Econophysics of adaptive power markets: When a market does not dampen fluctuations but amplifies them, Physical Review E 92 (1) (2015) 012815.
[56] D. Sornette, Predictability of catastrophic events: Material rupture, earthquakes, turbulence, financial crashes, and human birth, Proceedings of the National Academy of Sciences 99 (suppl 1) (2002) 2522–2529.
[57] V. Filimonov, D. Sornette, Quantifying reflexivity in financial markets: Toward a prediction of flash crashes, Physical Review E 85 (5) (2012) 056108.
[58] J. A. Collins, Failure of materials in mechanical design: analysis, prediction, prevention, John Wiley & Sons, 1993.
[59] J. Rasmussen, Human errors. a taxonomy for describing human malfunction in industrial installations, Journal of occupational accidents 4 (2-4) (1982) 311–333.
[60] M. Asch, T. Moore, R. Badia, M. Beck, P. Beckman, T. Bidot, F. Bodin, F. Cappello, A. Choudhary, B. de Supinski, et al., Big data and extreme-scale computing: Pathways to Convergence-Toward a shaping strategy for a future software and data ecosystem for scientific inquiry, The International Journal of High Performance Computing Applications 32 (4) (2018) 435–479.
[61] D. G. Ullman, B. D'Ambrosio, Taxonomy for classifying engineering decision problems and support systems, AI EDAM 9 (5) (1995) 427–438.
[62] P. Westphal, A. S. Sohal, Taxonomy of outsourcing decision models, Production Planning & Control 24 (4-5) (2013) 347–358.
[63] M. Farazmand, T. P. Sapsis, Extreme events: Mechanisms and prediction, Applied Mechanics Reviews 71 (5).
[64] R. Temam, Induced trajectories and approximate inertial manifolds, ESAIM: Mathematical Modelling and Numerical Analysis 23 (3) (1989) 541–561.
[65] J.-P. Eckmann, D. Ruelle, Ergodic theory of chaos and strange attractors, Review of Moderm Physics 57 (1985) 617–656.
[66] J. Milnor, On the concept of attractor, Communications in Mathematical Physics 99 (1985) 177–195.
[67] D. Ruelle, Small Random Perturbations of Dynamical Systems and the Definition of Attractors, Communications in Mathematical Physics 82 (1981) 137–151.
[68] Y. Tang, M. Hassanaly, V. Raman, B. Sforzo, S. Wei, J. M. Seitzman, Simulation of gas turbine ignition using Large eddy simulation approach, in: ASME Turbo Expo 2018, 2018, p. 76216.
[69] Y. Tang, M. Hassanaly, V. Raman, B. Sforzo, J. M. Seitzman, Numerical simulation of forced ignition of Jet-fuel/air using large eddy simulation (LES) and a tabulation-based ignition, in: AIAA Scitech 2019 Forum, 2019, p. 2242.
[70] Q. An, W. Y. Kwong, B. D. Geraedts, A. M. Steinberg, Coupled dynamics of lift-off and precessing vortex core formation in swirl flames, Combustion and Flame 168 (2016) 228–239.
[71] K. Oberleithner, M. Stöhr, S. H. Im, C. M. Arndt, A. M. Steinberg, Formation and flame-induced suppression of the precessing vortex core in a swirl combustor: experiments and linear stability analysis, Combustion and Flame 162 (8) (2015) 3100–3114.
[72] S. Barwey, M. Hassanaly, Q. An, V. Raman, A. Steinberg, Experimental data-based reduced-order model for analysis and prediction of flame transition in gas turbine combustors, Combustion Theory and Modelling 23 (6) (2019) 994–1020.
[73] S. T. Chong, M. Hassanaly, H. Koo, M. E. Mueller, V. Raman, K.-P. Geigle, Large eddy simulation of pressure and dilution-jet effects on soot formation in a model aircraft swirl combustor, Combustion and Flame 192 (2018) 452–472.
[74] H. Koo, M. Hassanaly, V. Raman, M. E. Mueller, K. P. Geigle, Large-eddy simulation of soot formation in a model gas turbine combustor, Journal of Engineering for Gas Turbines and Power 139 (3) (2017) 031503.
[75] V. Raman, R. O. Fox, Modeling of fine-particle formation in turbulent flames, Annual Review of Fluid Mechanics 48 (2016) 159–190.

