Predicting Rare Events via Large Deviations Theory: Rogue Waves and Motile Bacteria

Tobias Grafke, M. Cates, G. Dematteis, E. Vanden-Eijnden
Rare events matter

- Rare events are important if they are **extreme**
- Or **separation of scales** makes them common after all
- Underlying dynamics might be **very complex**, and analytical solutions are not available in most cases: Turbulence, Climate, chemical- or biological systems
- Direct numerical simulations (sampling) is **infeasible** because events are very rare
- Rare events are often **predictable**: Requires computational approaches based on LDT
The way rare events occur is often predictable — it is dominated by the least unlikely scenario — which is the essence of LDT.

Calculation of the least unlikely scenario (maximum likelihood pathway, MLP) reduces to a deterministic optimization problem.
The way rare events occur is often predictable — it is dominated by the least unlikely scenario — which is the essence of LDT.

Calculation of the least unlikely scenario (maximum likelihood pathway, MLP) reduces to a deterministic optimization problem.

Simple example: gradient systems (navigating a potential landscape), transitions between local energy minima happen through minimum energy paths (mountain pass transition).
The way rare events occur is often predictable — it is dominated by the least unlikely scenario — which is the essence of LDT.

Calculation of the least unlikely scenario (maximum likelihood pathway, MLP) reduces to a deterministic optimization problem.

Simple example: gradient systems (navigating a potential landscape), transitions between local energy minima happen through minimum energy paths (mountain pass transition).
The way rare events occur is often predictable — it is dominated by the *least unlikely* scenario — which is the essence of LDT.

Calculation of the least unlikely scenario (maximum likelihood pathway, MLP) reduces to a deterministic optimization problem.

Simple example: gradient systems (navigating a potential landscape), transitions between local energy minima happen through minimum energy paths (mountain pass transition).
The way rare events occur is often predictable — it is dominated by the least unlikely scenario — which is the essence of LDT.

Calculation of the least unlikely scenario (maximum likelihood pathway, MLP) reduces to a deterministic optimization problem.

Simple example: gradient systems (navigating a potential landscape), transitions between local energy minima happen through minimum energy paths (mountain pass transition).
A family of stochastic processes \( \{X^\varepsilon_t\}_{t \in [0,T]} \) with smallness-parameter \( \varepsilon \) (e.g. \( \varepsilon = 1/N \), or \( \varepsilon = k_B T \), etc) fulfils **large deviation principle**:

The probability that \( \{X^\varepsilon(t)\}_{t \in [0,T]} \) is close to a path \( \{\phi(t)\}_{t \in [0,T]} \) is

\[
P^\varepsilon \left\{ \sup_{0 \leq t \leq T} |X^\varepsilon(t) - \phi(t)| < \delta \right\} \asymp \exp \left( -\varepsilon^{-1} \mathcal{I}_T(\phi) \right) \text{ for } \varepsilon \to 0
\]

where \( \mathcal{I}_T(\phi) \) is the **rate function**.
Large deviation theory for stochastic processes

A family of stochastic processes \( \{X_t^\varepsilon\}_{t \in [0,T]} \) with smallness-parameter \( \varepsilon \) (e.g. \( \varepsilon = 1/N \), or \( \varepsilon = k_B T \), etc) fulfils **large deviation principle**: 

The probability that \( \{X^\varepsilon(t)\}_{t \in [0,T]} \) is close to a path \( \{\phi(t)\}_{t \in [0,T]} \) is

\[
P^\varepsilon \left\{ \sup_{0 \leq t \leq T} |X^\varepsilon(t) - \phi(t)| < \delta \right\} \asymp \exp \left( -\varepsilon^{-1} \mathcal{I}_T(\phi) \right) \text{ for } \varepsilon \to 0
\]

where \( \mathcal{I}_T(\phi) \) is the **rate function**.

The probability of hitting set \( A_z = \{x | F(x) = z\} \) is reduced to a **minimisation** problem

\[
P^\varepsilon \{X^\varepsilon(T) \in A_z | X^\varepsilon(0) = x \} \asymp \exp \left( -\varepsilon^{-1} \inf_{\phi: \phi(0)=x, F(\phi(T))=z} \mathcal{I}_T(\phi) \right)
\]
Large deviation theory for stochastic processes

A family of stochastic processes \( \{X_t^\varepsilon\}_{t \in [0,T]} \) with smallness-parameter \( \varepsilon \) (e.g. \( \varepsilon = 1/N \), or \( \varepsilon = k_B T \), etc) fulfils **large deviation principle**:

The probability that \( \{X^\varepsilon(t)\}_{t \in [0,T]} \) is close to a path \( \{\phi(t)\}_{t \in [0,T]} \) is

\[
  \mathcal{P}^\varepsilon \left\{ \sup_{0 \leq t \leq T} |X^\varepsilon(t) - \phi(t)| < \delta \right\} \asymp \exp \left(-\varepsilon^{-1} \mathcal{I}_T(\phi)\right) \quad \text{for } \varepsilon \to 0
\]

where \( \mathcal{I}_T(\phi) \) is the **rate function**.

