Science Advances

AAAS

advances.sciencemag.org/cgi/content/full/3/9/e1701533/DC1

Supplementary Materials for

A variational approach to probing extreme events in turbulent dynamical systems

Mohammad Farazmand and Themistoklis P. Sapsis

Published 22 September 2017, *Sci. Adv.* **3**, e1701533 (2017)
DOI: 10.1126/sciadv.1701533

The PDF file includes:

Other Supplementary Material for this manuscript includes the following:
(available at advances.sciencemag.org/cgi/content/full/3/9/e1701533/DC1)

- movie S1 (.mp4 format). The prediction of an extreme event in the Kolmogorov flow.

section S1. Derivation of the Euler-Lagrange equation

In this section, we detail the derivation of the Euler–Lagrange equations. We first form the constrained Lagrangian functional

$$\mathcal{L}_c(\boldsymbol{u}, \alpha, \boldsymbol{\beta}) := J(\boldsymbol{u}) + \langle \mathcal{K}(\boldsymbol{u}), \alpha \rangle_X + \boldsymbol{\beta} \cdot (\boldsymbol{C}(\boldsymbol{u}) - \boldsymbol{c}_0) \tag{S1}$$

where the function $\alpha : \Omega \to \mathbb{R}$ and the vector $\boldsymbol{\beta} = (\beta_1, \cdots, \beta_k) \in \mathbb{R}^k$ are the Lagrange multipliers. Taking the first variation of the constrained Lagrangian with respect to $\boldsymbol{u}$, we obtain

$$\begin{aligned}
\frac{\delta \mathcal{L}_c}{\delta \boldsymbol{u}}(\boldsymbol{v}) &:= \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \left[\mathcal{L}(\boldsymbol{u} + \varepsilon \boldsymbol{v}, \alpha, \boldsymbol{\beta}) - \mathcal{L}(\boldsymbol{u}, \alpha, \boldsymbol{\beta}) \right] \\
&= \mathrm{d}J(\boldsymbol{u}; \boldsymbol{v}) + \langle \mathcal{K}(\boldsymbol{v}), \alpha \rangle + \sum_{i=1}^{k} \beta_i \mathrm{d}C_i(\boldsymbol{u}; \boldsymbol{v}) \\
&= \langle J'(\boldsymbol{u}), \boldsymbol{v} \rangle + \langle \mathcal{K}^\dagger(\alpha), \boldsymbol{v} \rangle + \sum_{i=1}^{k} \beta_i \langle C_i'(\boldsymbol{u}), \boldsymbol{v} \rangle \\
&= \langle J'(\boldsymbol{u}) + \mathcal{K}^\dagger(\alpha) + \sum_{i=1}^{k} \beta_i C_i'(\boldsymbol{u}), \boldsymbol{v} \rangle
\end{aligned} \tag{S2}$$

Since the first variation $\delta \mathcal{L}_c / \delta \boldsymbol{u}$ must vanish for all $\boldsymbol{v}$, we obtain

$$J'(\boldsymbol{u}) + \mathcal{K}^\dagger(\alpha) + \sum_{i=1}^{k} \beta_i C_i'(\boldsymbol{u}) = \boldsymbol{0} \tag{S3}$$

Similarly, the first variations of the Lagrangian $\mathcal{L}_c$ with respect to the Lagrange multipliers α and $\boldsymbol{\beta}$ read

$$\frac{\delta \mathcal{L}_c}{\delta \alpha}(\tilde{\alpha}) = \langle \mathcal{K}(\boldsymbol{u}), \tilde{\alpha} \rangle_X, \quad \frac{\delta \mathcal{L}_c}{\delta \boldsymbol{\beta}} = \boldsymbol{C}(\boldsymbol{u}) - \boldsymbol{c}_0 \tag{S4}$$

Since they must vanish for all $\tilde{\alpha}$, we obtain the constraints $\mathcal{K}(\boldsymbol{u}) = 0$ and $\boldsymbol{C}(\boldsymbol{u}) = \boldsymbol{c}_0$.

section S2. The Navier-Stokes equation

S2.1 Preliminaries

Recall the Navier–Stokes equations

$$\partial_t \boldsymbol{u} = -\boldsymbol{u} \cdot \boldsymbol{\nabla} \boldsymbol{u} - \boldsymbol{\nabla} p + \nu \Delta \boldsymbol{u} + \boldsymbol{f} \tag{S5a}$$

$$\boldsymbol{\nabla} \cdot \boldsymbol{u} = 0 \tag{S5b}$$

with the Kolmogorov forcing $\boldsymbol{f}(\boldsymbol{x}) = \sin(k_f y)\boldsymbol{e}_1$ for some forcing wave number $\boldsymbol{k}_f = (0, k_f)$. In two dimensions, a divergence free velocity field $\boldsymbol{u} : \Omega \to \mathbb{R}^2$ admits the following Fourier series expansion

$$\boldsymbol{u}(\boldsymbol{x}, t) = \sum_{\boldsymbol{k} \in \mathbb{Z}^2} \frac{a(\boldsymbol{k}, t)}{k} \begin{pmatrix} k_2 \\ -k_1 \end{pmatrix} e^{\hat{i}\boldsymbol{k}\cdot\boldsymbol{x}} \tag{S6}$$

where $\boldsymbol{k} = (k_1, k_2)$, $k = |\boldsymbol{k}|$ and $\hat{i} = \sqrt{-1}$ (see Ref. (*33*)). Since the velocity field is real-valued, we have $a(-\boldsymbol{k}) = -\overline{a(\boldsymbol{k})}$.