[76] K. P. Geigle, R. Hadef, W. Meier, Soot formation and flame characterization of an aero-engine model combustor burning ethylene at elevated pressure, Journal of Engineering for Gas Turbines and Power 136 (2) (2014) 021505.
[77] T. Guiberti, L. Zimmer, D. Durox, T. Schuller, Experimental analysis of V-to M-shape transition of premixed CH4/H2/air swirling flames, in: ASME Turbo Expo 2013: Turbine Technical Conference and Exposition, American Society of Mechanical Engineers, 2013, pp. V01AT04A063–V01AT04A063.
[78] S. Candel, D. Durox, T. Schuller, J.-F. Bourgouin, J. P. Moeck, Dynamics of swirling flames, Annual review of fluid mechanics 46 (2014) 147–173.
[79] I. Chterev, C. W. Foley, D. Foti, S. Kostka, A. W. Caswell, N. Jiang, A. Lynch, D. R. Noble, S. Menon, J. M. Seitzman, T. Lieuwen, Flame and flow topologies in an annular swirling flow, Combustion Science and Technology 186 (8) (2014) 1041–1074.
[80] Y. Huang, V. Yang, Bifurcation of flame structure in a lean-premixed swirl-stabilized combustor: transition from stable to unstable flame, Combustion and Flame 136 (3) (2004) 383–389.
[81] F. Bouchet, A. Venaille, Statistical mechanics of two-dimensional and geophysical flows, Physics reports 515 (5) (2012) 227–295.
[82] H. Gotoda, H. Nikimoto, T. Miyano, S. Tachibana, Dynamic properties of combustion instability in a lean premixed gas-turbine combustor, Chaos: An Interdisciplinary Journal of Nonlinear Science 21 (1) (2011) 013124.
[83] M. P. Juniper, R. Sujith, Sensitivity and nonlinearity of thermoacoustic oscillations, Annual Review of Fluid Mechanics 50 (2018) 661–689.
[84] T. Poinsot, Prediction and control of combustion instabilities in real engines, Proceedings of the Combustion Institute 36 (2017) 1–28.
[85] Y. Huang, V. Yang, Dynamics and stability of lean-premixed swirl-stabilized combustion, Progress in energy and combustion science 35 (4) (2009) 293–364.
[86] S. Seo, Parametric study of lean premixed combustion instability in a pressurized model gas turbine combustor., Ph.D. thesis, The Pennsylvania State University (2000).
[87] H.-H. Chiu, J.-S. Huang, Multiple-state phenomena and hysteresis of a combusting isolated droplet, Atomization and Sprays 6 (1).
[88] P. P. Popov, A. Sideris, W. A. Sirignano, Low-Probability Events Leading to Rocket Engine Combustion Instability, AIAA Journal 55 (3) (2017) 919–929.
[89] B. Zhang, Y. Marzouk, B.-Y. Min, T. Sahai, Rare Event Simulation of a Rotorcraft System, in: 2018 AIAA Non-Deterministic Approaches Conference, 2018, p. 1181.
[90] J. D. Jakeman, R. Archibald, D. Xiu, Characterization of discontinuities in high-dimensional stochastic problems on adaptive sparse grids, Journal of Computational Physics 230 (10) (2011) 3977–3997.
[91] K. Sargsyan, C. Safta, B. Debusschere, H. Najm, Uncertainty quantification given discontinuous model response and a limited number of model runs, SIAM Journal on Scientific Computing 34 (1) (2012) B44–B64.
[92] R. Archibald, A. Gelb, R. Saxena, D. Xiu, Discontinuity detection in multivariate space for stochastic simulations, Journal of Computational Physics 228 (7) (2009) 2676–2689.
[93] R. Archibald, A. Gelb, J. Yoon, Polynomial fitting for edge detection in irregularly sampled signals and images, SIAM journal on numerical analysis 43 (1) (2005) 259–279.
