PHILOSOPHICAL TRANSACTIONS A

rsta.royalsocietypublishing.org

Review

Cite this article: Sapsis TP. 2018 New perspectives for the prediction and statistical quantification of extreme events in high-dimensional dynamical systems. *Phil. Trans. R. Soc. A* **376**: 20170133.
http://dx.doi.org/10.1098/rsta.2017.0133

Accepted: 24 May 2018

One contribution of 14 to a theme issue 'Nonlinear energy transfer in dynamical and acoustical systems'.

Subject Areas:
applied mathematics, chaos theory

Keywords:
prediction, heavy-tailed stochastic processes, transient instabilities, intermittency, rogue waves, turbulence

Author for correspondence:
Themistoklis P. Sapsis
e-mail: sapsis@mit.edu

New perspectives for the prediction and statistical quantification of extreme events in high-dimensional dynamical systems

Themistoklis P. Sapsis

Department of Mechanical Engineering, Massachusetts Institute of Technology, 77 Massachusetts Ave., Cambridge, MA 02139, USA

TPS, 0000-0003-0302-0691

We discuss extreme events as random occurrences of strongly transient dynamics that lead to nonlinear energy transfers within a chaotic attractor. These transient events are the result of finite-time instabilities and therefore are inherently connected with both statistical and dynamical properties of the system. We consider two classes of problems related to extreme events and nonlinear energy transfers, namely (i) the derivation of precursors for the short-term prediction of extreme events, and (ii) the efficient sampling of random realizations for the fastest convergence of the probability density function in the tail region. We summarize recent methods on these problems that rely on the simultaneous consideration of the statistical and dynamical characteristics of the system. This is achieved by combining available data, in the form of second-order statistics, with dynamical equations that provide information for the transient events that lead to extreme responses. We present these methods through two high-dimensional, prototype systems that exhibit strongly chaotic dynamics and extreme responses due to transient instabilities, the Kolmogorov flow and unidirectional nonlinear water waves.

This article is part of the theme issue 'Nonlinear energy transfer in dynamical and acoustical systems'.

1. Introduction

A plethora of dynamical systems in nature and engineering exhibit intermittent behaviour in the form of

THE ROYAL SOCIETY PUBLISHING

extreme events. These large excursions have significant consequences and are important to predict and characterize statistically. Extreme events are quantified through certain observables of the dynamical system, which exhibit sporadic bursts with values as large as several standard deviations and nonlinear transfers between different scales or modes. In addition, the existence of extreme events is inherently connected with the associated *probability density function* (PDF) of the observable, which is characterized by heavy-tail properties.

Extreme events are low probability realizations, but is every rare event also an extreme event? We start by making an informal distinction between the two. A rare event is, by definition, a very low-probability realization of a random experiment. Getting the same number from a fair dice six times in a row has very low probability. However, there is no special dynamical mechanism that is associated with its occurrence. It is the result of pure randomness. Such rare events have to be separated from the more special notion of extreme events, which, in the context of this work, will be considered as rare realizations caused by the synergistic action of randomness and dynamical instabilities. In particular, in our set-up, extreme events will be caused by finite-time instabilities triggered by the effects of randomness and characterized by nonlinear energy transfers between scales or modes. Finite-time instabilities refer to the case where a finite-time Lyapunov exponent of the system becomes positive but only for a finite-time interval [1]. In this sense, extreme events are a special case of rare events and therefore it is important to emphasize their special properties:

1. The fact that there is a dynamical mechanism involved often increases substantially the intensity of extreme events compared with rare events caused by a purely random process. This is manifested by the heavy tail characteristics of the corresponding probability distribution.
2. Extreme events are usually associated with complex tails in the PDF. This complexity reflects the fact that there is an underlying dynamical mechanism having the form of nonlinear energy transfer, which may not act uniformly for all intensities of the extreme event. For example, in an energy conserving system, such as nonlinear waves, although instabilities can lead to extreme events, these cannot obtain arbitrarily large magnitudes due to finite energy. Therefore, although nonlinearity can induce spatially localized extremes, which are reflected as heavy-tailed regimes in the PDF, the latter have bounded extent because of bounded energy. See figure 1*b* for an example.
3. Extreme events are often associated with a specific time scale over which they develop. This time scale is imposed by the Lyapunov exponent of the associated instability, the time scale related to the nonlinear energy transfer. By identifying the nature of the instability, one can track the corresponding Laypunov exponent and use this as a precursor for the upcoming extreme event. Therefore, for extreme events one may be able to search for appropriate precursors. This is not the case for rare events caused by a purely random process.
4. From an analysis point of view, rare events can be effectively studied using statistical tools. However, for extreme events such an analysis may not be able to capture the statistical complexity due to the presence of non-trivial dynamical phenomena. A blended analysis considering the dynamics and statistics is essential for this case.

In figure 1, we present a typical system exhibiting extreme events due to internal instabilities. The first component is a stochastic attractor or more generally a set where the system state lies most of the time, represented as the brown-shaded region. This can be formed due to persistent instabilities (chaotic dynamics), stochastic parameters or stochastic excitation of the system. The second component is an instability region, represented with green colour. When the dynamical system enters this neighbourhood, we have the rapid growth of certain observables, i.e. the formation of extreme events, due to nonlinear effects. These large excursions are manifested in the PDF of the observable as heavy-tailed regimes (figure 1*b*). Note that the instability domain

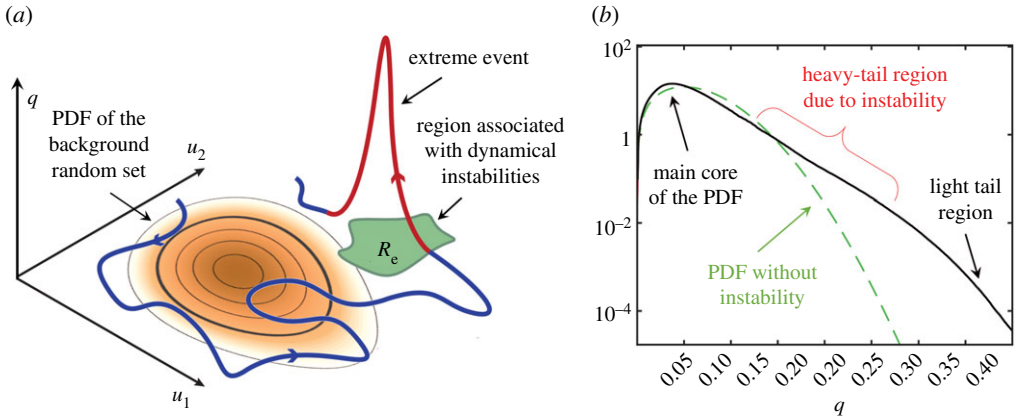

Figure 1. (*a*) Extreme events are associated with large excursions due to the random triggering of dynamical instabilities. The shaded region indicates the PDF associated with the attractor of the system or more generally the set where the system lies most of the time. The instability region is in green. (*b*) A typical heavy-tailed distribution for extreme events. The heavy-tail region has finite extent because instabilities cannot lead to arbitrarily large magnitude. On the other hand, for rare events the PDF has often a uniform tail behaviour. (Online version in colour.)

may have a finite extent in phase space and this will be reflected as a finite extent of the heavy-tail regime in the PDF.

A large number of dynamical systems exhibit extreme events due to transient instabilities, which are randomly triggered while the system evolves in its chaotic attractor. Such a chaotic attractor is formed, for example in turbulent fluid flows, due to persistent instabilities (i.e. positive Lyapunov exponents) and dissipation. For this case, a possible reason for the formation of extreme events is the random triggering of non-normal dynamics [2–4]. In nonlinear waves, we do not have a chaotic attractor but rather a set of possible states formed by the randomness induced by the dispersive mixing of phases (random superposition) and it can trigger nonlinear focusing phenomena that lead to extreme events [5,6]. Complex networks, such as networked populations [7] or communication networks, are another area where we have random perturbations that can trigger large changes in the pattern of the network, e.g. extinction diseases [8] or internal propagation of malware [9]. Similar phenomena can be observed in the mechanical systems excited by stochastic noise, such as parametric instabilities in ship rolling motions [10,11] or buckling of nonlinear beams under combined axial and transverse loads [12].