For the Kolmogorov forcing, the energy input rate satisfies

$$I(\boldsymbol{u}(t)) = -\mathrm{Im}[a(\boldsymbol{k}_f, t)] = -r(\boldsymbol{k}_f, t) \sin\left(\phi(\boldsymbol{k}_f, t)\right) \tag{S7}$$

where Im denotes the imaginary part and $a(\boldsymbol{k}, t) = r(\boldsymbol{k}, t) \exp(\hat{i}\phi(\boldsymbol{k}, t))$ is the Fourier coefficient with phase $\phi(\boldsymbol{k}, t) \in (-\pi, \pi]$ and amplitude $r(\boldsymbol{k}, t) \in \mathbb{R}^+$. For simplicity, we may omit the dependence of these variables on time t. For reasons that will become clear in the next section, we refer to the Fourier mode $a(\boldsymbol{k}_f, t)$ as the mean flow.

Examining equation (S7), the energy input I may grow through two mechanisms:

(1) The phase $\phi(\boldsymbol{k}_f)$ approaching $-\pi/2$,

(2) The amplitude $r(\boldsymbol{k}_f)$ growing.

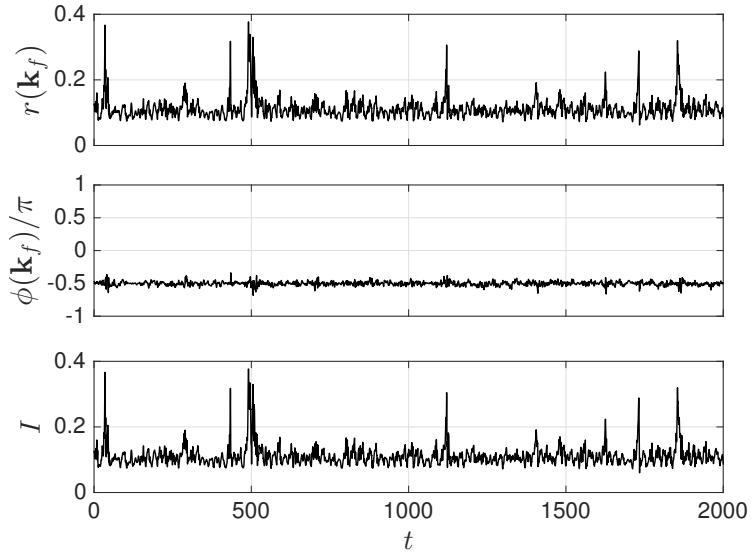

fig. S1. Evolution of the energy input solutions.
The evolution of the energy input $I = -\mathrm{Im}[a(\boldsymbol{k}_f)] = -r(\boldsymbol{k}_f)\sin[\phi(\boldsymbol{k}_f)]$, the phase $\phi(\boldsymbol{k}_f)$ of the mean flow and the amplitude $r(\boldsymbol{k}_f)$ of the mean flow along a typical trajectory of the Kolmogorov flow at $Re = 40$. Note that the phase of the external force is $-\pi/2$. The forcing wave number is $\boldsymbol{k}_f = (0, 4)$.

Noting that the phase of the external forcing is also $-\pi/2$, scenario (1) corresponds to an alignment between the phases of the external force and the mean flow $a(\boldsymbol{k}_f)$. It is therefore tempting to attribute the intermittent bursts of the energy input I to the intermittent alignments between the forcing $\boldsymbol{f}$ and the velocity field $\boldsymbol{u}$. This postulate, however, does not stand further scrutiny. Figure S1 shows the phase $\phi(\boldsymbol{k}_f, t)$ of the mean flow along a typical Kolmogorov trajectory $\boldsymbol{u}(t)$. This phase oscillates around $-\pi/2$ for all times. Note that $-\pi/2$ corresponds to perfect alignment between the mean flow and external forcing. Figure S1 also shows the evolution of the energy input I along the same trajectory. No positive correlation exists between intermittent growth of the mean flow energy I and the phase of the mean flow being $-\pi/2$. In fact, the phase $\phi(\boldsymbol{k}_f)$ seems to deviate from $-\pi/2$ during the bursts. Contrast this with the strong correlation between the growth of the energy input rate and the amplitude $r(\boldsymbol{k}_f)$ of the mean flow.
This observation shows that the intermittent energy input bursts are triggered through mechanism (2), that is the growth of the amplitude $r(\boldsymbol{k}_f)$. A similar observation is made at higher Reynolds numbers (not shown here). This growth of the mean flow amplitude, in turn, is possible through the internal transfer of energy via nonlinear terms as discussed below.

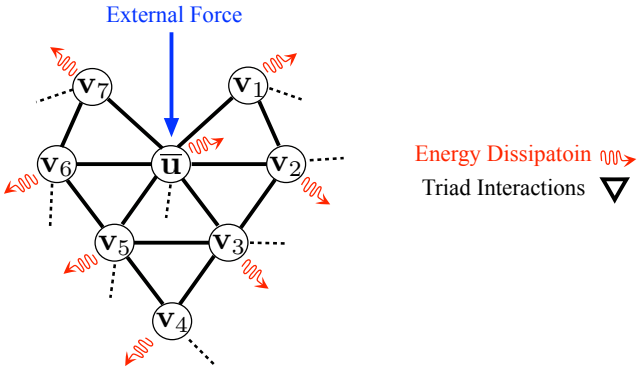

fig. S2. Triad interactions.
Schematic representation of the triad interactions of the Navier–Stokes equation.