[94] A. Gorodetsky, Y. Marzouk, Efficient localization of discontinuities in complex computational simulations, SIAM Journal on Scientific Computing 36 (6) (2014) A2584–A2610.
[95] J. Canny, A computational approach to edge detection, in: Readings in computer vision, Elsevier, 1987, pp. 184–203.
[96] N. Jankowski, M. Grochowski, Comparison of instances seletion algorithms i. algorithms survey, in: International conference on artificial intelligence and soft computing, Springer, 2004, pp. 598–603.
[97] D. R. Wilson, T. R. Martinez, Reduction techniques for instance-based learning algorithms, Machine learning 38 (3) (2000) 257–286.
[98] N. V. Chawla, K. W. Bowyer, L. O. Hall, W. P. Kegelmeyer, SMOTE: synthetic minority over-sampling technique, Journal of artificial intelligence research 16 (2002) 321–357.
[99] P. G. Constantine, Active subspaces: Emerging ideas for dimension reduction in parameter studies, SIAM, 2015.
[100] P. G. Constantine, M. Emory, J. Larsson, G. Iaccarino, Exploiting active subspaces to quantify uncertainty in the numerical simulation of the hyshot ii scramjet, Journal of Computational Physics 302 (2015) 1–20.
[101] W. Ji, Z. Ren, Y. Marzouk, C. K. Law, Quantifying kinetic uncertainty in turbulent combustion simulations using active subspaces, Proceedings of the Combustion Institute 37 (2) (2019) 2175–2182.
[102] W. Ji, J. Wang, O. Zahm, Y. M. Marzouk, B. Yang, Z. Ren, C. K. Law, Shared low-dimensional subspaces for propagating kinetic uncertainty to multiple outputs, Combustion and Flame 190 (2018) 146–157.
[103] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, S. Tarantola, Global sensitivity analysis: the primer, John Wiley & Sons, 2008.
[104] A. S. Tomlin, E. Agbro, V. Nevrlỳ, J. Dlabka, M. Vašinek, Evaluation of combustion mechanisms using global uncertainty and sensitivity analyses: a case study for low-temperature dimethyl ether oxidation, International Journal of Chemical Kinetics 46 (11) (2014) 662–682.
[105] E. Hébrard, A. S. Tomlin, R. Bounaceur, F. Battin-Leclerc, Determining predictive uncertainties and global sensitivities for large parameter systems: A case study for n-butane oxidation, Proceedings of the Combustion Institute 35 (1) (2015) 607–616.
[106] X. Jiang, Y. Tang, Z. Liu, V. Raman, Computational Modeling of Boundary Layer Flashback in a Swirling Stratified Flame Using a LES-Based Non-Adiabatic Tabulated Chemistry Approach, Entropy 23 (5) (2021) 567.
[107] W. Cousins, T. P. Sapsis, Reduced-order precursors of rare events in unidirectional nonlinear water waves, Journal of Fluid Mechanics 790 (2016) 368–388.
[108] M. Farazmand, T. P. Sapsis, Dynamical indicators for the prediction of bursting phenomena in high-dimensional systems, Physical Review E 94 (3) (2016) 032212.
[109] H. Gotoda, Y. Shinoda, M. Kobayashi, Y. Okuno, S. Tachibana, Detection and control of combustion instability based on the concept of dynamical system theory, Physical Review E 89 (2) (2014) 022910.
[110] M. Salem, S. Taheri, J. S. Yuan, Anomaly Generation Using Generative Adversarial Networks in Host-Based Intrusion Detection, in: 2018 9th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), IEEE, 2018, pp. 683–687.
[111] T. Schlegl, P. Seeböck, S. M. Waldstein, U. Schmidt-Erfurth, G. Langs, Unsupervised anomaly detection with generative adversarial net-