Numerous efforts have focused on the quantification of heavy tail characteristics through purely statistical approaches. Extreme value theory [13,14] deals with large deviations related to random variables exceeding a given threshold. The theory provides representations for the asymptotic distribution of extreme values from a set of ordered samples given by a random variable and the problem is transformed to estimation of the distribution parameters through maximum-likelihood methods. However, the complexity of the heavy tail characteristics associated with the dynamical mechanism does not always fit within the context of purely statistical approaches (such as statistical extrapolation techniques). In particular, a major limitation is the non-uniform behaviour of the tail due to the complex nature of the instabilities that cause the extreme events. Indeed, for many of the systems mentioned, the PDF has several regimes with different behaviour. Moreover, the application of extreme value theorems often requires a large number of samples which are hard to obtain.

On the other hand, we have *large deviations theory* (LDT) [15–18], a powerful method for the probabilistic quantification of extreme events in sequences of probability distributions. LDT has also been applied in the context of stochastic differential equations, known as Freidlin–Wentzell theory [19], as well as for stochastic partial differential equations [20–23]. LDT uses the dynamical

equations to approximate the asymptotic behaviour of the tail for any given up-crossing level by identifying the most probably initial condition that corresponds to a trajectory that crosses this level. This involves the solution of an optimization problem in the high-dimensional phase space of the dynamical system, which can be expensive to solve. Moreover, the method provides information only for the tail of the PDF and not its full form.

In this work, we rely on a *blended data-equations perspective* that uses a small amount of data as well as the equations of the dynamical system. This is motivated by the very nature of extreme events as described above, i.e. the fact that they form due to the synergistic action of randomness and dynamics and to this end we need information for both. The data information represents a rough description of the underlying stochastic attractor or set where the system lies most of the time (brown-shaded region in figure 1*a*) and is expressed as a small number of samples or second-order statistics. The equations, on the other hand, are used to collect information for the instability regions (green-shaded region in figure 1*a*), such as growth rate or trajectory information for these domains of the phase space, associated with strongly transient dynamics. Note that data-driven approaches may also be used to recover the dynamics for extreme event regions, allowing for the formulation of purely data-driven (equation-free) approaches that use the presented ideas. This is a very important topic that will be treated in the future.

Here, we consider two problems related to extreme events: (i) the short-term prediction problem given current information for the system state, and (ii) the determination of the PDF (including its tails) for quantities of interest given low-order statistics describing the system state. The structure of the paper is as follows. In §2, we present the mathematical formulation of the problem and discuss two prototype systems exhibiting extreme events, which we will use for the demonstration and assessment of the methods presented. In §3, we discuss the prediction problem through the derivation of appropriate precursors which signal the occurrence of an upcoming extreme event. Section 4 deals with the problem of quantification of PDF statistics with emphasis on the form of the tails in connection with the dynamics. In §5, we discuss future directions and conclusion.

2. Prototype systems for extreme events

We consider dynamical systems described by

$$\dot{u} = f(u,t), \quad u \in \mathbb{R}^n, \tag{2.1}$$

where $f : U \times I \to \mathbb{R}^n$ is a sufficiently smooth vector field. Let $u(t; t_0, u_0)$ be the trajectory of system (2.1) with the initial condition u_0. The linearized system around any given trajectory satisfies the variational equation,

$$\dot{v} = L(u(t),t)v, \quad v(t) \in \mathbb{R}^n, \tag{2.2}$$

where $L(u,t) := \nabla_u f(u,t)$. The scalar observable associated with extreme events has the form, $q(t) = T(u(t))$, where T is a functional of the system state.

We consider two prototype systems exhibiting extreme events (figure 2). The first is the Kolmogorov flow, a dissipative, chaotic dynamical system described by the two-dimensional Navier–Stokes equation driven by a monochromatic body forcing. At sufficiently high Reynolds numbers, this flow is known to exhibit intermittent bursts of energy dissipation rate [2,4], which are due to the internal transfers of energy through nonlinearities as opposed to phase locking with the external forcing [24] (figure 1). These extreme events are also manifested by the non-trivial behaviour of the tail for the PDF of the energy dissipation rate. The second prototype system is unidirectional nonlinear water waves described by a *modified nonlinear Schrodinger* (MNLS) equation [25]. Dispersion leads to the continuous mixing of phases between harmonics and the random accumulation of energy in space [5,6]. When energy localized over specific length scales exceeds a certain threshold, nonlinearity leads to the formation of extreme events, which are also demonstrated by the heavy tails in the PDF of the wave elevation (figure 1).

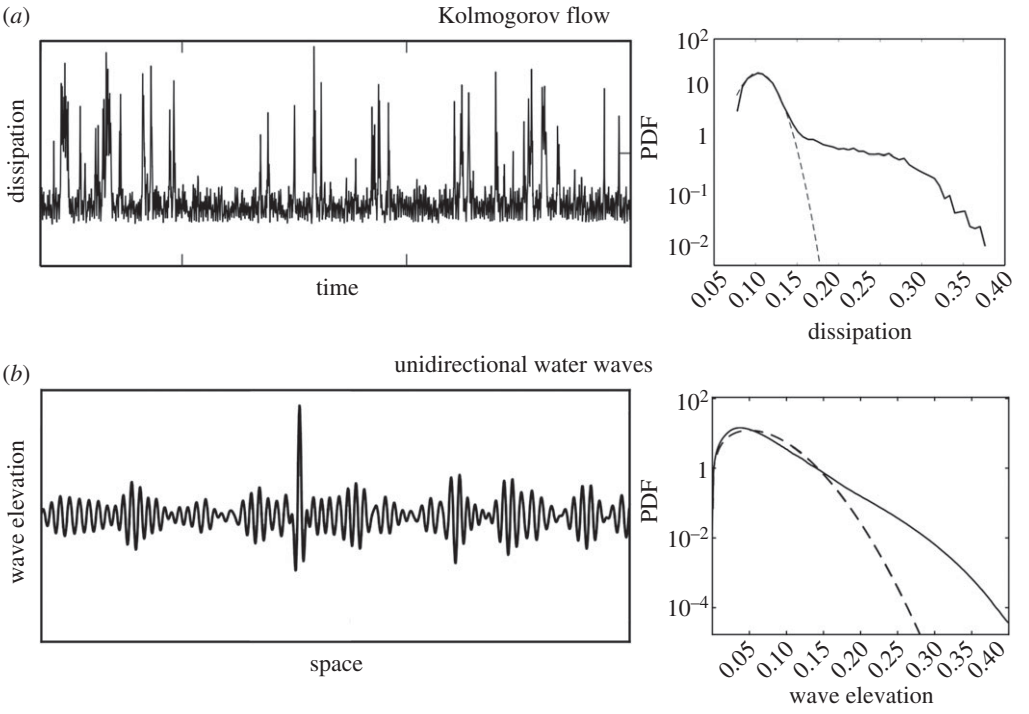

Figure 2. (*a*) Time series and PDF for the dissipation rate in the Kolmogorov flow. The dashed line shows the best Rayleigh PDF approximation. (*b*) Spatial profile associated with an extreme event in nonlinear water waves and the associated PDF for the local maxima.

In this work, we are interested in two problems related to the occurrence of extreme events:

1. *Prediction problem*: We aim to derive functionals of the state of the system, $\alpha(t) = \alpha(u(t))$, that probe future extreme occurrences of the quantity of interest, $q(t)$, over a given prediction time window.
2. *Quantification problem*: Develop efficient sampling algorithms that provide the statistics of the quantity of interest, $q(t)$, with an emphasis on the tail region using a small number of samples.

3. Prediction problem: precursors of extreme events

A critical step towards the understanding and prediction of extreme events is the description of the subspace of associated modes. Knowing what modes are related to the extreme event dynamics and the associated energy transfers allows for the formulation of efficient strategies for their prediction and control. This is a challenging task, however, as this subspace is not necessarily connected to the modes that obtain important energy during the occurrence of an extreme event, but primarily with the modes that trigger the extreme event. Here, we give an overview of two approaches that accomplish this goal.

The first one, the *optimally time-dependent* (OTD) mode [1,26,27] aims to describe, along a trajectory of an infinite-dimensional dynamical system, the directions associated with the strongest growth, i.e. the largest Lyapunov exponent over finite-time intervals. The result is a time-dependent subspace that reveals at every time instant the most unstable directions of the system over finite-times. These are the modes associated with the transient instabilities

and subsequent nonlinear energy transfers that lead to the formation of extreme events. The advantage of the method is that the resulting modes inherently depend on the system state. In this way, the method not only provides the modes associated with the largest growth over finite-times, but it gives only the relevant ones for the current state of the system.