S2.2 Nonlinear triad interactions

The velocity field $\boldsymbol{u}(\boldsymbol{x},t)$ can be written in the general Fourier-type expansion

$$\boldsymbol{u}(\boldsymbol{x},t) = \overline{\boldsymbol{u}}(\boldsymbol{x},t) + \sum_{j=1}^{\infty} \alpha_j(t)\boldsymbol{v}_j(\boldsymbol{x}) \tag{S8}$$

where $\overline{\boldsymbol{u}}$ is the statistical mean and $\{\boldsymbol{v}_j\}$ is a set of prescribed functions that form a complete basis for the function space X $(= L^2(\Omega))$. Under certain assumptions which are met by the Navier–Stokes equation (S5), the energy is injected into the mean flow $\overline{\boldsymbol{u}}$ by the external forcing $\boldsymbol{f}$ (see Refs. (*12, 13*)). The nonlinear term $\boldsymbol{u}\cdot\nabla\boldsymbol{u}$, coupling the mean flow and the modes $\boldsymbol{v}_j$, redistributes the injected energy to all modes $\boldsymbol{v}_j$. This nonlinear term conserves the total energy of the system. At the same time, each mode dissipates energy due to the viscous term $\nu\Delta\boldsymbol{u}$ (see fig. S2, for an illustration). A convenient choice of the basis $\{\boldsymbol{v}_j\}$ is problem dependent. Here, we choose the conventional Fourier basis as described in equation (S6). In case $\boldsymbol{f}$ is the Kolmogorov forcing, the symmetries of the system dictate $\overline{\boldsymbol{u}}(\boldsymbol{x},t) = \alpha_0(t)\boldsymbol{f}(\boldsymbol{x}) = \alpha_0(t)\sin(k_f y)\boldsymbol{e}_1$ (see, e.g., (*18, 34*)) .
In order to make the above statements more explicit, we write the Navier–Stokes equation in the Fourier space. Following (*39*), we have

$$\partial_t\hat{u}_i(\boldsymbol{k}) = -\hat{\imath}P_{ij}(\boldsymbol{k})\sum_{\boldsymbol{p}+\boldsymbol{q}=\boldsymbol{k}} q_m\hat{u}_m(\boldsymbol{p})\hat{u}_j(\boldsymbol{q}) - \nu k^2\hat{u}_i(\boldsymbol{k}) + \hat{f}_i(\boldsymbol{k}) \tag{S9}$$

where the hat signs denote the Fourier transform, $P_{ij}(\boldsymbol{k}) = \delta_{ij} - k_ik_j/k^2$ is the Leray projection onto the space of diverge-free vector fields and the convention of summation over repeated indices is used. Equation (S9) can be written more explicitly as

$$\partial_t\hat{u}_1(\boldsymbol{k}) = -\hat{\imath}\sum_{\boldsymbol{p}+\boldsymbol{q}=\boldsymbol{k}} q_m\hat{u}_m(\boldsymbol{p})\left[\left(1-\frac{k_1^2}{k^2}\right)\hat{u}_1(\boldsymbol{q}) - \frac{k_1k_2}{k^2}\hat{u}_2(\boldsymbol{q})\right] - \nu k^2\hat{u}_1(\boldsymbol{k}) + \hat{f}_1(\boldsymbol{k}) \tag{S10a}$$

$$\partial_t\hat{u}_2(\boldsymbol{k}) = -\hat{\imath}\sum_{\boldsymbol{p}+\boldsymbol{q}=\boldsymbol{k}} q_m\hat{u}_m(\boldsymbol{p})\left[-\frac{k_1k_2}{k^2}\hat{u}_1(\boldsymbol{q}) + \left(1-\frac{k_2^2}{k^2}\right)\hat{u}_2(\boldsymbol{q})\right] - \nu k^2\hat{u}_2(\boldsymbol{k}) + \hat{f}_2(\boldsymbol{k}) \tag{S10b}$$

Recall the Fourier expansion (S6) which implies $\hat{u}_1(\boldsymbol{k}) = k_2 a(\boldsymbol{k})/k$ and $\hat{u}_2(\boldsymbol{k}) = -k_1 a(\boldsymbol{k})/k$. Upon substitution in equation (S10b) and noting that

$$q_m \hat{u}_m(\boldsymbol{p}) = \frac{q_1 p_2 - q_2 p_1}{p} a(\boldsymbol{p})$$

and

$$\hat{f}_1(\boldsymbol{k}) = \frac{1}{2} e^{-\hat{i}\frac{\pi}{2}} \delta_{\boldsymbol{k},\boldsymbol{k}_f} + \frac{1}{2} e^{+\hat{i}\frac{\pi}{2}} \delta_{\boldsymbol{k},-\boldsymbol{k}_f}, \quad \hat{f}_2(\boldsymbol{k}) = 0$$

we obtain

$$\dot{a}(\boldsymbol{k}) = -\hat{i} \sum_{\boldsymbol{p}+\boldsymbol{q}=\boldsymbol{k}} \frac{(q_1 p_2 - q_2 p_1)(k_1 q_1 + k_2 q_2)}{p\, q\, k} a(\boldsymbol{p}) a(\boldsymbol{q}) - \nu k^2 a(\boldsymbol{k}) + \frac{1}{2} e^{-\hat{i}\frac{\pi}{2}} \left(\delta_{\boldsymbol{k},\boldsymbol{k}_f} + \delta_{\boldsymbol{k},-\boldsymbol{k}_f}\right) \tag{S11}$$