The probability of hitting set \( A_z = \{x | F(x) = z\} \) is reduced to a **minimisation** problem

\[
  \mathcal{P}^\varepsilon \left\{ X^\varepsilon(T) \in A_z | X^\varepsilon(0) = x \right\} \asymp \exp \left(-\varepsilon^{-1} \inf_{\phi : \phi(0) = x, F(\phi(T)) = z} \mathcal{I}_T(\phi)\right)
\]

Here, \( \asymp \) is log-asymptotic equivalence, i.e.

\[
  \lim_{\varepsilon \to 0} \varepsilon \log \mathcal{P}^\varepsilon = -\inf_{\phi \in \Phi} \mathcal{I}_T(\phi) \quad \text{with e.g. } \Phi = \{\{x\}_{t \in [0,T]} | x(0) = x, F(x(T)) = z\}
\]
In particular consider SDE (diffusion) for $X_t^\varepsilon \in \mathbb{R}^n$, 

$$dX_t^\varepsilon = b(X_t^\varepsilon) \, dt + \sqrt{\varepsilon} \sigma dW_t,$$

with “drift” $b : \mathbb{R}^n \to \mathbb{R}^n$ and “noise” with covariance $\chi = \sigma \sigma^T$, we have 

$$I_T(\phi) = \frac{1}{2} \int_0^T |\dot{\phi} - b(\phi)|^2 \chi \, dt = \int_0^T L(\phi, \dot{\phi}) \, dt,$$

for Lagrangian $L(\phi, \dot{\phi})$ (follows by contraction from Schilder’s theorem).

We are interested in 

$$\phi^* = \argmin_{\phi \in \mathcal{C}} \int_0^T L(\phi, \dot{\phi}) \, dt$$

which is the maximum likelihood pathway (MLP).
Physicists approach: Path integral formalism

Consider

\[ \dot{x} = b(x) + \eta \]

with white noise \( \eta \) with covariance

\[ \langle \eta_i(t) \eta_j(t') \rangle = \epsilon \chi_{ij} \delta(t - t') \]
Consider\[\dot{x} = b(x) + \eta\]

with white noise \(\eta\) with covariance
\[\langle \eta_i(t)\eta_j(t') \rangle = \epsilon \chi_{ij} \delta(t-t')\]

then
\[P(\{\eta\}) \sim \int D[\eta] e^{-\frac{1}{2\epsilon} \int \eta \chi^{-1} \eta \, dt}\]

but \(x = x[\eta]\), with \(\eta = \dot{x} - b(x)\), so that (ignoring Jacobian)
\[P(\{x\}) \sim \int D[x] e^{-\frac{1}{2\epsilon} \int (\dot{x} - b(x))^2 \chi \, dt} \sim \int D[x] e^{-\frac{1}{\epsilon} I_T(x)}\]
Physicists approach: Path integral formalism

Consider

\[ \dot{x} = b(x) + \eta \]

with white noise \( \eta \) with covariance

\[ \langle \eta_i(t)\eta_j(t') \rangle = \epsilon \chi_{ij} \delta(t - t') \]

then

\[ \mathcal{P}(\{\eta\}) \sim \int D[\eta] e^{-\frac{1}{2\epsilon} \int \eta \chi^{-1} \eta \, dt} \]

but \( x = x[\eta] \), with \( \eta = \dot{x} - b(x) \), so that (ignoring Jacobian)

\[ \mathcal{P}(\{x\}) \sim \int D[x] e^{-\frac{1}{2\epsilon} \int |\dot{x} - b(x)|^2 \, dt} \sim \int D[x] e^{-\frac{1}{\epsilon} I_T(x)} \]

Approximate path integral for \( \epsilon \to 0 \) via saddle point approximation,

\[ \frac{\delta I}{\delta \phi^*} = 0, \quad \text{Instanton, semi-classical trajectory} \]
Consider \( \dot{x} = b(x) + \eta \) with white noise \( \eta \) with covariance
\[
\langle \eta_i(t) \eta_j(t') \rangle = \epsilon \chi_{ij} \delta(t - t')
\]
then
\[
P(\{\eta\}) \sim \int \mathcal{D}[\eta] e^{-\frac{1}{2\epsilon} \int \eta \chi^{-1} \eta dt}
\]
but \( x = x[\eta] \), with \( \eta = \dot{x} - b(x) \), so that (ignoring Jacobian)
\[
P(\{x\}) \sim \int \mathcal{D}[x] e^{-\frac{1}{2\epsilon} \int |\dot{x} - b(x)|^2 dt} \sim \int \mathcal{D}[x] e^{-\frac{1}{\epsilon} I_T(x)}
\]
Approximate path integral for \( \epsilon \to 0 \) via **saddle point approximation**,\[
\frac{\delta I}{\delta \phi^*} = 0, \quad \text{(Instanton, semi-classical trajectory)}
\]
Rate function \( \leftrightarrow \) Action, MLP \( \leftrightarrow \) Instanton, LDP \( \leftrightarrow \) Hamiltonian principle
Maximum likelyhood pathway and rare events

Main problem

Find the **maximum likelyhood pathway** (MLP) $\phi^*$ realizing an event, i.e. such that

$$I_T(\phi^*) = \inf_{\phi \in \mathcal{C}} I_T(\phi)$$

where $\mathcal{C}$ is the set of trajectories that fulfil our constraints.
Maximum likelyhood pathway and rare events

Main problem

Find the **maximum likelyhood pathway** (MLP) \( \phi^* \) realizing an event, i.e. such that

\[
I_T(\phi^*) = \inf_{\phi \in \mathcal{C}} I_T(\phi)
\]

where \( \mathcal{C} \) is the set of trajectories that fulfil our constraints.