On the other hand, an important drawback of OTD modes is their computational cost as these are time-dependent modes and they have to be resolved as the system evolves. To this end, it is important to formulate a method that maintains the attractive features of the OTD modes, namely to describe state-relevant modes associated with finite-time instabilities, but to circumvent the time-dependent character of the OTD modes. This led us to the adoption of a probabilistic approach, the high-likelihood unstable modes, where, instead of computing the modes associated with the most unstable directions for the current system state, we compute the most unstable directions over finite-times for a constrained set of states associated with important probability of occurrence [24]. The result is a set of modes that describe the most unstable directions for states that have important probability of occurrence. Tracking just these modes provides an effective strategy for the prediction of extreme events [24].

(a) Optimally time-dependent modes

The OTD modes represent a reduced-order set of r time-dependent orthonormal modes, $U(t) = [u_1(t), u_2(t), \ldots, u_r(t)]$, that minimize the difference between the action of the linearized operator $L(u(t), t)$ on $u_i(t)$ and the rate of change of $u_i(t)$. More specifically, the functional that is minimized is given by

$$\mathcal{F}(\dot{u}_1, \dot{u}_2, \ldots, \dot{u}_r) = \sum_{i=1}^{r} \left\| \frac{\partial u_i(t)}{\partial t} - L(u(t), t)u_i(t) \right\|^2, \tag{3.1}$$

with the constraint that the time-dependent basis is orthonormal, $\langle u_i(t), u_j(t) \rangle = \delta_{ij}$, $i, j = 1, \ldots, r$. We emphasize that the minimization of the function (3.1) is considered only with respect to the time-derivative (rate of change) of the basis, $\dot{U}(t)$, instead of the basis $U(t)$ itself. This is because we do *not* want to optimize the subspace that the operator is acting on, but rather find an optimal set of vectors, $\dot{U}(t)$, that best approximates the linearized dynamics in the subspace U. We then solve the resulting equations and compute $U(t)$. We will refer to these modes as the OTD modes, and the space that these modes span as the OTD subspace. By using the minimization principle and taking into account the orthonormality constraint, it can be shown [26] that the OTD subspace is described by the following evolution equations:

$$\frac{\partial U}{\partial t} = LU - UU^{\mathrm{T}}LU. \tag{3.2}$$

The OTD modes possess the following favourable properties for the description of extreme events.

1. For the case of a time-independent operator L, the subspace spanned by the columns of $U(t)$ converges asymptotically to the modes associated with the most unstable directions of the operator L [26].
2. For the general case of a time-dependent operator $L(t)$, under mild conditions, the subspace $U(t)$ aligns with the r most dominant left CauchyGreen strain eigenvectors exponentially fast. To this end, the OTD subspace captures the r largest finite-time Lyapunov exponents [1].

The first property shows how the OTD modes capture persistent instabilities and thus provides a connection with the *dynamic mode decomposition* (DMD) modes for the special case of steady linearized dynamics. DMD was proposed by Schmid [28] for extracting a linear approximation to the flow map of a nonlinear dynamical system. The resulting modes have proved insightful in the analysis of fluid flows [29,30] and shown to have intricate connections to the Koopman and Fourier modes of time periodic solutions [31,32]. However, the real benefit of OTD

modes is their property to capture transient episodes of intense growth, as illustrated by the second property. In particular, they provide the directions associated with the largest finite-time Lyapunov exponents, which is essential for the inherently transient character of extreme events, independently of whether these are caused by exponential or non-normal growth. In other words, computation of the finite-time Lyapunov exponents of the full system can be performed using the projected variational equation to the OTD directions. We emphasize that non-normal growth cannot be quantified through standard methods, such as the eigenvalues or the singular values of the linearized operator $L(u,t)$ (see figure 1 in [1]). For this reason, OTD modes are particularly attractive and robust in capturing transient growth phenomena. OTD modes share some fundamental characteristics of other stability measures. The time varying nature of the OTD modes distinguishes them from the Lyapunov vectors [33,34] and Oseledec subspaces [35,36]. The relation between the OTD modes and the finite-time Lyapunov vectors is examined in [1]. Their connection with covariant Lyapunov vectors [36] remains to be explored. The OTD modes also describe the behaviour of the *dynamically orthogonal* (DO) modes used for uncertainty quantification of high-dimensional dynamical systems [37,38]. This connection has been studied in [1].

We present results illustrating the predictive properties of the OTD modes. A detailed discussion of OTD modes as precursors of extreme events can be found in [27]. For a given scalar quantity of interest $q(t)$, we define

$$\bar{q}(t) = \max_{\tau \in [t+t_i, t+t_f]} q(\tau), \tag{3.3}$$

where $0 < t_i < t_f$ are prescribed numbers that control the time window of prediction in the future. The time window expresses the time scale it takes for the extreme event to grow and depends on the intensity of the instability. The aim is to parametrize the expected future value of the quantity of interest with respect to an indicator $\alpha(t)$. In this case, the indicator will be expressed using measures associated with the OTD subspace. To this end, we use the joint PDF of $\bar{q}$ and α, denoted as $p_{\bar{q},\alpha}$. The conditional PDF of $\bar{q}$ (conditioned on α) is then given by

$$p(\bar{q} \mid \alpha) = \frac{p_{\bar{q},\alpha}(\bar{q}, \alpha)}{p_\alpha(\alpha)}, \tag{3.4}$$

where p_α is the PDF of the indicator α. Roughly speaking, $p(\bar{q}(t) = \bar{q}_0 \mid \alpha(t) = \alpha_0)$ denotes the likelihood of the maximum of the scalar q over the time interval $[t + t_i, t + t_f]$ being q_0 given that the value of α at time t is α_0.

In [27], the predictive skills of OTD modes for extreme events in the Kolmogorov flow as well as in nonlinear waves described by the MNLS equation were studied. For the Kolmogorov flow, the quantity of interest q is the energy dissipation rate, D, and the predictor α is chosen as the largest eigenvalue, λ_1, of the symmetric part of the reduced linearized operator within the OTD subspace, $L_r = \langle U, LU \rangle$ for $r = 8$. For this system, the prediction window is set to $t_i = 3$ and $t_f = 5$. For the nonlinear waves problem, the quantity of interest is the spatial local maximum of the wave field, $\max_x |u(x,t)|$, and the indicator is chosen to be the spatial maximum of the first OTD mode, $\max_x |u_1(x,t)|$. The prediction window in this case is set as $t_i = 25$ and $t_f = 26$. The joint PDF $p_{\bar{q},\alpha}$ is approximated for each case from a large set of numerical simulations and the conditional PDF is then computed through the Bayesian relation (3.4). The conditional PDF for each case is shown in figure 3*a*,*c*. In both cases, the shape of the PDF suggests a successful parametrization with respect to the values of the indicator in regions corresponding to high probability for extreme events and regions of low probability for extreme events. For reference, we also show in figure 3*b* the corresponding PDF using predictors built from DMD modes. For this case, however, the indicator has very limited skill on parametrizing effectively the two regions. This is not a surprise taking into account that DMD modes are designed to capture persistent instabilities, i.e. instabilities associated with long-time behaviour rather than finite-time episodes. Note, however, that variants of DMD such as the multi-resolution DMD [39] and the Hankel alternative view of Koopman