We rewrite the above equation more compactly

$$\dot{a}(\boldsymbol{k}) = \hat{i} \sum_{\boldsymbol{p}+\boldsymbol{q}=\boldsymbol{k}} \frac{\mu(\boldsymbol{p},\boldsymbol{q})(\boldsymbol{k}\cdot\boldsymbol{q})}{p\, q\, k} a(\boldsymbol{p}) a(\boldsymbol{q}) - \nu k^2 a(\boldsymbol{k}) + \frac{1}{2} e^{-\hat{i}\frac{\pi}{2}} \left(\delta_{\boldsymbol{k},\boldsymbol{k}_f} + \delta_{\boldsymbol{k},-\boldsymbol{k}_f}\right) \tag{S12}$$

where $\boldsymbol{k}\cdot\boldsymbol{q} = k_1 q_1 + k_2 q_2$ and $\mu(\boldsymbol{p},\boldsymbol{q}) := p_1 q_2 - p_2 q_1$ is the two-form measuring the surface area of the parallelogram with sides $\boldsymbol{p}$ and $\boldsymbol{q}$. Writing the modes in terms of their amplitudes and phases, $a(\boldsymbol{k}) = r(\boldsymbol{k}) \exp[\hat{i}\phi(\boldsymbol{k})]$, and using equation (S12), we obtain

$$\begin{aligned} \dot{r}(\boldsymbol{k}) =& \frac{1}{2} \cos\left[\frac{\pi}{2} + \phi(\boldsymbol{k})\right] \left(\delta_{\boldsymbol{k},\boldsymbol{k}_f} + \delta_{\boldsymbol{k},-\boldsymbol{k}_f}\right) - \nu k^2 r(\boldsymbol{k}) \\ &+ \sum_{\boldsymbol{p}+\boldsymbol{q}=\boldsymbol{k}} \frac{\mu(\boldsymbol{p},\boldsymbol{q})(\boldsymbol{k}\cdot\boldsymbol{q})}{p\, q\, k} r(\boldsymbol{p}) r(\boldsymbol{q}) \sin\left[\phi(\boldsymbol{k}) - \phi(\boldsymbol{p}) - \phi(\boldsymbol{q})\right] \end{aligned} \tag{S13a}$$

$$\begin{aligned} \dot{\phi}(\boldsymbol{k}) =& -\frac{1}{2} \frac{1}{r(\boldsymbol{k})} \sin\left[\frac{\pi}{2} + \phi(\boldsymbol{k})\right] \left(\delta_{\boldsymbol{k},\boldsymbol{k}_f} + \delta_{\boldsymbol{k},-\boldsymbol{k}_f}\right) \\ &+ \sum_{\boldsymbol{p}+\boldsymbol{q}=\boldsymbol{k}} \frac{\mu(\boldsymbol{p},\boldsymbol{q})(\boldsymbol{k}\cdot\boldsymbol{q})}{p\, q\, k} \frac{r(\boldsymbol{p}) r(\boldsymbol{q})}{r(\boldsymbol{k})} \cos\left[\phi(\boldsymbol{k}) - \phi(\boldsymbol{p}) - \phi(\boldsymbol{q})\right] \end{aligned} \tag{S13b}$$

Note that $a(-\boldsymbol{k}) = -\overline{a(\boldsymbol{k})}$ implies $\phi(-\boldsymbol{k}) = \pi - \phi(\boldsymbol{k})$.
We now focus on the amplitude of the mean flow $r(\boldsymbol{k}_f)$ (and its corresponding conjugate at $\boldsymbol{k} = -\boldsymbol{k}_f$). The negative definite term $-\nu k_f^2 r(\boldsymbol{k}_f)$ representing the dissipation acts to decrease the mean flow amplitude. This decay is counteracted by the external forcing $\frac{1}{2}\cos\left[\frac{\pi}{2} + \phi(\boldsymbol{k}_f)\right]$. Recall from fig. S1 that the phase $\phi(\boldsymbol{k}_f)$ oscillates around $-\pi/2$ for all times, implying $\cos\left[\frac{\pi}{2} + \phi(\boldsymbol{k}_f)\right] > 0$. The complications arise from the summation term in (S13a) which couples the mean flow to all other modes that form the wave vector triads, $\boldsymbol{p} + \boldsymbol{q} = \boldsymbol{k}_f$. The contribution from these other modes depends on the amplitudes, $r(\boldsymbol{p})$ and $r(\boldsymbol{q})$, and the relative phases $\phi(\boldsymbol{k}_f) - \phi(\boldsymbol{p}) - \phi(\boldsymbol{q})$. Even the modes that do not form a triad with $\boldsymbol{k}_f$, affect the mean flow amplitude indirectly through their coupling to the modes that do form a triad with $\boldsymbol{k}_f$ (see the schematic fig. S2).

S2.3 Derivation of Euler-Lagrange equation for Navier–Stokes

We first derive the functional J corresponding to the Navier–Stokes equation and the energy input rate I. For the function space X we set $X = L^2(\Omega)$ assuming that the state $\boldsymbol{u}$ belongs

to the space of square integrable vector fields. By definition, we have $J(\boldsymbol{u}) = dI(\boldsymbol{u}; \mathcal{N}(\boldsymbol{u}))$ which implies

$$\begin{aligned} J(\boldsymbol{u}) &= \frac{1}{|\Omega|}\int_\Omega (-\boldsymbol{u}\cdot\boldsymbol{\nabla}\boldsymbol{u} - \boldsymbol{\nabla}p + \nu\Delta\boldsymbol{u} + \boldsymbol{f})\cdot\boldsymbol{f}\,\mathrm{d}\boldsymbol{x} \\ &= \frac{1}{|\Omega|}\int_\Omega [\boldsymbol{u}\cdot(\boldsymbol{u}\cdot\boldsymbol{\nabla}\boldsymbol{f}) + \nu\boldsymbol{u}\cdot\Delta\boldsymbol{f}]\,\mathrm{d}\boldsymbol{x} + \frac{1}{|\Omega|}\|\boldsymbol{f}\|_2^2 \end{aligned} \tag{S14}$$

where we used integration by parts. The term involving the pressure p vanishes since the forcing is divergence free, $\boldsymbol{\nabla}\cdot\boldsymbol{f} = 0$. Since $\|\boldsymbol{f}\|_2$ is constant, we can safely omit the second term and let

$$J(\boldsymbol{u}) = \frac{1}{|\Omega|}\int_\Omega [\boldsymbol{u}\cdot(\boldsymbol{u}\cdot\boldsymbol{\nabla}\boldsymbol{f}) + \nu\boldsymbol{u}\cdot\Delta\boldsymbol{f}]\,\mathrm{d}\boldsymbol{x}$$