Knowledge of the optimal trajectory gives us

1. **Probability** of event, \( \mathcal{P} \sim \exp \left(-\epsilon^{-1} I_T(\phi^*)\right) \)

2. Most likely **occurence**, \( \phi^* \) itself (allows for prediction, exploring causes, etc.)

3. Most effective way to force event (optimal control), **optimal fluctuation**
Example: Ornstein-Uhlenbeck

**Ornstein-Uhlenbeck process**

\[ du = b(u)\, dt + dW, \quad b(u) = -\gamma u, \quad \gamma > 0 \]

Consider extreme events with \( u(T) = z \) (so \( F(u) = u(T) \)).

The **instanton** is

\[ u^*(t) = z e^{\gamma(t-T)} \left( \frac{1 - e^{-2\gamma t}}{1 - e^{-2\gamma T}} \right), \]

obtained from **constrained optimization**

\[
\inf_{\{u_t\} \in \mathcal{U}_z} \mathcal{I}_T(z) = \inf_{\{u_t\} \in \mathcal{U}_z} \frac{1}{2} \int_0^T |\dot{u} + \gamma u|^2 \, dt
\]

over the set

\[ \mathcal{U}_z = \left\{ \{u_t\} \mid F(u_T) = z \right\} \]
Example: Ornstein-Uhlenbeck

**Ornstein-Uhlenbeck** process

\[ du = b(u) \, dt + dW \, , \quad b(u) = -\gamma u \, , \quad \gamma > 0 \]

Consider extreme events with \( u(T) = z \) (so \( F(u) = u(T) \)).

The **instanton** is

\[ u^*(t) = z e^{\gamma(t-T)} \left( \frac{1 - e^{-2\gamma t}}{1 - e^{-2\gamma T}} \right) , \]

obtained from **constrained optimization**

\[
\inf_{\{u_t\} \in U_z} \mathcal{I}_T(z) = \inf_{\{u_t\} \in U_z} \int_0^T |\dot{u} + \gamma u|^2 \, dt
\]

over the set

\[ U_z = \left\{ \{u_t\} \mid F(u_T) = z \right\} \]
Example: Ornstein-Uhlenbeck

**Ornstein-Uhlenbeck** process

\[ du = b(u) \, dt + dW, \quad b(u) = -\gamma u, \quad \gamma > 0 \]

Consider extreme events with \( u(T) = z \) (so \( F(u) = u(T) \)).

The **instanton** is

\[ u^*(t) = ze^{\gamma(t-T)} \left( \frac{1 - e^{-2\gamma t}}{1 - e^{-2\gamma T}} \right), \]

obtained from **constrained optimization**

\[ \inf_{\{u_t\} \in \mathcal{U}_z} \mathcal{I}_T(z) = \inf_{\{u_t\} \in \mathcal{U}_z} \frac{1}{2} \int_0^T |\dot{u} + \gamma u|^2 \, dt \]

over the set

\[ \mathcal{U}_z = \left\{ \{u_t\} \mid F(u_T) = z \right\} \]
Reversible systems and gradient flows

Special case: Systems in **detailed balance**. For example,

\[ dX_t^\epsilon = -\nabla U(X_t^\epsilon) \, dt + \sqrt{2\epsilon} \, dW_t \]

Then

\[ I_T(\phi) = \frac{1}{4} \int_0^T |\dot{\phi} + \nabla U|^2 \, dt \]

is minimized either by \( \dot{\phi} = -\nabla U \) (“sliding” down-hill) or

\[ I_T(\phi) = \frac{1}{4} \int_0^T |\dot{\phi} + \nabla U|^2 \, dt = \frac{1}{4} \int_0^T |\dot{\phi} - \nabla U|^2 \, dt + \int_0^T \nabla U \cdot \dot{\phi} \, dt \]

\[ = U(\phi_{\text{end}}) - U(\phi_{\text{start}}) \quad \text{if we choose} \quad \dot{\phi} = \nabla U \]

which is the **time-reversed** down-hill path. Easy algorithms exist*.

---

Example: Pendulum

Consider \textbf{damped pendulum}

\[
\begin{align*}
    dx &= v \, dt + \sigma \, dW_x, \\
    dv &= -\sin(x) \, dt - \gamma v \, dt + \sigma \, dW_v
\end{align*}
\]
Example: Pendulum

Consider damped pendulum

\[
\begin{align*}
    dx &= v \, dt + \sigma \, dW_x, \\
    dv &= -\sin(x) \, dt - \gamma v \, dt + \sigma \, dW_v
\end{align*}
\]
Consider **damped pendulum**

\[
\begin{align*}
    dx &= v \, dt + \sigma \, dW_x, \\
    dv &= -\sin(x) \, dt - \gamma v \, dt + \sigma \, dW_v
\end{align*}
\]
Example: Pendulum

Consider damped pendulum

\[
\begin{cases}
    dx = v \, dt + \sigma \, dW_x, \\
    dv = -\sin(x) \, dt - \gamma v \, dt + \sigma \, dW_v
\end{cases}
\]
Hamiltonian formalism

Main problem

Find the **maximum likelyhood pathway** (MLP) $\phi^*$ realizing an event, i.e. such that

$$I_T(\phi^*) = \inf_{\phi \in C} I_T(\phi)$$

where $C$ is the set of trajectories that fulfil our constraints.