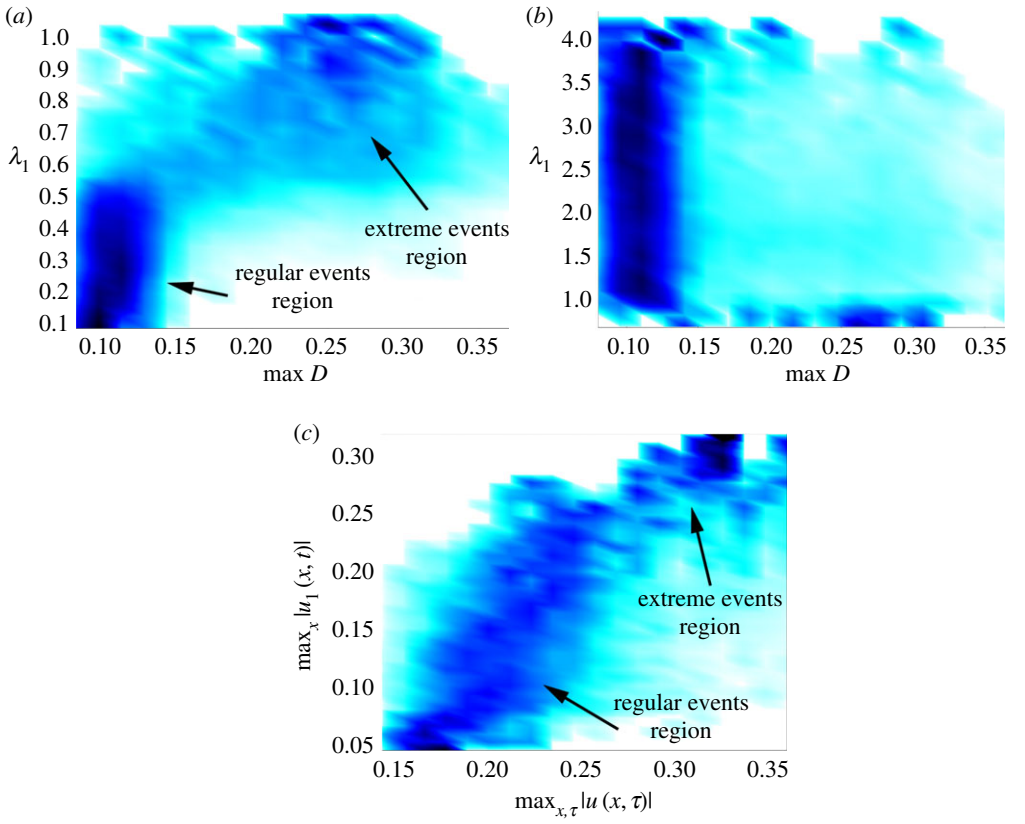

Figure 3. Conditional PDF of the quantity of interest within the prediction window (horizontal axis) with respect to a given value of the indicator on the current time instant (vertical axis). The two left plots (*a*,*b*) refer to extreme events in the Kolmogorov flow using OTD (*a*) and DMD (*b*) modes, while the right plot (*c*) is about extreme nonlinear waves. A conditional PDF that is able to map each value of the predictor to a distinct neighbourhood for the quantity of interest indicates a good predictor (cases (*a*) and (*c*)). This is not the case for (*b*). Reproduced with permission from Farazmand & Sapsis [27]. (Online version in colour.)

modes [40] have been shown to perform satisfactorily in capturing intermittent events in certain chaotic systems.

(b) High-likelihood unstable modes

Despite their favourable properties, OTD modes can be expensive to compute especially for high-dimensional systems. To this end, we seek a formulation that will provide us with *static* modes associated with finite-time instabilities that have high likelihood. This is along the philosophy of OTD modes that give the most unstable directions around the current system state, but it does not require to evolve the modes.

For a general dynamical system, there can be many states that evolve to an extreme event. However, there are only few of those that have high likelihood to occur if the system state 'lives' within the chaotic attractor (or more generally a low-dimensional set as it is the case of the nonlinear waves problem). For this reason, it is not effective to build precursors based exclusively on stability arguments. This led to the formulation of a variational theory [24], where the objective is the determination of modes associated with intense finite-time instabilities under a *feasibility constraint* which guarantees that these modes are associated with states that have high likelihood.

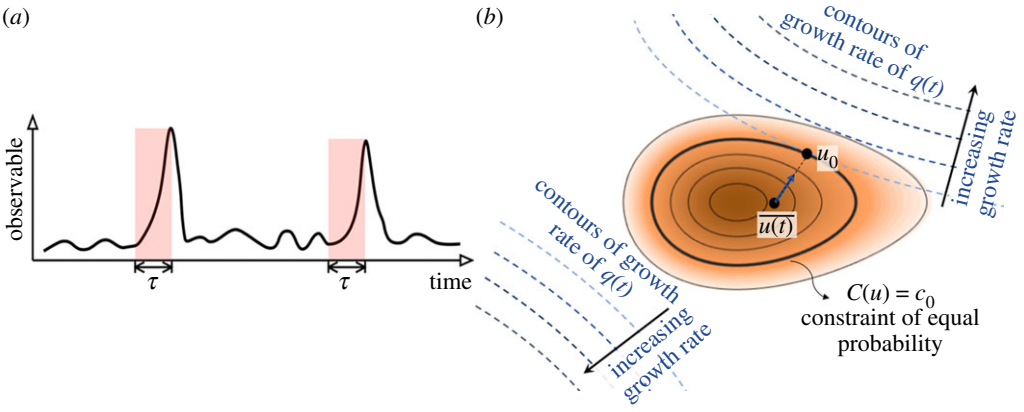

Figure 4. (*a*) A depiction of intermittent bursts of an observable. The highlighted regions mark an approximation of the growth phase of the extreme events. (*b*) The shaded region indicates the probability density function of the system in phase space. The dark contour indicates one possible feasibility constraint associated with given probability of occurrence. The dashed blue lines indicate contours of equal finite-time growth for the quantity of interest q. The optimization problem determines all states that satisfy the feasibility constraint and have the most intense growth. (Online version in colour.)

The constraint is expressed in terms of observed low-order statistics of the system and the growth rate is optimized based on the governing equations of the system. This constraint is of particular importance as it excludes 'exotic' states with rapid growth rates for the quantity of interest, but negligible probability of being observed in practice.

This method is formulated as a constrained optimization problem for which we give a brief overview (see [24] for details). Assume that there is a typical time scale τ over which the bursts in the observable $q(t)$ develop (figure 4*a*). We seek initial conditions u_0 whose associated observable $q(t) = T(u(t))$ attains a maximal growth over the finite-time interval τ. More precisely, we seek the solutions to the constrained optimization problem,

$$\sup_{u_0}(T(u(t_0+\tau)) - T(u(t_0))) \tag{3.5a}$$

and

$$\text{where } \begin{cases} u(t) \text{ satisfies (2.1)}, \\ C(u_0) = c_0, \end{cases} \tag{3.5b}$$

where the optimization variable is the initial condition $u(t_0) = u_0$ of system (2.1). The set of critical states are required to satisfy the constraints in (3.5*b*) in order to enforce two important properties. The first property ensures that $u(t)$ obeys the governing equation (2.1) as opposed to being an arbitrary one-parameter family of functions. The second property $C(u_0) = c_0$ (where C is a co-dimension-k constraint) is enforced to ensure the non-zero probability of occurrence, i.e. states that are sufficiently close to the chaotic background attractor or set of high probability. The set of probabilistically feasible states can generally be described by exploiting basic physical properties of the chaotic attractor or set, such as average energy along different components of the state space or the second-order statistics.

The optimization problem is illustrated graphically in figure 4*b*. The shaded region indicates the PDF of the system state in phase space—it is a representation of the chaotic set for which we have a rough description through e.g. second-order statistics. The bold contour indicates one possible feasibility constraint associated with a given probability of occurrence; this is formulated

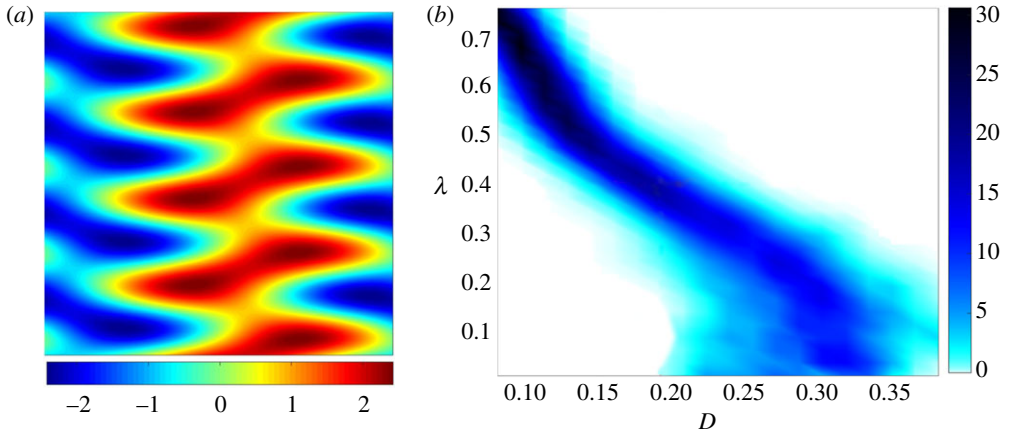

Figure 5. (*a*) High-likelihood unstable mode for the Kolmogorov flow shown in terms of the vorticity field. (*b*) Conditional PDF for the dissipation rate with respect to the precursor obtained from the computed mode. Figure reproduced with permission from Farazmand & Sapsis [24]. (Online version in colour.)

in terms of second-order statistical information. The dashed blue lines indicate contours of equal finite-time growth for the quantity of interest $q(t)$. While there are rapidly growing solutions in two regions of the phase space (lower left and upper right), only those in the upper-right region are associated with important probability of occurrence. 'The optimization problem identifies the states that satisfy the feasibility constraint and have the most intense growth. It provides directions in phase space of intense growth and high-likelihood.' An indicator is then built by measuring the distance of the system state from this critical state. This can be done by measuring the projected distance along the vector $u_0 - \bar{u}$ (figure 4*b*).