Next we compute the Gâteaux differential of J. By definition, we have

$$\begin{aligned} \mathrm{d}J(\boldsymbol{u};\boldsymbol{v}) &= \frac{1}{|\Omega|}\int_\Omega [\boldsymbol{v}\cdot(\boldsymbol{u}\cdot\boldsymbol{\nabla}\boldsymbol{f}) + \boldsymbol{u}\cdot(\boldsymbol{v}\cdot\boldsymbol{\nabla}\boldsymbol{f}) + \nu\,\boldsymbol{v}\cdot\Delta\boldsymbol{f}]\,\mathrm{d}\boldsymbol{x} \\ &= \frac{1}{|\Omega|}\int_\Omega \left[\left(\boldsymbol{\nabla}\boldsymbol{f} + \boldsymbol{\nabla}\boldsymbol{f}^\top\right)\boldsymbol{u} + \nu\Delta\boldsymbol{f}\right]\cdot\boldsymbol{v}\,\mathrm{d}\boldsymbol{x} \end{aligned}$$

On the other hand, by Riesz representation theorem, we have $\mathrm{d}J(\boldsymbol{u};\boldsymbol{v}) = \langle J'(\boldsymbol{u}), \boldsymbol{v}\rangle_{L^2}$ which implies

$$J'(\boldsymbol{u}) = \frac{1}{|\Omega|}\left[\left(\boldsymbol{\nabla}\boldsymbol{f} + \boldsymbol{\nabla}\boldsymbol{f}^\top\right)\boldsymbol{u} + \nu\Delta\boldsymbol{f}\right] \tag{S15}$$

Similarly, the Gâteaux differential of the constraint $C(\boldsymbol{u}) = \langle A\boldsymbol{u}, A\boldsymbol{u}\rangle_{L^2}/(2|\Omega|)$ is given by

$$\mathrm{d}C(\boldsymbol{u};\boldsymbol{v}) = \frac{1}{|\Omega|}\langle A\boldsymbol{u}, A\boldsymbol{v}\rangle_{L^2} = \frac{1}{|\Omega|}\langle A^\dagger A\boldsymbol{u}, \boldsymbol{v}\rangle_{L^2} = \langle C'(\boldsymbol{u}), \boldsymbol{v}\rangle_{L^2} \tag{S16}$$

implying $C'(\boldsymbol{u}) = A^\dagger A\boldsymbol{u}/|\Omega|$. Finally, the adjoint of the divergence operator, $\mathcal{K} = \boldsymbol{\nabla}\cdot$, with respect to the L^2 inner product is the gradient operator, $\mathcal{K}^\dagger = -\boldsymbol{\nabla}$. Substituting the above in the Euler–Lagrange equation (S3) and (S4), we obtain

$$\left(\boldsymbol{\nabla}\boldsymbol{f} + \boldsymbol{\nabla}\boldsymbol{f}^\top\right)\boldsymbol{u} + \nu\Delta\boldsymbol{f} - \boldsymbol{\nabla}\alpha + \beta A^\dagger A\boldsymbol{u} = \boldsymbol{0} \tag{S17a}$$

$$\boldsymbol{\nabla}\cdot\boldsymbol{u} = 0 \tag{S17b}$$

$$\frac{1}{|\Omega|}\int_\Omega \frac{|A\boldsymbol{u}|^2}{2}\mathrm{d}\boldsymbol{x} = c_0 \tag{S17c}$$

A few remarks about equations (S17) are in order: (i) The PDE (S17a) is inhomogeneous due to the term $\nu\Delta\boldsymbol{f} = -\nu k_f^2\sin(k_f y)\boldsymbol{e}_1$. (ii) The equations are nonlinear in the constraint (S17c). (iii) With the Kolmogorov forcing $\boldsymbol{f} = \sin(k_f y)\boldsymbol{e}_1$, the translations $\boldsymbol{u}(x,y) \mapsto \boldsymbol{u}(x+\ell, y)$, with $\ell \in \mathbb{R}$, are a symmetry transformation of equations (S17). That is, if $\boldsymbol{u}(x,y)$ solves (S17), so does $\tilde{\boldsymbol{u}}(x,y) = \boldsymbol{u}(x+\ell, y)$ for all $\ell \in \mathbb{R}$.

section S3. Newton iterations

In this section, we outline the Newton iterations for solving the system (S17). Define

$$\mathcal{F}(\boldsymbol{u},\alpha,\beta)=\begin{pmatrix}\left(\boldsymbol{\nabla f}+\boldsymbol{\nabla f}^{\top}\right)\boldsymbol{u}+\nu\Delta\boldsymbol{f}-\boldsymbol{\nabla}\alpha+\beta A^{\dagger}A\boldsymbol{u}\\ \boldsymbol{\nabla}\cdot\boldsymbol{u}\\ \int_{\Omega}|A(\boldsymbol{u})|^{2}\mathrm{d}\boldsymbol{x}-2|\Omega|c_{0}\end{pmatrix} \tag{S18}$$