Obtained through direct **numerical minimisation**,
Hamiltonian formalism

Main problem

Find the **maximum likelyhood pathway** (MLP) $\phi^*$ realizing an event, i.e. such that

$$I_T(\phi^*) = \inf_{\phi \in \mathcal{C}} I_T(\phi)$$

where $\mathcal{C}$ is the set of trajectories that fulfil our constraints.

Obtained through direct **numerical minimisation**, or through **Hamiltonian**

$$H(x, p) = \sup_y \{ y p - L(x, y) \} \overset{\text{FW}}{=} b(x)p + \frac{1}{2}p \chi p$$

so that $(\phi^*, \theta^*)$ fulfil **equations of motion**

$$\begin{cases}
\dot{\phi} = \nabla_\theta H(\phi, \theta) & \overset{\text{FW}}{\Rightarrow} \dot{\phi} = b(\phi) + \chi \theta \\
\dot{\theta} = -\nabla_\phi H(\phi, \theta) & \overset{\text{FW}}{\Rightarrow} \dot{\theta} = -\nabla b(\phi)^T \theta
\end{cases}$$
Finding the minimizer

Algorithm†‡:

\begin{align*}
\dot{\phi} &= b(\phi) + \chi \theta \\
\dot{\theta} &= -\nabla b(\phi)^T \theta
\end{align*}

Advantages:

- Fits with the boundary conditions
- Simple time-integration scheme applicable (Runge-Kutta)
- No higher derivatives of $H(\phi, \theta)$
- This is essentially computing the gradient via the adjoint formalism

Finding the minimizer

Algorithm\(^\S\), \(^\¶\):

\[
t = 0 \quad t = T
\]

Problem for PDEs: Memory, e.g. 2D

\[
2 \times 1024 \times 1024 \times 10^4 \approx 10^{10}
\]


Finding the minimizer

Algorithm:\n\[ t = 0 \quad \text{to} \quad t = T \]

Problem for PDEs: Memory, e.g. 2D

- For \( \theta \): Store only \( \chi\theta \) instead of \( \theta \), \( 1024^2 \rightarrow 64^2 \)

---


Finding the minimizer

Algorithm\textsuperscript{§,¶}:
\[ t = 0 \quad \text{to} \quad t = T \]

Problem for PDEs: **Memory**, e.g. 2D

\[ \frac{2}{(\phi, \theta)} \times 1024 \times 1024 \times 10^4 \approx 10^{10} \]

- For \( \theta \): Store only \( \chi \theta \) instead of \( \theta \), \( 1024^2 \rightarrow 64^2 \)
- For \( \phi \): Recursive solution in \( \phi \), \( O(N_t) \rightarrow O(\log N_t) \)


Finding the minimizer

Algorithm\textsuperscript{§,¶}:
\[
t = 0 \quad t = T
\]

Problem for PDEs: Memory, e.g. 2D

\[
\frac{2}{(\phi, \theta)} \times 1024 \times 1024 \times 10^4 \approx 10^{10}
\]

- For \( \theta \): Store only \( \chi \theta \) instead of \( \theta \), \( 1024^2 \rightarrow 64^2 \)
- For \( \phi \): Recursive solution in \( \phi \), \( O(N_t) \rightarrow O(\log N_t) \)


Finding the minimizer

Algorithm:\n\[ t = 0 \quad \text{to} \quad t = T \]

Problem for PDEs: **Memory**, e.g. 2D

- For \( \theta \): Store only \( x_\theta \) instead of \( \theta \), \( 1024^2 \to 64^2 \)
- For \( \phi \): Recursive solution in \( \phi \), \( \mathcal{O}(N_t) \to \mathcal{O}(\log N_t) \)

\(2 \times 1024 \times 1024 \times 10^4 \approx 10^{10}\)

---


Finding the minimizer

Algorithm:\n\[ t = 0 \quad \text{to} \quad t = T \]

\[ \phi \quad \text{and} \quad \theta \]

Problem for PDEs: Memory, e.g. 2D

\[ 2 \times 1024 \times 1024 \times 10^4 \approx 10^{10} \]

- For \( \theta \): Store only \( \chi \theta \) instead of \( \theta \), \( 1024^2 \rightarrow 64^2 \)
- For \( \phi \): Recursive solution in \( \phi \), \( O(N_t) \rightarrow O(\log N_t) \)


Finding the minimizer

Algorithm\textsuperscript{§,¶}: 
\[ t = 0 \quad \text{to} \quad t = T \]

\[ \phi \quad \theta \]

Problem for PDEs: \textbf{Memory}, e.g. 2D

\[ 2 \times 1024 \times 1024 \times 10^4 \approx 10^{10} \]