works to guide marker discovery, in: International conference on information processing in medical imaging, Springer, 2017, pp. 146–157.
[112] E. Kaiser, B. R. Noack, L. Cordier, A. Spohn, M. Segond, M. Abel, G. Daviller, J. Östh, S. Krajnović, R. K. Niven, Cluster-based reduced-order modelling of a mixing layer, Journal of Fluid Mechanics 754 (2014) 365–414.
[113] C. Rao, A. Ray, S. Sarkar, M. Yasar, Review and comparative evaluation of symbolic dynamic filtering for detection of anomaly patterns, Signal, Image and Video Processing 3 (2) (2009) 101–114.
[114] Lorenz, Deterministic Nonperiodic flow, Journal of the Atmospheric Sciences 20 (1963) 130–141.
[115] Kalnay, Atmospheric Modeling, Data Assimilation and Predictability, Cambridge University Press, 2003.
[116] O. Métais, M. Lesieur, Statistical predictability of decaying turbulence, Journal of the Atmospheric Sciences 43 (1986) 857–870.
[117] P. J. Blonigan, M. Farazmand, T. P. Sapsis, Are extreme dissipation events predictable in turbulent fluid flows?, Physical Review Fluids 4 (4) (2019) 044606.
[118] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, Generative adversarial nets, in: Advances in neural information processing systems, 2014, pp. 2672–2680.
[119] D. P. Kingma, M. Welling, Auto-encoding variational Bayes, arXiv preprint arXiv:1312.6114.
[120] J. Burkardt, M. Gunzburger, H.-C. Lee, POD and CVT-based reduced-order modeling of Navier–Stokes flows, Computer Methods in Applied Mechanics and Engineering 196 (1-3) (2006) 337–355.
[121] Q. Du, V. Faber, M. Gunzburger, Centroidal Voronoi tessellations: Applications and algorithms, SIAM review 41 (4) (1999) 637–676.
[122] T. M. Schneider, B. Eckhardt, J. Vollmer, Statistical analysis of coherent structures in transitional pipe flow, Physical Review E 75 (6) (2007) 066313.
[123] D. Sculley, Web-scale k-means clustering, in: Proceedings of the 19th international conference on World wide web, 2010, pp. 1177–1178.
[124] M.-S. Oh, J. O. Berger, Adaptive importance sampling in Monte Carlo integration, Journal of Statistical Computation and Simulation 41 (3-4) (1992) 143–168.
[125] A. Owen, Y. Zhou, Safe and effective importance sampling, Journal of the American Statistical Association 95 (449) (2000) 135–143.
[126] R. Y. Rubinstein, Optimization of computer simulation models with rare events, European Journal of Operational Research 99 (1) (1997) 89–112.
[127] P.-T. De Boer, D. P. Kroese, S. Mannor, R. Y. Rubinstein, A tutorial on the cross-entropy method, Annals of operations research 134 (1) (2005) 19–67.
[128] B. Peherstorfer, T. Cui, Y. Marzouk, K. Willcox, Multifidelity importance sampling, Computer Methods in Applied Mechanics and Engineering 300 (2016) 490–509.
[129] B. Peherstorfer, B. Kramer, K. Willcox, Multifidelity preconditioning of the cross-entropy method for rare event simulation and failure probability estimation, SIAM/ASA Journal on Uncertainty Quantification 6 (2) (2018) 737–761.
[130] J. Wouters, F. Bouchet, Rare event computation in deterministic chaotic systems using genealogical particle analysis, Journal of Physics A: Mathematical and Theoretical 49 (2016) 374002.
[131] M. A. Mohamad, W. Cousins, T. P. Sapsis, A probabilistic decomposition-synthesis method for the quantification of rare events due to internal instabilities, Journal of Computational Physics 322 (2016) 288–308.
[132] M. A. Mohamad, T. P. Sapsis, Sequential sampling strategy for extreme event statistics in nonlinear dynamical systems, Proceedings of the National Academy of Sciences 115 (44) (2018) 11138–11143.
[133] J. Morio, M. Balesdent, D. Jacquemart, C. Vergé, A survey of rare event simulation methods for static input–output models, Simulation Modelling Practice and Theory 49 (2014) 287–304.
[134] G. Rubino, B. Tuffin, Rare Event Simulation using Monte Carlo Methods, John Wiley & Sons, 2009.
[135] F. Bouchet, J. Rolland, J. Wouters, Rare event sampling methods (2019).
[136] D. Siegmund, Importance sampling in the Monte Carlo study of sequential tests, The Annals of Statistics (1976) 673–684.
[137] M.-S. Oh, J. O. Berger, Integration of multimodal functions by Monte Carlo importance sampling, Journal of the American Statistical Association 88 (422) (1993) 450–456.
[138] S. T. Tokdar, R. E. Kass, Importance sampling: a review, Wiley Interdisciplinary Reviews: Computational Statistics 2 (1) (2010) 54–60.
[139] V. Rao, R. Maulik, E. Constantinescu, M. Anitescu, A Machine-Learning-Based Importance Sampling Method to Compute Rare Event Probabilities, in: International Conference on Computational Science, Springer, 2020, pp. 169–182.
[140] L. Dinh, J. Sohl-Dickstein, S. Bengio, Density estimation using real NVP, arXiv preprint arXiv:1605.08803.
[141] T. Müller, B. McWilliams, F. Rousselle, M. Gross, J. Novák, Neural importance sampling, ACM Transactions on Graphics (TOG) 38 (5) (2019) 1–19.
[142] C. Gao, S. Höche, J. Isaacson, C. Krause, H. Schulz, Event generation with normalizing flows, Physical Review D 101 (7) (2020) 076002.
[143] M. Villen-Altamirano, J. Villen-Altamirano, RESTART: A method for accelerating rare event simulations, Analysis 3 (3).
[144] H. Kahn, T. E. Harris, Estimation of particle transmission by random sampling, National Bureau of Standards applied mathematics series 12 (1951) 27–30.
[145] M. J. J. Garvels, The splitting method in rare event simulation, Ph.D. thesis, University of Twente, Enschede, Netherlands (2000).
[146] C. Jegourel, A. Legay, S. Sedwards, Importance splitting for statistical model checking rare properties, in: International Conference on Computer Aided Verification, Springer, 2013, pp. 576–591.
[147] F. Bouchet, J. Rolland, E. Simonnet, Rare event algorithm links transitions in turbulent flows with activated nucleations, Physical Review Letters 122 (7) (2019) 074502.
[148] M. Hassanaly, V. Raman, A self-similarity principle for the computation of rare event probability, Journal of Physics A: Mathematical and Theoretical 52 (49) (2019) 495701.
[149] M. Inubushi, M. U. Kobayashi, S.-i. Takehiro, M. Yamada, Covariant Lyapunov analysis of chaotic Kolmogorov flows, Physical Review E 85 (1) (2012) 016331.
[150] M. Xu, M. R. Paul, Covariant Lyapunov vectors of chaotic Rayleigh-Bénard convection, Physical Review E 93 (6) (2016) 062208.
[151] M. Balesdent, J. Morio, J. Marzat, Recommendations for the tuning of rare event probability estimators, Reliability Engineering & System Safety 133 (2015) 68–78.