The zeros of $\mathcal{F}$ coincide with the solutions of (S17). We find these zeros numerically using damped Newton iterations

$$\boldsymbol{u}_{n+1}=\boldsymbol{u}_{n}+\epsilon\tilde{\boldsymbol{u}},\quad \alpha_{n+1}=\alpha_{n}+\epsilon\tilde{\alpha},\quad \beta_{n+1}=\beta_{n}+\epsilon\tilde{\beta} \tag{S19}$$

At each iteration, the Newton direction $(\tilde{\boldsymbol{u}},\tilde{\alpha},\tilde{\beta})$ is obtained as the solution of the linear equation

$$\mathcal{L}(\boldsymbol{u}_{n},\alpha_{n},\beta_{n};\tilde{\boldsymbol{u}},\tilde{\alpha},\tilde{\beta})=-\mathcal{F}(\boldsymbol{u}_{n},\alpha_{n},\beta_{n}) \tag{S20}$$

where $\mathcal{L}(\boldsymbol{u},\alpha,\beta;\cdot,\cdot,\cdot)$ is the Gateaux differential of $\mathcal{F}$ at $(\boldsymbol{u},\alpha,\beta)$ and is given explicitly as

$$\mathcal{L}(\boldsymbol{u},\alpha,\beta;\tilde{\boldsymbol{u}},\tilde{\alpha},\tilde{\beta})=\begin{pmatrix}\left(\boldsymbol{\nabla f}+\boldsymbol{\nabla f}^{\top}\right)\tilde{\boldsymbol{u}}-\boldsymbol{\nabla}\tilde{\alpha}+\tilde{\beta}A^{\dagger}A\boldsymbol{u}+\beta A^{\dagger}A\tilde{\boldsymbol{u}}\\ \boldsymbol{\nabla}\cdot\tilde{\boldsymbol{u}}\\ 2\int_{\Omega}A(\boldsymbol{u})\cdot A(\tilde{\boldsymbol{u}})\mathrm{d}\boldsymbol{x}\end{pmatrix} \tag{S21}$$

The solution of the linear PDE (S20) is approximated by the generalized minimal residual (GMRES) algorithm (*40*). At each iteration, the step size $\epsilon\in(0,1]$ is adjusted to achieve maximal decrease in the error $\|\mathcal{F}(\boldsymbol{u}_{n+1},\alpha_{n+1},\beta_{n+1})\|_{L^2}$ (*41*). The standard Newton iterations correspond to $\epsilon=1$.

section S4. Sensitivity to parameters

Recall that the constraint $\int_{\Omega}|\boldsymbol{\nabla u}|^{2}\mathrm{d}\boldsymbol{x}/(2|\Omega|)=c_{0}$ enforces a constant energy dissipation rate. This constraint is motivated by the fact that, away from extreme bursts, the energy dissipation rate D exhibits small oscillations around its mean value. Nonetheless, D is not exactly constant, prompting the question whether the optimal solution is robust with respect to small perturbations to the constant c_0.

To examine this robustness, we have computed the optimal solution for a wide range of parameters c_0. We find that the optimal solution is in fact robust even with respect to relatively large variations in the parameter c_0. Figure S3, for instance, shows the optimal solution for three different values of c_0 at $Re=40$ and 100 (the results are similar for $Re=60$ and 80).

The insensitivity of the optimal solution with respect to the constant c_0 also implies that the equality constraint $\int_{\Omega}|\boldsymbol{\nabla u}|^{2}\mathrm{d}\boldsymbol{x}/(2|\Omega|)=c_{0}$ can be replaced with an inequality constraint of the form $c_1\le\int_{\Omega}|\boldsymbol{\nabla u}|^{2}\mathrm{d}\boldsymbol{x}/(2|\Omega|)\le c_2$. For a wide range of values for $c_2>c_0>c_1>0$, the optimal solutions corresponding to the two constraints will be similar.

section S5. Computing the probability of extreme events

We approximate the conditional PDFs using the following steps. For any two observables λ and γ, we assume that their joint probability density function $p_{\gamma,\lambda}$ exists such that

$$\mathcal{P}(\gamma_1\le\gamma\le\gamma_2,\lambda_1\le\lambda\le\lambda_2)=\int_{\gamma_1}^{\gamma_2}\int_{\lambda_1}^{\lambda_2}p_{\gamma,\lambda}(\gamma',\lambda')\mathrm{d}\lambda'\mathrm{d}\gamma' \tag{S22}$$

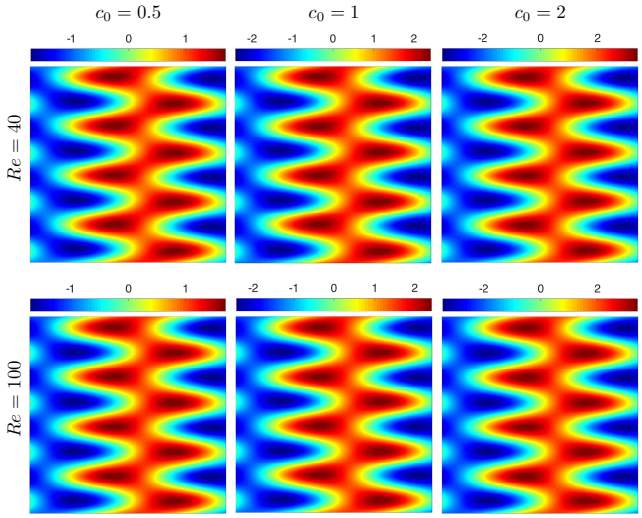

fig. S3. Sensitivity of the optimal solutions.
The global optimal solutions with $c_0 = 0.5$, 1 and 2 at $Re = 40$ and $Re = 100$.