- For $\theta$: Store only $\chi$ instead of $\theta$, $1024^2 \rightarrow 64^2$
- For $\phi$: Recursive solution in $\phi$, $O(N_t) \rightarrow O(\log N_t)$


Finding the minimizer

Algorithm\(^\S,\¶\):
\[
t = 0 \quad \rightarrow \quad t = T
\]

\[
\phi \quad \theta
\]

Problem for PDEs: **Memory**, e.g. 2D

\[
2 \times 1024 \times 1024 \times 10^4 \approx 10^{10}
\]

- For \(\theta\): Store only \(\chi \theta\) instead of \(\theta\), \(1024^2 \rightarrow 64^2\)
- For \(\phi\): Recursive solution in \(\phi\), \(\mathcal{O}(N_t) \rightarrow \mathcal{O}(\log N_t)\)


Finding the minimizer

Algorithm:\n
\[ t = 0 \rightarrow t = T \]

Problem for PDEs: **Memory**, e.g. 2D

\[ \begin{array}{c}
\phi \\
\theta
\end{array} \]

\[ 2 \times 1024 \times 1024 \times 10^4 \approx 10^{10} \]

- For θ: Store only \( \chi_\theta \) instead of \( \theta \), \( 1024^2 \rightarrow 64^2 \)
- For \( \phi \): Recursive solution in \( \phi \), \( O(N_t) \rightarrow O(\log N_t) \)

---


Finding the minimizer

Algorithm:\n\begin{align*}
\phi_{\theta_{t=0}} &= \phi_{t=T} \\
\theta_{t=0} &= \theta_{t=T}
\end{align*}

Problem for PDEs: Memory, e.g. 2D

\[2 \times 1024 \times 1024 \times 10^4 \approx 10^{10}\]

- For \(\theta\): Store only \(\chi_{\theta}\) instead of \(\theta\), \(1024^2 \rightarrow 64^2\)
- For \(\phi\): Recursive solution in \(\phi\), \(O(N_t) \rightarrow O(\log N_t)\)

\[\text{\textsuperscript{5}}\text{Antonio Celani, Massimo Cencini, and Alain Noullez. “Going forth and back in time: a fast and parsimonious algorithm for mixed initial/final-value problems”. In: } \textit{Physica D: Nonlinear Phenomena} \textit{195.3} (2004), pp. 283–291.\]

Finding the minimizer

Algorithm\textsuperscript{§,¶}:
\[ t = 0 \quad \rightarrow \quad t = T \]

\[ \phi \]
\[ \theta \]

Problem for PDEs: Memory, e.g. 2D
\[ 2 \times 1024 \times 1024 \times 10^4 \approx 10^{10} \]

- For \( \theta \): Store only \( \chi \theta \) instead of \( \theta \), \( 1024^2 \rightarrow 64^2 \)
- For \( \phi \): Recursive solution in \( \phi \), \( O(N_t) \rightarrow O(\log N_t) \)


Finding the minimizer

Algorithm\textsuperscript{§,¶}:

\begin{align*}
t &= 0 \\
\phi \theta_t &= 0 \\
\theta &= T
\end{align*}

\[ t = 0 \quad \text{to} \quad t = T \]

Problem for PDEs: \textbf{Memory}, e.g. 2D

\[
2 \times 1024 \times 1024 \times 10^4 \approx 10^{10}
\]

- For \( \theta \): Store only \( \chi \theta \) instead of \( \theta \), \( 1024^2 \rightarrow 64^2 \)
- For \( \phi \): Recursive solution in \( \phi \), \( \mathcal{O}(N_t) \rightarrow \mathcal{O}(\log N_t) \)


Finding the minimizer

Algorithm\textsuperscript{§, ¶}:

\[ t = 0 \quad \text{to} \quad t = T \]

\[ \phi \quad \theta \]

Problem for PDEs: Memory, e.g. 2D

\[ \frac{2}{\phi, \theta} \times 1024 \times 1024 \times 10^4 \approx 10^{10} \]

- For $\theta$: Store only $\chi \theta$ instead of $\theta$, $1024^2 \rightarrow 64^2$
- For $\phi$: Recursive solution in $\phi$, $O(N_t) \rightarrow O(\log N_t)$


Finding the minimizer

Algorithm\textsuperscript{§,¶}:

\[ t = 0 \]

\[ t = T \]

\[ \phi \]

\[ \theta \]

Problem for PDEs: Memory, e.g. 2D

\[ \frac{2}{\phi, \theta} \times 1024 \times 1024 \times 10^4 \approx 10^{10} \]

- For \( \theta \): Store only \( \chi \theta \) instead of \( \theta \), \( 1024^2 \rightarrow 64^2 \)
- For \( \phi \): Recursive solution in \( \phi \), \( O(N_t) \rightarrow O(\log N_t) \)


Finding the minimizer

Algorithm:\n\[ t = 0 \quad t = T \]

\[ \phi \quad \theta \]

Problem for PDEs: Memory, e.g. 2D

\[ 2 \times 1024 \times 1024 \times 10^4 \approx 10^{10} \]

- For \( \theta \): Store only \( \chi \theta \) instead of \( \theta \), \( 1024^2 \rightarrow 64^2 \)
- For \( \phi \): Recursive solution in \( \phi \), \( O(N_t) \rightarrow O(\log N_t) \)

