RARE AND EXTREME EVENTS IN CLIMATE DYNAMICS: SAMPLING USING RARE EVENT ALGORITHMS AND MACHINE LEARNING

Freddy Bouchet (CNRS and ENS de Lyon)

With:

a) Eric Simonnet and Joran Rolland (Jupiter's abrupt climate change)

b) Francesco Ragone and Jeroen Wouters (rare event algorithms in climate models)

b) Dario Lucente, George Miloshevich and Francesco Ragone (extreme heat waves)

c) Valerian Jacques-Dumas, George Miloshevich, Francesco Ragone, Pierre Borgnat and Patrice Abry (prediction of extreme heat waves with deep neural networks)

d) Dario Lucente, Joran Rolland and Corentin Herbert (coupling rare event algorithms and machine learning)

ICTAM 2020, August 2021

ENS DE LYON

Lab PHYS ENS de LYON

European Research Council
Established by the European Commission

IDEX LYON UNIVERSITÉ DE LYON

https://labos1point5.org

Labos 1point5

Le collectif Les initiatives Les ressources Nous rejoindre

Réduire l'empreinte de nos activités de recherche sur l'environnement

À propos

Labos 1point5 est un collectif de membres du monde académique, de toutes disciplines et sur tout le territoire, partageant un objectif commun :
mieux comprendre et réduire l'impact des activités de recherche scientifique sur l'environnement, en particulier sur le climat.

Reduce the environmental footprint of our researches.

https://labos1point5.org

How to build a strategy for the ecology transition at ENS de Lyon

Proposition by the ecology transition group at ENS de Lyon:

Chapter 1 Ecology transition at ENS de Lyon : how to build together a cultural change?

Chapter 2 Quantifying the environmental impact of ENS de Lyon

Chapter 3 How to reduce the impact of building and infrastructures

Chapter 4 Impact of travels

Chapter 5 Daily life and environment (recycling, wastes, transport, bikes, ...)

Chapter 6 Environmental impact of digital technologies

Chapter 7 Teaching, research, and the environment

Lien Web : lettre d'interpelation des collegues et de la direction (01/2021)

Lien Web : construire une stratégie (01/2021)

Rare event algorithms for climate dynamics
Outline

I) Introduction: rare events do matter - rare event algorithms

II) Rare events algorithms for predicting Jupiter's abrupt climate change

III) Rare events algorithms for predicting extreme heat waves

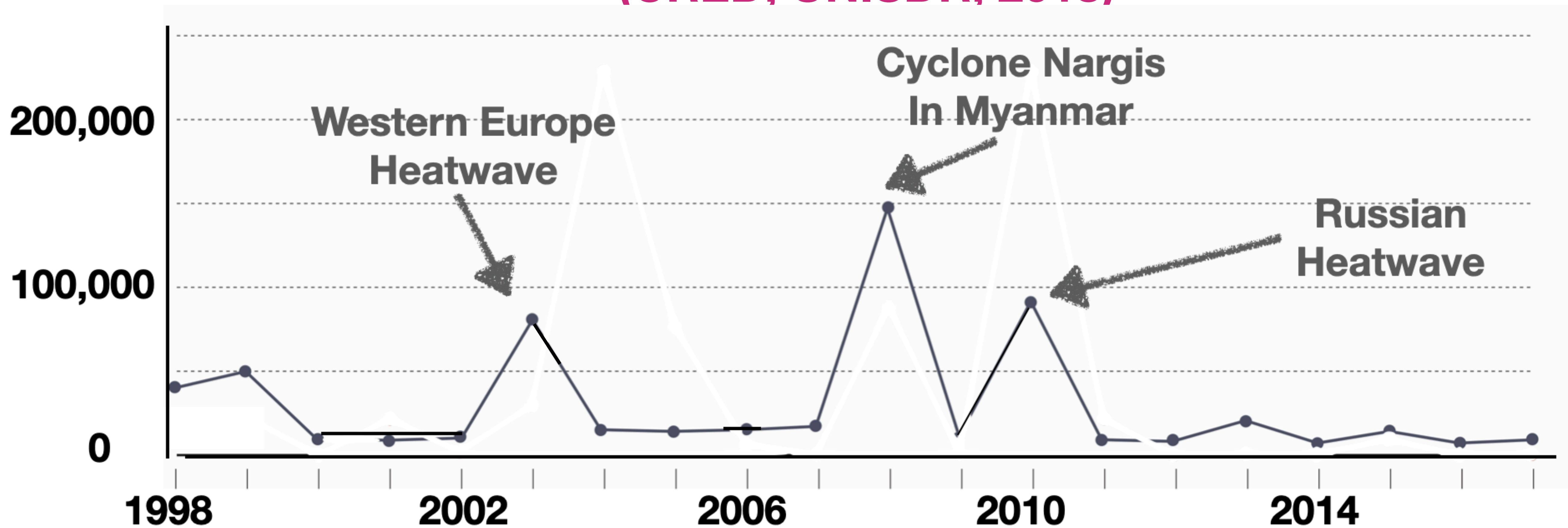
IV) Coupling rare event algorithms with machine learning