Similarly, we also assume that the observable λ has a probability density p_λ. Once the PDF p_λ and the joint PDF $p_{\gamma,\lambda}$ are approximated using direct numerical simulations, the conditional PDF $p_{\gamma|\lambda}$ can be evaluated by the Bayesian formula

$$p_{\gamma|\lambda} = \frac{p_{\gamma,\lambda}}{p_\lambda}$$

Computation of the extreme event probability P_{ee} from the conditional probability is straightforward. Let γ_e denote the threshold such that $\gamma > \gamma_e$ denotes an extreme event. Then by definition, we have

$$P_{ee}(\lambda_0) = \mathcal{P}\left(\gamma > \gamma_e \middle| \lambda = \lambda_0\right) = \int_{\gamma_e}^{\infty} p_{\gamma|\lambda}(\gamma'|\lambda_0)\mathrm{d}\gamma' \tag{S23}$$

where γ' is a dummy integration variable. In the present paper, the variable λ is the indicator $|a(1,0)|$ and the variable γ is the future maximum of the energy dissipation rate, $\gamma(t) = D_m(t) = \max_{\tau\in[t+t_i,t+t_f]} D(\boldsymbol{u}(\tau))$. At each Reynolds number, the joint probability $p_{\gamma,\lambda}$ is approximated from the 100,000 computed data points on a 20×30 grid over the (γ, λ) plane.

section S6. Supporting computational results

In this section, we present the numerical results for Reynolds numbers $Re = 40$, 60, 80 and 100. The relevant parameters and variables are summarized in table S1. At each Reynolds number, the statistics are computed from long trajectory data of length 10,000 time units. The states (i.e. the velocity fields $\boldsymbol{u}$) are saved along these trajectories at every 0.1 time units, amounting to a combined 100,000 distinct states at each Reynolds number. Before recording any data, we evolved random initial conditions for 500 time units to ensure the decay of transients.

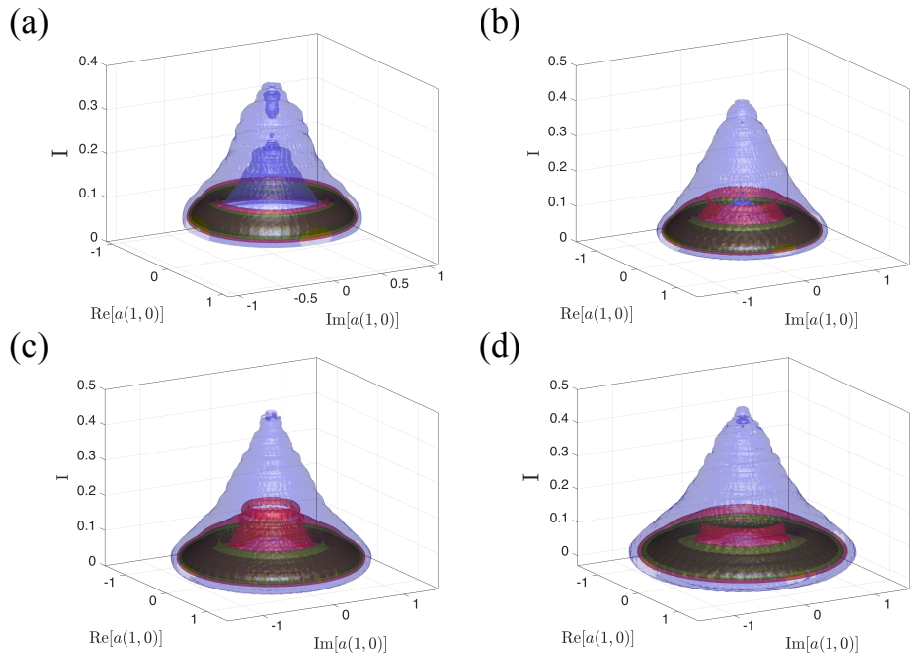

fig. S4. Joint PDFs for higher Reynolds numbers.
The joint PDF of the energy input rate I, Re$[a(1,0)]$ and Im$[a(1,0)]$ at $Re = 40$ (a), $Re = 60$ (b) $Re = 80$ (c) $Re = 100$ (d). The PDFs show that small values of $|a(1,0)|$ correlate strongly with the large values of the energy input rate I.

table S1. Simulation parameters.
Simulation parameters including the Reynolds number Re, the resolution $N \times N$, the mean $\mathbb{E}[D]$ and the standard deviation $\sqrt{\mathbb{E}[D^2] - \mathbb{E}[D]^2}$ of the energy dissipation rate D. The eddy turn-over time t_e and the prediction time t_i are reported in terms of non-dimensional time units. The percentage of hits, correct rejections and the false positives and negatives of the extreme event predictions are also reported. The rate of successful predictions (RSP) and the rate of successful rejections (RSR) are computed from formula (S24) and (S25).

Re	40	60	80	100
N	128	256	256	256
$\mathbb{E}[D]$	0.1168	0.1159	0.1010	0.0903
$\sqrt{\mathbb{E}[D^2] - \mathbb{E}[D]^2}$	0.0384	0.0465	0.0369	0.0295
t_e	0.46	0.38	0.35	0.33
t_i	1.0	1.0	1.0	1.0
t_f	2.0	2.0	2.0	1.5
Hits	5.60%	17.7%	15.3%	11.3%
Correct Rejection	93.3%	77.8%	78.5%	81.7%
False Negatives	0.26%	2.3%	3.5%	4.3%
False Positives	0.85%	2.1%	2.6%	2.6%
RSP	95.6%	88.4%	81.2%	72.3%
RSR	99.1%	97.4%	96.8%	96.9%

The Navier–Stokes equations are solved numerically with a standard pseudo-spectral code with $N \times N$ Fourier modes and 2/3 dealiasing (*42*) and a forth-order Runge–Kutta scheme for the temporal evolution. For $Re = 60$, 80 and 100, we use 256×256 Fourier modes to fully resolve the velocity fields. At $Re = 40$, however, this resolution is unnecessarily high and hence we use 128×128 modes.