Finding the minimizer

Algorithm $^8, ^9$:

\[ t = 0 \quad \text{to} \quad t = T \]

\[ \phi \quad \theta \]

Problem for PDEs: Memory, e.g. 2D

\[ 2 \times 1024 \times 1024 \times 10^4 \approx 10^{10} \]

- For $\theta$: Store only $\chi \theta$ instead of $\theta$, $1024^2 \rightarrow 64^2$
- For $\phi$: Recursive solution in $\phi$, $O(N_t) \rightarrow O(\log N_t)$

---


Finding the minimizer

Algorithm:
\[ t = 0 \quad \text{to} \quad t = T \]

\[ \phi \theta \]

Problem for PDEs: Memory, e.g. 2D

\[ 2 \times 1024 \times 1024 \times 10^4 \approx 10^{10} \]

- For \( \theta \): Store only \( \chi \theta \) instead of \( \theta \), \( 1024^2 \rightarrow 64^2 \)
- For \( \phi \): Recursive solution in \( \phi \), \( \mathcal{O}(N_t) \rightarrow \mathcal{O}(\log N_t) \)


Finding the minimizer

Algorithm:

\[ t = 0 \quad \text{to} \quad t = T \]

Problem for PDEs: **Memory**, e.g. 2D

- For \( \theta \): Store only \( \chi \theta \) instead of \( \theta \), \( 1024^2 \rightarrow 64^2 \)
- For \( \phi \): Recursive solution in \( \phi \), \( O(N_t) \rightarrow O(\log N_t) \)

---

\( \text{§ Antonio Celani, Massimo Cencini, and Alain Noullez. “Going forth and back in time: a fast and parsimonious algorithm for mixed initial/final-value problems”. In: Physica D: Nonlinear Phenomena 195.3 (2004), pp. 283–291.} \)

\( \text{¶ Tobias Grafke, Rainer Grauer, and Stephan Schindel. “Efficient Computation of Instantons for Multi-Dimensional Turbulent Flows with Large Scale Forcing”. In: Communications in Computational Physics 18.03 (Sept. 2015), pp. 577–592. ISSN: 1991-7120. DOI: 10.4208/cicp.031214.200415a.} \)
Finding the minimizer

Algorithm\textsuperscript{§,¶}:

\[ t = 0 \quad \text{to} \quad t = T \]

\[ \phi \quad \theta \]

Problem for PDEs: Memory, e.g. 2D

\[ 2 \times 1024 \times 1024 \times 10^4 \approx 10^{10} \]

- For \( \theta \): Store only \( \chi\theta \) instead of \( \theta \), \( 1024^2 \rightarrow 64^2 \)
- For \( \phi \): Recursive solution in \( \phi \), \( \mathcal{O}(N_t) \rightarrow \mathcal{O}(\log N_t) \)


Finding the minimizer

Algorithm: $t = 0$ $t = T$

Problem for PDEs: **Memory**, e.g. 2D

- For $\theta$: Store only $\chi \theta$ instead of $\theta$, $1024^2 \rightarrow 64^2$
- For $\phi$: Recursive solution in $\phi$, $\mathcal{O}(N_t) \rightarrow \mathcal{O}(\log N_t)$

---


Finding the minimizer

Algorithm:\[ t = 0 \quad \text{to} \quad t = T \]

Problem for PDEs: **Memory**, e.g. 2D

- For \( \theta \): Store only \( \chi \theta \) instead of \( \theta \), \( 1024^2 \rightarrow 64^2 \)
- For \( \phi \): Recursive solution in \( \phi \), \( O(N_t) \rightarrow O(\log N_t) \)

---

\[
2 \times 1024 \times 1024 \times 10^4 \approx 10^{10}
\]

---


Finding the minimizer

Algorithm:\n\[
\phi_t = 0, \quad t = 0 \quad \text{to} \quad t = T
\]

Problem for PDEs: Memory, e.g. 2D

\[
2 \times 1024 \times 1024 \times 10^4 \approx 10^{10}
\]

- For \( \theta \): Store only \( \chi_\theta \) instead of \( \theta \), \( 1024^2 \rightarrow 64^2 \)
- For \( \phi \): Recursive solution in \( \phi \), \( \mathcal{O}(N_t) \rightarrow \mathcal{O}(\log N_t) \)
- This is known as “checkpointing” in PDE optimization
- Additionally, bi-orthogonal wavelets to store fields

---


Finding the minimizer

![Graph showing memory usage vs. resolution for different optimization strategies.](image)

- **Memory Usage (in MB)**
  - **no optimization**
  - **\(\chi_{\theta}\)**
  - **recursive**
  - **\(\chi_{\theta} \& \) recursive**

- **Resolution**
  - 128
  - 256
  - 512
  - 1024
  - 2048
  - 4096
  - 8192

Tobias Grafke
Predicting Rare Events via Large Deviations Theory
Application: Extreme gradients in Burgers equation

Evolution of Burgers shocks:

\[ u_t + uu_x - \nu u_{xx} = \eta \]

with

\[ \langle \eta \eta' \rangle = \delta(t - t') \chi(x - x') \]

Compute

\[ \mathcal{P} \{ u_x(0, 0) > z | u(x, -T) = 0 \} \]

**Question:** What is the most likely evolution from \( u(x) = 0 \) at \( t = -\infty \), such that at the end (i.e. \( t = 0 \)) we have a high gradient in the origin \( u_x(x = 0, t = 0) = z \) (shock)?