We emphasize that there are some fundamental differences between this approach and the general spirit of LDT, where the search is over the full space of initial conditions associated with trajectories that cross a given level; from those, we pick the one with the highest likelihood. In the described approach, we specify a desired likelihood level and we search for the initial condition with the highest growth. In contrast with LDT, the optimization problem here is tractable even for very high dimensions since we are constrained to look over a smaller set—one associated with feasible solutions. Also, it is natural to apply order reduction methods because the set of feasible sets often can be described through a reduced set of variables. The details for the numerical solution of the optimization problem can be found in [24].

We review results related to the application of this method to the two prototype problems. For the Kolmogorov flow, the high dimensionality of the space of solutions leads us to seek for modes associated with intense growth *instantaneously*, so that we have a tractable computational cost. For this case, we consider the limit of $\tau \to 0$ and the growth rate can be computed analytically from the governing equations. The feasibility constraint is formulated in terms of the dissipation rate. In particular, we seek only the states with dissipation rate equal to its statistical mean value (obtained from a short simulation of the system), so that we always remain close to the chaotic attractor. In figure 5*a*, we present the solution of the optimization process. This is the state that has the most intense growth rate in terms of the quantity of interest and has dissipation rate consistent with the chaotic attractor. The magnitude of the projection of the velocity field to this mode, $\lambda(t)$, provides the precursor for this case. The conditional PDF for the dissipation rate D with a prediction window extending from $t_i = 1$ to $t_f = 2$ is shown in figure 5*b*. Note that the prediction horizon can be extended and a detailed study can be found in [24]. We can clearly observe that the derived precursor parametrizes effectively the range of extreme events providing an efficient and parsimonious method for their prediction.

For the water waves problem [41,42] the background chaotic set is not formed due to dissipation, i.e. we do not have a chaotic attractor but rather a random set caused by the random phase between harmonics. We first approximate the stochastic set by a reduced order model. In this way, the space where the optimization is performed is low-dimensional (two dimensional for unidirectional wave-groups [43] and three dimensional for two-dimensional directional waves [42]). Specifically, the space of initial conditions with non-zero likelihood of occurrence can be parametrized as localized wave-groups having random amplitude, A and random length, L [41]:

$$u(x,t) = A \operatorname{sech}\left(\frac{x}{L}\right).$$

Their joint PDF can be computed directly from the wave elevation spectrum (see [41,42] for details), i.e. second-order statistical information. In figure 6*a*, we present this joint PDF for a given sea spectrum. This provides an efficient representation of the random background set. Although there are numerous wave-groups that constitute the full wave-field, it is a valid assumption to consider their evolution independent from each other, at least for dynamical regimes associated with realistic steepness and amplitude [41,42]. To this end, we quantify the evolution of each wave-group as if this was uncoupled to the surrounding waves. In figure 6*b*, we present the normalized (with respect to the initial amplitude) space–time maximum of the solution, $(1/A)\max_{x,t} |u(x,t)|$, for each initial wave-group. While for small initial amplitudes we have linear dynamics (no amplification), for larger initial wave-groups we have amplification due to nonlinear focusing effects. We specify a given likelihood of occurrence and within this set of wave-group amplitudes and lengths we identify the pair with length L^*, corresponding to the largest amplification.

We formulate the precursor for extreme waves by utilizing the projection of the wavefield to this wave-group with critical length scale, L^*:

$$Y_0(x^*,t) = 2L^* \left\langle \operatorname{sech}\left(\frac{x-x^*}{L^*}\right), u(x,t) \right\rangle, \tag{3.6}$$

where the factor $2L^*$ is a normalization constant. In figure 6*c*, we present the conditional PDF of the maximum wave elevation in a future time instant within a prediction window extending from $t_i = 50$ to $t_f = 350$, given the magnitude of the precursor $|Y_0|$ at the current time instant. The fact that the high-likelihood unstable mode is localized in space also allows for the determination of the neighbourhood, x^*, where the extreme event will occur in the near future. Note that compared with the OTD method the precursors based on the highest-likelihood unstable mode provides with longer warning times. This is because OTD modes need a small amount of time to align to the currently most unstable modes. The modes presented have been used for predicting extreme events in nonlinear water waves with high success rates and low false-positive rates [41]. Moreover, in [42] a similar analysis was employed for two-dimensional directional waves.

4. Quantification problem: efficient computation of the probability density function

The next problem we are interested in involves the parsimonious statistical quantification of extreme events. This is an essential task for design and operational purposes but also challenging because it requires a large number of samples for the quantity of interest. Following our formulation, we assume that we have a coarse description of the chaotic attractor or background random set through second-order statistics information or a few samples. We will discuss two recent approaches for the quantification of statistics focusing on extreme event properties. In the first case, we will additionally assume that we have a description of the instability region that leads to the formation of extreme events (green region in figure 1). This will result in a semi-analytical formulation with very high accuracy for the tails of the PDF [44,45]. However, for cases where we do not know explicitly the instability region, we will use a machine learning approach that leads to an efficient sampling strategy for extreme event statistics [46].

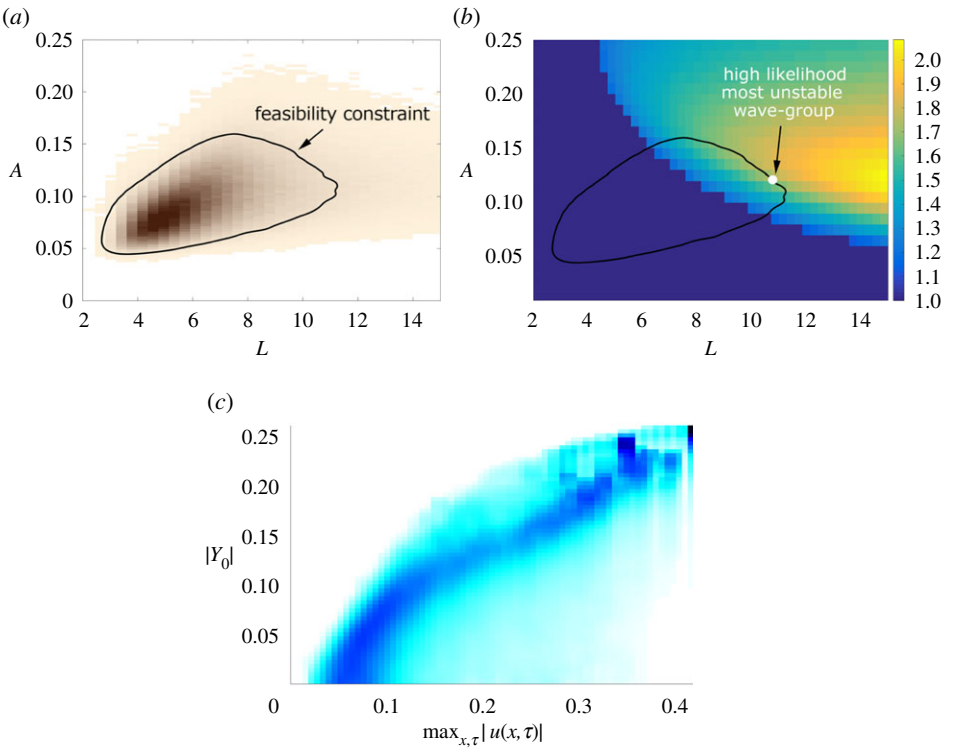

Figure 6. (*a*) Joint PDF of amplitude A and length scale L of the random wave-groups that represent the chaotic set of initial conditions, consistent with the spectrum of the random wave field. (*b*) Space–time maximum of the wave amplitude, normalized by the initial amplitude, for elementary wave-groups. (*c*) Conditional PDF of the maximum wave elevation within the prediction horizon, given the magnitude of the precursor. (Online version in colour.)

(a) Probabilistic decomposition-synthesis method

The idea behind the probabilistic decomposition-synthesis (PDS) method is to rely on an ergodicity assumption for the system dynamics and then employ (i) a special treatment (statistical quantification) of the regimes associated with instabilities which are responsible for the heavy-tail character of the PDF and (ii) second-order statistics for the probabilistic core of the PDF. We first compute the conditional PDF statistics when the system state goes through the instability region (simulating a few carefully selected initial conditions) and then merge this information with the low-order statistical description of the chaotic attractor or background random set. This idea was first developed for oscillators with parametric instabilities (multiplicative noise) having the form of correlated stochastic processes [44]. In this case, the coarse description of the attractor was given by a Langevin approximation of the original equation in the stable regime, while the unstable bursts were approximated analytically. The idea was further developed for stochastic PDEs in [45] and here we give an overview of this analysis.