Figure S4 shows the joint PDFs of the mode $a(1, 0)$ versus the energy input I. At all Reynolds numbers the joint PDFs have a cone shape reflecting the fact that small values of $|a(1, 0)|$ correspond to large values of the energy input rate.

As in $Re = 40$, we use the evolution of $|a(1, 0)|$ to predict an upcoming burst of the energy dissipation D. Figure S5 shows the computational results at higher Reynolds numbers. For $Re = 60$, 80 and 100, we set the threshold D_e for the extreme energy dissipation to be the mean plus one standard deviation of the energy dissipation. The measured mean $\mathbb{E}[D]$ and standard deviation $\sqrt{\mathbb{E}[D^2] - \mathbb{E}[D]^2}$ are reported in table S1. The corresponding extreme dissipation thresholds D_e are marked by vertical red dashed lines in the middle panel of fig. S5. The horizontal dashed line marks the critical λ_0 for which $P_{ee} = 0.5$, that is 50% probability of an upcoming extreme event.

We recall from the main body of the paper that the four quadrants in the conditional PDFs (middle column of fig. S5) correspond to:

(I) Correct rejection ($P_{ee} < 0.5$ and $D_m(t) < D_e$): Correct prediction of no upcoming extremes.

(II) False positives ($P_{ee} > 0.5$ but $D_e(t) < D_e$): The indicator predicts an upcoming extreme event but no extreme event actually takes place.

(III) Hit ($P_{ee} > 0.5$ and $D_m(t) > D_e$): Correct prediction of an upcoming extreme.

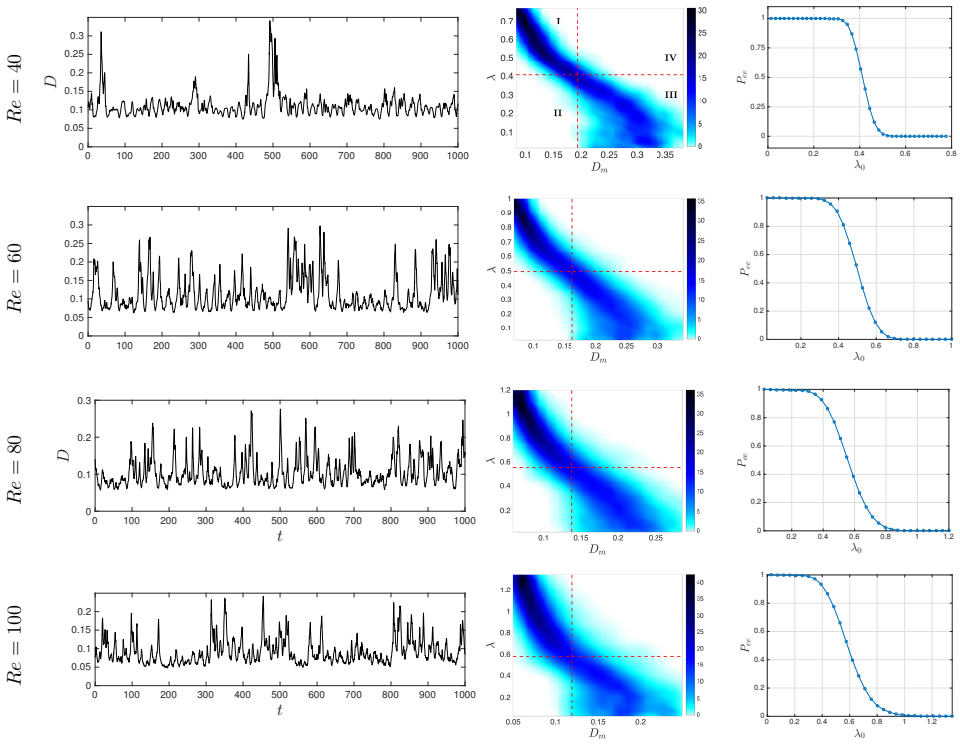

fig. S5. Prediction of intermittent bursts at higher Reynolds numbers.
The prediction of intermittent bursts of the energy dissipation rate D at Reynolds numbers $Re = 40, 60, 80$ and 100. Left column: Time series of the energy dissipation along a typical trajectory. Middle column: The conditional density $p(D_m|\lambda)$. Right column: Probability of extreme events P_{ee}.

(IV) False negatives ($P_{ee} < 0.5$ but $D_m(t) > D_e$): An extreme event takes place but the indicator fails to predict it.

Table S1 also shows the results of the extreme event prediction. In order to quantify the success of these predictions, we define

$$\text{Rate of Successful Predictions (RSP)} = \frac{\text{Hits}}{\text{Hits+False Negatives}} \tag{S24}$$

which measures the ratio of the number of extreme events that were successfully predicted to the total number of extreme events. Similarly, the quantity,

$$\text{Rate of Successful Rejections (RSR)} = \frac{\text{Correct Rejections}}{\text{Correct Rejections+False Positives}} \tag{S25}$$

measures the ratio of the number of non-extreme events that were correctly rejected to the total number of non-extreme events.

movie S1. The prediction of an extreme event in the Kolmogorov flow. The indicator value drops below the computed ethreshold (dashed line) indicating an upcoming burst in energy dissipation .