[152] J. A. Vastano, R. D. Moser, Short-time Lyapunov exponent analysis and the transition to chaos in Taylor-Couette flow, Journal of Fluid Mechanics 233 (1991) 83–118.

[153] K. Braman, T. A. Oliver, V. Raman, Adjoint-based sensitivity analysis of flames, Combustion Theory and Modelling 19 (1) (2015) 29–56.

[154] A. Orlitsky, A. T. Suresh, Y. Wu, Optimal prediction of the number of unseen species, Proceedings of the National Academy of Sciences 113 (47) (2016) 13283–13288.

[155] M. Hassanaly, V. Raman, Numerical convergence of the Lyapunov spectrum computed using low Mach number solvers, Journal of Computational Physics 386 (2019) 467–485.

[156] M. Hassanaly, V. Raman, Ensemble-LES analysis of perturbation response of turbulent partially-premixed flames, Proceedings of the Combustion Institute 37 (2) (2019) 2249–2257.

[157] M. Akram, M. Hassanaly, V. Raman, A priori analysis of reduced description of dynamical systems using approximate inertial manifolds, Journal of Computational Physics 409 (2020) 109344.

[158] R. Karnatak, H. Kantz, S. Bialonski, Early warning signal for interior crises in excitable systems, Physical Review E 96 (4) (2017) 042211.

[159] I. Tobasco, D. Goluskin, C. R. Doering, Optimal bounds and extremal trajectories for time averages in nonlinear dynamical systems, Physics Letters A 382 (6) (2018) 382–386.

[160] I. Tobasco, C. R. Doering, Optimal wall-to-wall transport by incompressible flows, Physical Review Letters 118 (26) (2017) 264502.

[161] R. Kerswell, Nonlinear nonmodal stability theory, Annual Review of Fluid Mechanics 50 (2018) 319–345.

[162] C. R. Doering, P. Constantin, Energy dissipation in shear driven turbulence, Physical review letters 69 (11) (1992) 1648.

[163] R. Buizza, T. N. Palmer, The Singular-Vector Structure of the Atmospheric Global Circulation, Journal of the Atmospheric Sciences 52 (1995) 1434–1456.

[164] L. N. Howard, Bounds on flow quantities, Annual Review of Fluid Mechanics 4 (1) (1972) 473–494.

[165] M. L. Wasserman, J. C. Slattery, Upper and lower bounds on the drag coefficient of a sphere in a power-model fluid, AIChE Journal 10 (3) (1964) 383–388.

[166] Z. Lin, J.-L. Thiffeault, C. R. Doering, Optimal stirring strategies for passive scalar mixing, Journal of Fluid Mechanics 675 (2011) 465–476.

[167] G. Mathew, I. Mezić, L. Petzold, A multiscale measure for mixing, Physica D: Nonlinear Phenomena 211 (1-2) (2005) 23–46.

[168] T. Grafke, R. Grauer, T. Schäfer, Instanton filtering for the stochastic Burgers equation, Journal of Physics A: Mathematical and Theoretical 46 (6) (2013) 062002.