For the ergodic dynamical system (2.1), we assume that the second-order statistics (i.e. a coarse representation of the chaotic background set) and the instability region, R_e (green region in figure 1), are both also known. We simulate or analytically describe a few trajectories initiated in the unstable region. Using those, we approximate the conditional PDF associated with extreme events, $\rho(q \mid \text{extreme events})$. This conditional PDF describes the statistics of the system during transient instabilities that lead to extreme events. We use the second-order statistics of the background random set through a Gaussian approximation to compute the overall probability

of occurrence of the unstable region, R_e, i.e. the probability that the system will go through this region:

$$\mathbb{P}_r = \mathbb{P}(\text{extreme events}), \quad \text{where } \mathbb{P} \text{ denotes probability.}$$

This probability expresses how frequently the system exhibits extreme events due to the transient instabilities. Using the Gaussian approximation for the conditional statistics of the system state, u, when this is away from the unstable region, R_e, we obtain the conditional statistics of $q = T(u)$ when we have no extreme events (e.g. using derived distributions). This will give the conditional PDF, $\rho(q|\text{regular events})$. Note that this PDF also includes rare occurrences of large magnitude which are not caused by instabilities. The final step is the probabilistic synthesis of this information and this is done through a total probability argument:

$$\rho(q) = \rho(q|\text{extreme events})\mathbb{P}_r + \rho(q|\text{regular events})(1 - \mathbb{P}_r). \tag{4.1}$$

The first term expresses the contribution of extreme events due to internal instabilities and it is the heavy-tailed part of the distribution for q. The second term expresses the contribution of the background random set and accounts for the main probability mass in the PDF for q. The decomposition separates statistical quantities according to the total probability law through conditioning on different dynamical regimes. It provides a partition in terms of the Gaussian 'core' due to the background random set or chaotic attractor and the heavy tails caused by the intermittent bursts.

The framework has been applied to mechanical systems, such as the Mathieu equation excited by a correlated stochastic process [11], optimal mitigation of extreme forcing events using nonlinear springs [47] and the quantification of extreme nonlinear water wave statistics [45]. In figure 7, we present the results of the PDS method for the MNLS equation in the context of nonlinear water waves. The quantity of interest in this case is the local spatial maxima of the wave field. We use the statistical information we employed in the previous section for the derivation of precursors, i.e. the PDF for the amplitudes and lengths of the formed wave-groups, obtained from second-order statistics. The red line defines the unstable region where wave-groups undergo nonlinear focusing. By performing a few simulations of the full equations initiated in the unstable regime, we obtain an approximation of the conditional statistics for extreme events. The conditional statistics for regular events are obtained analytically and have the form of a Rayleigh distribution. We apply the total probability law and obtain the PDF shown in figure 7*b*, which compares favourably with direct Monte Carlo simulations several standard deviations away from zero. The computational cost of the PDS method is extremely small compared with the direct Monte Carlo method as it involves only a few simulations of the full equations for a short time interval. Nevertheless, the PDS method suffers from an important drawback and that is the need to know explicitly the unstable regime. This difficulty is overcome through the formulation of the adaptive sequential sampling scheme presented next.

(b) Adaptive sequential sampling

To overcome the need for explicit knowledge of the unstable region, we develop a sequential sampling strategy of the attractor or background random set that provides with the PDF of the local maxima of the quantity of interest over a specific time window, $\bar{q}(t) = \max_\tau q(t) \triangleq \bar{T}(u(t))$, using only a small number of simulations of the system. Specifically, the ergodic property allows us to obtain extreme event statistics by sampling the attractor or background random set over different initial conditions and performing short time simulations (over a time window of $O(\tau)$) to characterize the quantity of interest. We assume a coarse knowledge of the attractor or background random set, i.e. that u has known second-order statistics. The details of the scheme are presented in [46] and here we give a summary of the approach.

Given a set of $n-1$ samples $\{\hat{u}_1, \hat{u}_2, \ldots, \hat{u}_{n-1}\}$ for which we have performed short time simulation of the full equations and we have quantified the local maxima of the quantity of interest, we aim to identify the next point of the attractor or background random set that we

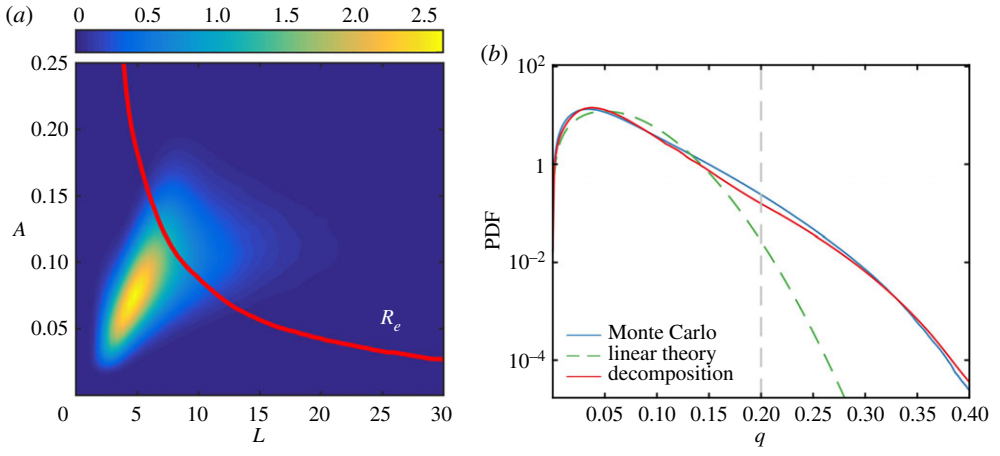

Figure 7. (*a*) Joint PDF of amplitude A and length scale L of the random wave-groups that represent the background random set. The red line denotes the unstable region, R_e i.e. the wave-groups that undergo nonlinear focusing. (*b*) PDF of the wave elevation maxima using the PDS method and comparison with direct Monte Carlo method. The green dashed line indicates the conditional PDF for regular events and the grey line is 4σ. Reproduced with permission from Mohamad *et al.* [45]. (Online version in colour.)

should sample in order to achieve fastest convergence of the PDF for $\bar{q}$. What properties this point should have? One potential strategy could be to sample points that have high probability according to the second-order statistics that we have available for the system state (i.e. sample the dark shaded region in figure 1*a*). But, there is no guarantee that these high probability points will be associated with large magnitude for $\bar{q}$, so we can capture the extreme event statistics accurately. Another strategy would be to sample points associated with large value of $\bar{q}$ (green region in figure 1*a*), but we do not assume that these regions are known, as the map $\bar{T}: u \to \bar{q}$ is unknown, and even if we knew them these may have very small probability of occurrence to contribute to the extreme event statistics, i.e. they may be away from the brown region.

Our approach will focus on sampling points that have high probability but also important contribution to the form of the tail. In particular, we first use the existing samples to machine learn the map $\bar{T}: u \to \bar{q}$. We also estimate the local error caused by the sparse sampling of this map. Then, we formulate a criterion which leads to sampling points that (i) have high probability, and (ii) cause important error reduction for the estimated map in regions where the value of the map is expected to be large (and therefore contributes to the form of the tail).

More specifically we follow these steps (see also figure 8 for a graphical description of the steps):

1. We use a Gaussian process regression (GPR) scheme, based on the existing set of samples $\mathcal{D}_{n-1} = \{\hat{u}_i, \bar{T}(\hat{u}_i)\}_{i=1}^{n-1}$. To estimate $\bar{T}$ we place a Gaussian process (GP) prior over $\bar{T}$ and consider the function values as a realization of a GP. This gives us the posterior mean $\bar{T}_{n-1}(u)$ and variance $\sigma_{n-1}(u)$.
2. We hypothesize a new sample $\hat{u}^*$ which we want to determine according to the properties above, in order to optimally sample the system. We compute the new error map $\sigma_n(u; \hat{u}^*)$. As we do not wish to directly sample the system before we finalize this new point, we assume that the exact value of the map is given by the best linear unbiased estimator, using the $n-1$ points we have already analysed through the GPR scheme, i.e. $\bar{T}(\hat{u}^*) \simeq \bar{T}_{n-1}(\hat{u}^*)$.
3. Using the second-order statistics for u, we compute the PDFs of the upper and lower bounds of the estimated map (denoted as $\rho_n^{\pm}(s; \hat{u}^*)$) i.e. the PDFs of $\bar{T}_{n-1}(u) \pm \sigma_n(u; \hat{u}^*)$. These PDFs take into account the addition of the hypothetical point $\hat{u}^*$.