Grafke, Grauer, Schäfer, and Vanden-Eijnden 2015
Evolution of Burgers shocks:

\[ u_t + uu_x - \nu u_{xx} = \eta \]

with

\[ \langle \eta \eta' \rangle = \delta(t - t') \chi(x - x') \]

\textbf{Question}: What is the most likely evolution from \( u(x) = 0 \) at \( t = -\infty \), such that at the end (i.e. \( t = 0 \)) we have a high gradient in the origin \( u_x(x = 0, t = 0) = z \) (shock)?

\[ \mathcal{P} \{ u_x(0, 0) > z | u(x, -T) = 0 \} \]

Grafke, Grauer, Schäfer, and Vanden-Eijnden 2015
Application: Extreme gradients in Burgers turbulence

\[ u(x) \]

-10 -8 -6 -4 -2 0 2 4 6 8 10

Grafke, Grauer, and Schäfer 2013

\[ u(x) \]
Application: Extreme gradients in Burgers turbulence

\[ H(u, \theta) = \int (\theta \cdot (u \cdot \nabla u - \nu \nabla^2 u) + \frac{1}{2} \theta \chi \ast \theta) \, dx \]
Application: Active matter phase separation

Bacteria show complex collective behavior

- have **active propulsion**, i.e. a free-swimming (planktonic) stage
- are able to sense their environment through **quorum sensing**
- stick to surfaces in **biofilms**

![Image of bacteria](image1)

E. Coli: active propulsion & biofilms

![Image of biofilm](image2)
Bacteria show complex collective behavior

- have **active propulsion**, i.e. a free-swimming (planktonic) stage
- are able to sense their environment through **quorum sensing**
- stick to surfaces in **biofilms**

Model bacteria as $N$ **agents** with

- **active** Brownian motion, i.e. velocity vector diffuses on a sphere,
- **density-dependent** diffusion constant,
- and **birth/death**
Application: Active matter phase separation

Bacteria show complex collective behavior

- have **active propulsion**, i.e. a free-swimming (planktonic) stage
- are able to sense their environment through **quorum sensing**
- stick to surfaces in **biofilms**

Model bacteria as $N$ agents with

- **active** Brownian motion, i.e. velocity vector diffuses on a sphere,
- **density dependent** diffusion constant,
- and **birth/death**

Then take LDT for $N \to \infty$

$$H(\rho, \theta) = \int \left( \theta \partial_x (D_e(\rho) \partial_x \rho - \rho D(\rho) \partial_x (\delta^2 \partial_x^2 \rho + \theta)) + \alpha \rho (e^{\theta} - 1) + \alpha \rho^2 / \rho_0 (e^{-\theta} - 1) \right) dx$$

E. Coli: active propulsion & biofilms
Application: Active matter phase separation

Complex collective behaviour for simple active agents:

**Propulsion and Reproduction**

- When $\rho_0 < \rho_S$, **planktonic** phase is robust.

- When $\rho_S < \rho_0 < \rho_c$, particles oscillate between **biofilm** and **planktonic** phase.

- When $\rho_c < \rho_0$, biofilms are **metastable**. They *rarely* disperse and reform by dying out.

- Full **phase diagram** depends on carrying capacity $\rho_0$ and domain size $\delta^{-1}$.

---

Problem of **Rogue waves**:  
- Creation mechanism not understood  
- Probability unknown (but $> \text{Gaussian}$)  
- Measurements difficult (you might not be able to tell the tale)
Application: Extreme ocean surface waves

Problem of Rogue waves:
- Creation mechanism not understood
- Probability unknown (but > Gaussian)
- Measurements difficult (you might not be able to tell the tale)

Strategy:
- Random data from observation as input

\[
\begin{align*}
\frac{\partial}{\partial t} u + \frac{1}{2} \frac{\partial^2}{\partial x^2} u + i \frac{8}{\pi} \frac{\partial^3}{\partial x^3} u - \frac{1}{16} \frac{\partial^4}{\partial x^4} u + i \frac{2}{\pi} |u|^2 u + \frac{3}{2} |u|^2 \frac{\partial}{\partial x} u + \frac{1}{4} u^2 \frac{\partial}{\partial x} u^* - i \frac{2}{\pi} |\frac{\partial}{\partial x} u|^2 = 0
\end{align*}
\]

Tobias Grafke Predicting Rare Events via Large Deviations Theory
Application: Extreme ocean surface waves

Problem of Rogue waves:
- Creation mechanism not understood
- Probability unknown (but $>\text{Gaussian}$)
- Measurements difficult (you might not be able to tell the tale)

Strategy:
- Random data from observation as input
- Accurate dynamical system to extrapolate output (MNLS)

\[
\partial_t u + \frac{1}{2} \partial_x u + \frac{i}{8} \partial_x^2 u - \frac{1}{16} \partial_x^3 u + \frac{i}{2} |u|^2 u + \frac{3}{2} |u|^2 \partial_x u + \frac{1}{4} u^2 \partial_x u^* - \frac{i}{2} |\partial_x||u|^2 = 0
\]