[169] M. Hassanaly, Y. Tang, S. Barwey, V. Raman, Data-driven Analysis of Relight variability of Jet Fuels induced by Turbulence, Combustion and Flame 225 (2020) 453–467.

[170] M. Inubushi, S.-i. Takehiro, M. Yamada, Regeneration cycle and the covariant Lyapunov vectors in a minimal wall turbulence, Physical Review E 92 (2) (2015) 023022.

[171] Z. Bai, S. L. Brunton, B. W. Brunton, J. N. Kutz, E. Kaiser, A. Spohn, B. R. Noack, Data-driven methods in fluid dynamics: Sparse classification from experimental data, in: Whither Turbulence and Big Data in the 21st Century?, Springer, 2017, pp. 323–342.

[172] G. Benettin, L. Galgani, A. Giorgilli, J.-M. Strelcyn, Lyapunov characteristic exponents for smooth dynamical systems; a method for computing all of them. part 1: Theory, Meccanica 15 (1980) 21–30.

[173] M. Hassanaly, V. Raman, Perturbation Dynamics in Turbulent Flames, in: 55th AIAA Aerospace Sciences Meeting, 2017, p. 1100.

[174] I. Shimada, T. Nagashima, A numerical approach to ergodic problem of dissipative dynamical systems, Progress of theoretical physics 61 (6) (1979) 1605–1616.

[175] F. Ginelli, P. Poggi, A. Turchi, H. Chaté, R. Livi, A. Politi, Characterizing dynamics with covariant Lyapunov vectors, Physical Review Letters 99 (2007) 130601.

[176] M. Xu, M. R. Paul, Chaotic Rayleigh-Bénard convection with finite sidewalls, Physical Review E 98 (1) (2018) 012201.

[177] A. Mishra, S. Saha, M. Vigneshwaran, P. Pal, T. Kapitaniak, S. K. Dana, Dragon-king-like extreme events in coupled bursting neurons, Physical Review E 97 (6) (2018) 062311.

[178] P. H. Nardelli, N. Rubido, C. Wang, M. S. Baptista, C. Pomalaza-Raez, P. Cardieri, M. Latva-aho, Models for the modern power grid, The European Physical Journal Special Topics 223 (12) (2014) 2423–2437.

[179] H. Gotoda, H. Kobayashi, K. Hayashi, Chaotic dynamics of a swirling flame front instability generated by a change in gravitational orientation, Physical Review E 95 (2) (2017) 022201.

[180] F. Lu, K. K. Lin, A. J. Chorin, Data-based stochastic model reduction for the Kuramoto–Sivashinsky equation, Physica D: Nonlinear Phenomena 340 (2017) 46–57.

[181] M. Hassanaly, V. Raman, Lyapunov spectrum of forced homogeneous isotropic turbulent flows, Physical Review Fluids 4 (11) (2019) 114608.

[182] J. Schlüter, H. Pitsch, P. Moin, Large-eddy simulation inflow conditions for coupling with Reynolds-averaged flow solvers, AIAA journal 42 (3) (2004) 478–484.

[183] M. Klein, A. Sadiki, J. Janicka, A digital filter based generation of inflow data for spatially developing direct numerical or large eddy simulations, Journal of Computational Physics 186 (2) (2003) 652–665.

[184] D. Haworth, Progress in probability density function methods for turbulent reacting flows, Progress in Energy and combustion Science 36 (2) (2010) 168–259.

[185] S. B. Pope, A Monte Carlo method for the PDF equations of turbulent reactive flow, Combustion Science and Technology.

[186] V. Raman, H. Pitsch, A consistent LES/filtered-density function formulation for the simulation of turbulent flames with detailed chemistry, Proceedings of the Combustion Institute 31 (2006) 1711–1719.

[187] P.-D. Nguyen, L. Vervisch, V. Subramanian, P. Domingo, Multidimensional flamelet-generated manifolds for partially premixed combustion, Combustion and Flame 157 (1) (2010) 43–61.

[188] B. Fiorina, Accounting for complex chemistry in the simulations of future turbulent combustion systems, in: 57th AIAA Aerospace Sciences Meeting, 2019, p. 0995.

[189] M. Ihme, Requirements Towards Predictive Simulations of Turbulent Combustion, in: 57th AIAA Aerospace Sciences Meeting, 2019, p. 0996.

[190] M. E. Mueller, A Computationally Efficient Turnkey Approach to Turbulent Combustion Modeling: From Elusive Fantasy to Impending Reality, in: 57th AIAA Aerospace Sciences Meeting, 2019, p. 0994.