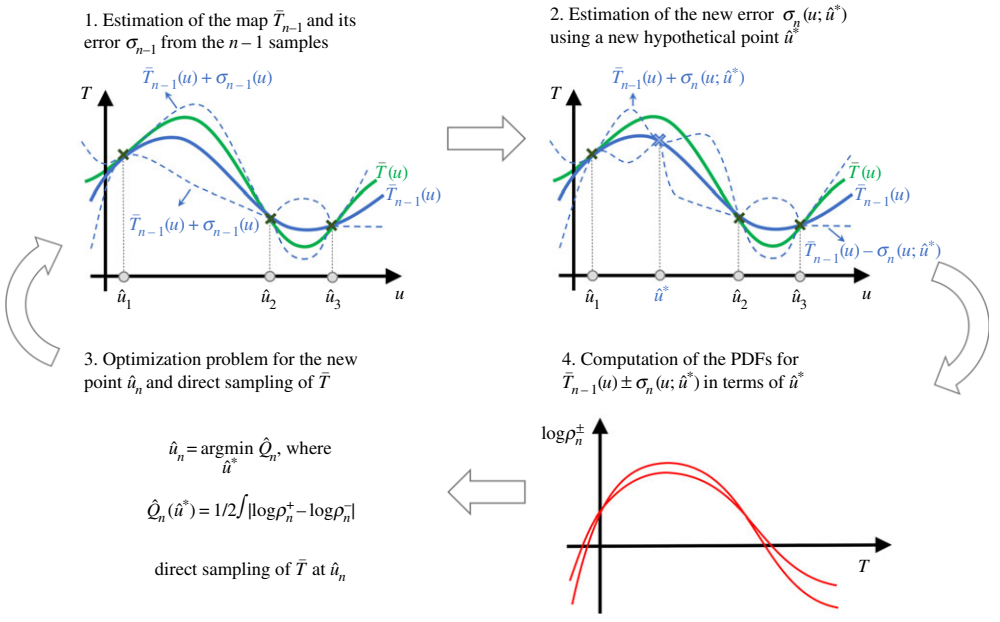

Figure 8. Summary of the adaptive sampling algorithm. The green curve represents the exact map $\bar{T}(u)$ while the blue curves are the GPR estimates using the $n-1$ samples. The dashed curves represent the error region within one standard deviation. (Online version in colour.)

4. The criterion we use to select the 'next-best' point is based on the following L_1 distance between the PDFs

$$\hat{Q}_n(\hat{u}^*) \triangleq \frac{1}{2}\int |\log \rho_n^+(s;\hat{u}^*) - \log \rho_n^-(s;\hat{u}^*)|\, \mathrm{d}s. \tag{4.2}$$

The next sample point $\hat{u}_n$ is chosen so that Q_n is minimized. We finally perform a direct sampling of the system at the new point.

After a sufficient number of points, we can estimate the PDF of $\bar{q}$ using the estimated map and the second-order statistics of u. Note that the logarithm used in the distance emphasizes accuracy for the tail region. In [46], we analytically study the asymptotic behaviour of the criterion and demonstrate the favourable properties of the method on capturing the tail region of the PDF using a small number of samples. In addition, we discuss the details for the implementation and numerical treatment of the optimization problem in the last step.

We emphasize that in the sequential sampling strategy knowledge of the instability region is not needed, in contrast to the PDS method. On the other hand, the adaptive sampling strategy relies on the solution of an expensive optimization problem. Depending on the dimensionality of the problem we may also need to apply order reduction techniques for representing the attractor or background random set in a low enough dimensional space so that the problem is numerically tractable. Nonetheless, the method is valuable for systems where samples are very expensive to obtain and thus is worth solving an optimization problem for each sample. The method has been demonstrated in high dimensional hydromechanical systems excited by random water waves [46] and it allows for the accurate quantification of the PDF for the quantity of interest several standard deviations away from the mean using $O(10)$ samples.

5. Conclusion

We have discussed recent approaches for the prediction and statistical quantification of extreme events in complex dynamical systems characterized by high dimensionality. We have considered systems where transient instabilities and the associated nonlinear energy transfers are responsible for the formation of extreme events. For all the methods presented, the underlying idea is to combine a coarse description of the attractor or background random set (expressed through second-order statistics or a few realizations) with some dynamical information concerning the instabilities of the system that cause the extreme events (expressed through carefully selected simulations of the full equations). The reviewed approaches use both equations and data.

For the prediction problem, we have presented two complementary approaches that allow for the short-term prediction of extreme events by identifying the subspaces related to the transient instabilities. These approaches should be important for applications beyond prediction, such as filtering and control. On the other hand, the developed strategies for statistical quantification allow for the inexpensive computation of the heavy-tailed characteristics for the PDF of interest. Such approaches pave the way for the optimization and design of systems using their higher-order statistical properties, a computational task that has been formidable using traditional methods such as Monte Carlo-based methods. Some results along this direction are presented in [47].

Future directions include the formulation of fully data-driven methods for system for which equations may not be available or they are characterized by important model error. Such efforts should not rely on the purely statistical analysis of the data but rather on the data-driven identification of the underlying dynamics [48–51] and then the subsequent utilization of the discovered models with the same data in a blended, data-driven, model-assisted philosophy.

Data accessibility. No supporting data are used.

Competing interests. I declare I have no competing interests.

Funding. This work has been supported through several grants over the last 5 years including the ONR grant nos. N00014-14-1-0520, N00014-15-1-2381, N00014-17-1-2676, the ARO grant nos. 66710-EG-YIP, W911NF-17-1-0306, the AFOSR grant no. FA9550-16-1-0231 and a grant from the Naval Engineering Education Center.

Acknowledgements. The author thanks his students and postdocs involved in the reviewed work, Dr W. Cousins, Dr M. Farazmand, Dr M. Mohamad, Dr H. Babaee and Dr H.K. Joo.

References

1. Babaee H, Farazmand M, Haller G, Sapsis TP. 2017 Reduced-order description of transient instabilities and computation of finite-time Lyapunov exponents. *Chaos* **27**, 063103. (doi:10.1063/1.4984627)
2. Platt N, Sirovich L, Fitzmaurice N. 1991 An investigation of chaotic Kolmogorov flows. *Phys. Fluids A* **3**, 681–696. (doi:10.1063/1.858074)
3. Chandler J, Kerswell RR. 2013 Invariant recurrent solutions embedded in a turbulent two-dimensional Kolmogorov flow. *J. Fluid Mech.* **722**, 554–595. (doi:10.1017/jfm.2013.122)
4. Farazmand M. 2016 An adjoint-based approach for finding invariant solutions of Navier-Stokes equations. *J. Fluid Mech.* **795**, 278–312. (doi:10.1017/jfm.2016.203)
5. Osborne AR. 2001 The random and deterministic dynamics of 'rogue waves' in unidirectional, deep-water wave trains. *Marine Struct.* **14**, 275–293. (doi:10.1016/S0951-8339(00)00064-2)
6. Cousins W, Sapsis TP. 2014 Quantification and prediction of extreme events in a one-dimensional nonlinear dispersive wave model. *Phys. D* **280**, 48–58. (doi:10.1016/j.physd.2014.04.012)
7. Nieddu GT, Billings L, Kaufman JH, Forgoston E, Bianco S. 2017 Extinction pathways and outbreak vulnerability in a stochastic Ebola model. *J. R. Soc. Interface* **14**, 20160847. (doi:10.1098/rsif.2016.0847)
8. Shaw LB, Schwartz IB. 2010 Enhanced vaccine control of epidemics in adaptive networks. *Phys. Rev. E* **81**, 046120. (doi:10.1103/PhysRevE.81.046120)
9. Lloyd AL, May RM. 2001 How viruses spread among computers and people. *Science* **292**, 1316–1317. (doi:10.1126/science.1061076)