Tobias Grafke
Predicting Rare Events via Large Deviations Theory

JONSWAP spectrum
Problem of **Rogue waves**:
- Creation mechanism not understood
- Probability unknown (but $> \text{Gaussian}$)
- Measurements difficult (you might not be able to tell the tale)

**Strategy:**
- **Random data** from observation as input
- Accurate **dynamical system** to extrapolate output (MNLS)
- Use LDT to obtain tails of height distribution

$$\partial_t u + \frac{1}{2} \partial_x u + \frac{i}{8} \partial^2_x u - \frac{1}{16} \partial^3_x u + \frac{i}{2} |u|^2 u + \frac{3}{2} |u|^2 \partial_x u + \frac{1}{4} u^2 \partial_x u^* - \frac{i}{2} |\partial_x| u|^2 = 0$$
Application: Extreme ocean surface waves

**rough sea** \((H_s = 3.3 \text{ m}, \text{BFI} = 0.34)\)

**high sea** \((H_s = 8.2 \text{ m}, \text{BFI} = 0.85)\)

**Probability distribution of spatial maximum of surface height**

**Monte-Carlo simulation (dots)**
Application: Extreme ocean surface waves

**rough sea** \(H_s = 3.3 \text{ m}, BFI = 0.34\)

**high sea** \(H_s = 8.2 \text{ m}, BFI = 0.85\)

*Probability distribution of spatial maximum of surface height*

Comparison between **Monte-Carlo** simulation (dots) and **Large deviation theory** (lines)

Tobias Grafke  Predicting Rare Events via Large Deviations Theory
Application: Extreme ocean surface waves

- $H_s = 3.3 \text{ m}$
- $H = 8.5 \text{ m}$
- $t = 20 \text{ min}$

- $H_s = 3.3 \text{ m}$
- $H = 5.5 \text{ m}$
- $t = 10 \text{ min}$

- $H_s = 3.3 \text{ m}$
- $H = 4.3 \text{ m}$
- $t = 0 \text{ min}$


Tobias Grafke Predicting Rare Events via Large Deviations Theory
LDT as WKB approximation

Consider Markov jump process with generator $\mathcal{L}$, s.t.

\[
\frac{\partial}{\partial t} f = \mathcal{L}^\dagger f \quad \text{(forward Kolmogorov, Fokker-Planck, Master eqn)}
\]
\[
\frac{\partial}{\partial t} f = \mathcal{L} f \quad \text{(backward Kolmogorov)}
\]

e.g. for diffusion above, $\mathcal{L} = b \cdot \nabla + \frac{1}{2} \varepsilon \nabla \nabla$

For WKB approximation, $f \sim \exp(\varepsilon^{-1} S)$, BKE becomes to leading order

\[
\frac{\partial}{\partial t} f = b \cdot \nabla S + \frac{1}{2} (\nabla S)^2
\]

which is a Hamilton-Jacobi equation,

\[
\frac{\partial}{\partial t} f = H(x, \nabla S), \quad H(x, p) = b \cdot p + \frac{1}{2} p^2
\]

This is the LDT Hamiltonian from before(!), but works for all MJP

- for additive Gaussian SDE
- for Lévy processes
- other cases, i.e. stochastic averaging

- for multiplicative Gaussian SDE
- for jump process

Tobias Grafke
Predicting Rare Events via Large Deviations Theory
Challenges: Infinite transition time and geometric rate function

We actually want the most probable event **regardless of duration**.

Drop the restriction of a pre-defined transition time $T$:

$$I(\tilde{\phi}) = \inf_{T \in (0, \infty)} \inf_{\phi} I_T(\phi)$$

Possibly attains minimum at $T \to \infty$. 

Effecitvely: Reduce minimisation over all paths to finding geodesic of the associated (almost Finsler) metric.


Tobias Grafke Predicting Rare Events via Large Deviations Theory
We actually want the most probable event \textbf{regardless of duration}.

Drop the restriction of a pre-defined transition time $T$:

$$I(\tilde{\phi}) = \inf_{T \in (0, \infty)} \inf_{\phi} I_T(\phi)$$

Possibly attains minimum at $T \to \infty$. Since $H(\phi, \theta) = h = \text{cst}$, we have

$$\int L(\phi, \dot{\phi}) \, dt = \int \sup_{\theta} \left( \langle \dot{\phi}, \theta \rangle - H(\phi, \theta) \right) \, dt = \sup_{\theta : H(\phi, \theta) = h} \int \langle \dot{\phi}, \theta \rangle \, dt + hT$$

Effectively:

Reduce minimisation over all paths to finding \textbf{geodesic} of the associated (almost Finsler) \textbf{metric}.

\textit{Heymann, Vanden-Eijnden (2008), Grafke, Schäfer, Vanden-Eijnden (2017)}
Summary

Main theme

Obtain statistics of and structures for rare events by numerically computing large deviation minimisers for spatially extended systems.

Challenges:

- Analytic solutions not available
- Needs PDE constrained optimisation (on GPUs)
- Simplification necessary through nature of problem

Applications:

- Fluid dynamic
- Non-equilibrium stat. mech.
- Rogue waves