10. Belenky VL, Sevastianov NB. 2007 *Stability and safety of ships: risk of capsizing*. Jersey City, NJ: The Society of Naval Architects and Marine Engineers.
11. Mohamad MA, Sapsis TP. 2016 Probabilistic response and rare events in Mathieu's equation under correlated parametric excitation. *Ocean Eng. J.* **120**, 289–297. (doi:10.1016/j.oceaneng.2016.03.008)
12. Nayfeh AH, Mook DT. 1984 *Nonlinear oscillations*. New York, NY: Wiley-Interscience.
13. Nicodemi M. 2009 Extreme value statistics. In *Encyclopedia of complexity and systems science* (ed. RA Meyers), E:3317. New York, NY: Springer.
14. Coles SG 2001 *An introduction to statistical modeling of extreme values*. Springer Series in Statistics. New York, NY: Springer.
15. Varadhan SRS. 1984 *Large deviations and applications*. Philadelphia, PA: SIAM.
16. Varadhan SRS. 2008 Special invited paper: large deviations. *Ann. Probab.* **36**, 397–419. (doi:10.1214/07-AOP348)
17. Dembo A, Zeitouni O. 2000 *Large deviations techniques and applications*. New York: Springer-Verlag.
18. Stroock D. 1984 *An introduction to the theory of large deviations*. New York, NY: Springer.
19. Freidlin MI, Wentzell AD. 1998 *Random perturbations of dynamical systems*, 2nd edn. Berlin, Germany: Springer.
20. Chow P-L. 1992 Some parabolic Ito equations. *Commun. Pure Appl. Math.* **45**, 97–120. (doi:10.1002/cpa.3160450105)
21. Sowers R. 1992 Large Deviations for a reaction diffusion equation with non-Gaussian perturbations. *Ann. Probab.* **20**, 504–537. (doi:10.1214/aop/1176989939)
22. Sritharan SS, Sundar P. 2006 Large deviations for the two-dimensional Navier-Stokes equations with multiplicative noise. *Stoch. Process. Appl.* **116**, 1636–1659. (doi:10.1016/j.spa.2006.04.001)
23. Dematteis G, Grafke T, Vanden-Eijnden E. 2018 Rogue waves and large deviations in deep sea. *Proc. Natl Acad. Sci. USA* **115**, 855–860. (doi:10.1073/pnas.1710670115)
24. Farazmand M, Sapsis TP. 2017 A variational approach to probing extreme events in turbulent dynamical systems. *Sci. Adv.* **3**, e1701533. (doi:10.1126/sciadv.1701533)
25. Trulsen K, Dysthe KB. 1996 A modified nonlinear Schrödinger equation for broader bandwidth gravity waves on deep water. *Wave Motion* **24**, 281–289. (doi:10.1016/S0165-2125(96)00020-0)
26. Babaee H, Sapsis TP. 2016 A minimization principle for the description of modes associated with finite-time instabilities. *Proc. R. Soc. A* **472**, 20150779. (doi:10.1098/rspa.2015.0779)
27. Farazmand M, Sapsis TP. 2016 Dynamical indicators for the prediction of bursting phenomena in high-dimensional systems. *Phys. Rev. E* **032212**, 1–31. (doi:10.1103/PhysRevE.94.032212)
28. Schmid PJ. 2010 Dynamic mode decomposition of numerical and experimental data. *J. Fluid Mech.* **656**, 5–28. (doi:10.1017/S0022112010001217)
29. Schmid PJ, Li L, Juniper MP, Pust O. 2011 Applications of the dynamic mode decomposition. *Theor. Comput. Fluid Dyn.* **25**, 249–259. (doi:10.1007/s00162-010-0203-9)
30. Schmid PJ. 2011 Application of the dynamic mode decomposition to experimental data. *Exp. Fluids.* **50**, 1123–1130. (doi:10.1007/s00348-010-0911-3)
31. Rowley W, Mezic I, Bagheri S, Schlatter P, Henningson DS. 2009 Spectral analysis of nonlinear flows. *J. Fluid Mech.* **641**, 115–127. (doi:10.1017/S0022112009992059)
32. Chen KK, Tu JH, Rowley CW. 2012 Variants of dynamic mode decomposition, boundary condition, Koopman, and Fourier analyses. *J. Nonlinear Sci.* **22**, 887–915. (doi:10.1007/s00332-012-9130-9)
33. Egolf DA, Melnikov IlV, Pesch W, Ecke RE. 2000 Mechanisms of extensive spatiotemporal chaos in Rayleigh–Bénard convection. *Nature* **404**, 733–736. (doi:10.1038/35008013)
34. Farazmand M, Haller G. 2013 Attracting and repelling Lagrangian coherent structures from a single computation. *Chaos* **23**, 023101. (doi:10.1063/1.4800210)
35. Oseledec VI. 1968 A multiplicative ergodic theorem. Lyapunov characteristic numbers for dynamical systems. *Trans. Moscow Math. Soc* **19**, 197–231.
36. Ginelli F, Poggi P, Turchi A, Chaté H, Livi R, Politi A. 2007 Characterizing dynamics with covariant Lyapunov vectors. *Phys. Rev. Lett.* **99**, 130601. (doi:10.1103/PhysRevLett.99.130601)
37. Sapsis TP, Lermusiaux PFJ. 2009 Dynamically Orthogonal field equations for continuous stochastic dynamical systems. *Phys. D* **238**, 2347–2360. (doi:10.1016/j.physd.2009.09.017)

38. Sapsis TP. 2013 Attractor local dimensionality, nonlinear energy transfers, and finite-time instabilities in unstable dynamical systems with applications to 2D fluid flows. *Proc. R. Soc. A* **469**, 20120550. (doi:10.1098/rspa.2012.0550)
39. Kutz JN, Fu X, Brunton SL. 2015 Multi-resolution dynamic mode decomposition. *SIAM J. Appl. Dyn. Syst.* **15**, 713–735. (doi:10.1137/15M1023543)
40. Brunton SL, Brunton BW, Proctor JL, Kaiser E, Kutz JN. 2017 Chaos as an intermittently forced linear system. *Nat. Commun.* **8**, 19. (doi:10.1038/s41467-017-00030-8)
41. Cousins W, Sapsis TP. 2016 Reduced-order precursors of rare events in unidirectional nonlinear water waves. *J. Fluid. Mech.* **790**, 368–388. (doi:10.1017/jfm.2016.13)
42. Farazmand M, Sapsis TP. 2017 Reduced-order prediction of rogue waves in two-dimensional deep-water waves. *J. Comput. Phys.* **340**, 418–434. (doi:10.1016/j.jcp.2017.03.054)
43. Cousins W, Sapsis TP. 2015 Unsteady evolution of localized unidirectional deep-water wave groups. *Phys. Rev. E - Stat. Nonlinear Soft Matter Phys.* **91**, 063204. (doi:10.1103/PhysRevE.91.063204)
44. Mohamad MA, Sapsis TP. 2015 Probabilistic description of extreme events in intermittently unstable systems excited by correlated stochastic processes. *SIAM ASA J. Uncertain. Quantification* **3**, 709–736. (doi:10.1137/140978235)
45. Mohamad MA, Cousins W, Sapsis TP. 2016 A probabilistic decomposition-synthesis method for the quantification of rare events due to internal instabilities. *J. Comput. Phys.* **322**, 288–308. (doi:10.1016/j.jcp.2016.06.047)
46. Mohamad MA, Sapsis TP. 2018 A sequential sampling strategy for extreme event statistics in nonlinear dynamical systems. (arXiv:1804.07240)
47. Joo HK, Mohamad MA, Sapsis TP. 2017 Extreme events and their optimal mitigation in nonlinear structural systems excited by stochastic loads: application to ocean engineering systems. *Ocean Eng.* **142**, 145–160. (doi:10.1016/j.oceaneng.2017.06.066)
48. Wan ZY, Vlachas PR, Koumoutsakos P, Sapsis TP. 2018 Data-assisted reduced-order modeling of extreme events in complex dynamical systems. *PLoS ONE* **13**, e0197704. (doi:10.1371/journal.pone.0197704)
49. Vlachas PR, Byeon W, Wan ZY, Sapsis TP, Koumoutsakos P. 2018 Data-driven forecasting of high-dimensional chaotic systems with long-short term memory networks. *Proc. R. Soc. A* **474**, 20170844. (doi:10.1098/rspa.2017.0844)
50. Wan ZY, Sapsis TP. 2017 Reduced-space Gaussian process regression for data-driven probabilistic forecast of chaotic dynamical systems. *Phys. D: Nonlinear Phenom.* **345**, 40–55. (doi:10.1016/j.physd.2016.12.005)
51. Brunton SL, Proctor JL, Kutz JN. 2016 Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proc. Natl Acad. Sci. USA* **113**, 3932–3937. (doi:10.1073/pnas.1517384113)