V) Predicting extreme heat waves and committor functions using deep neural networks

The few most extreme climate events have more impact than all the others

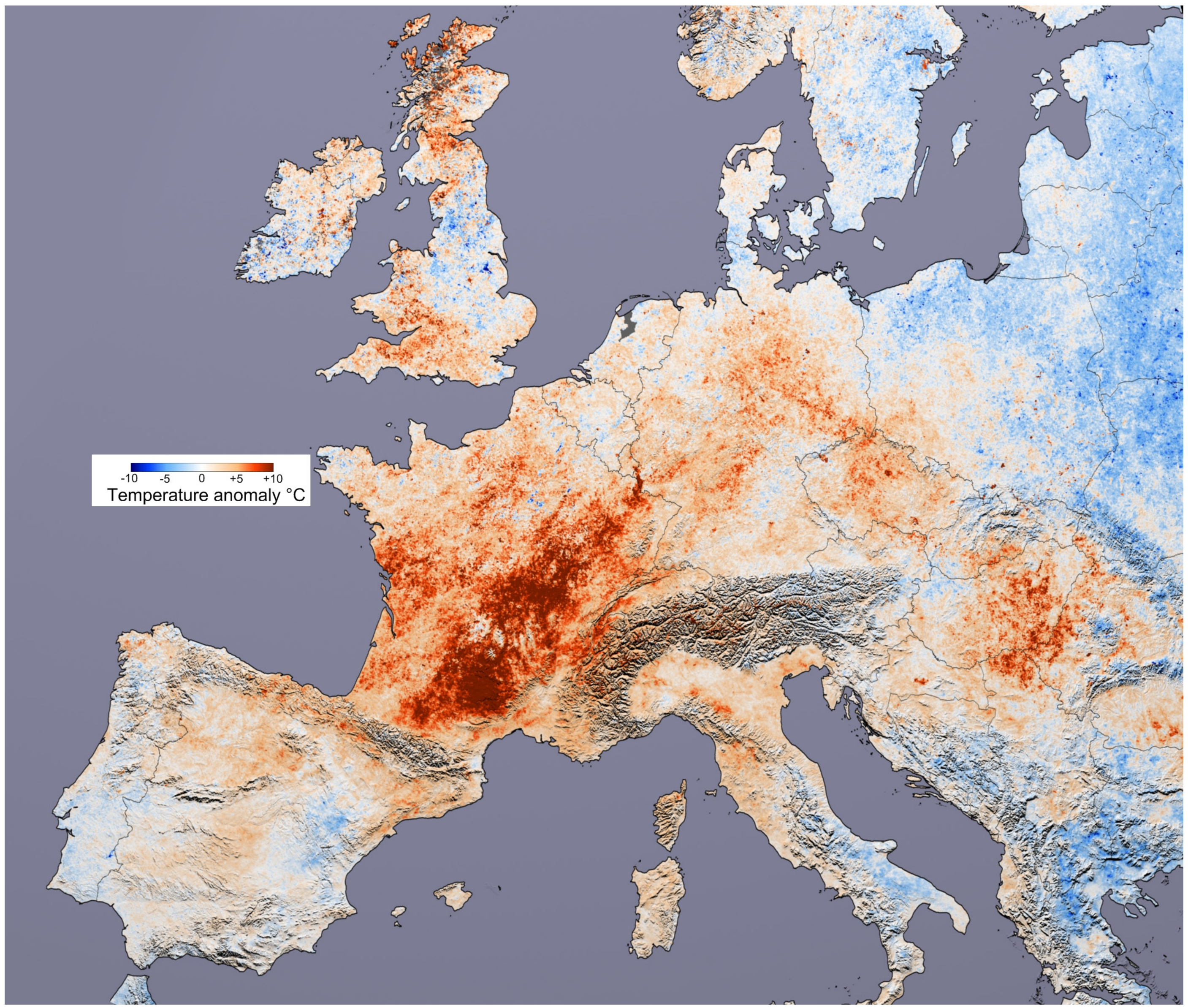
Annual deaths by major climate related disaster

(CRED, UNISDR, 2018)



We need to study extremely rare events. This is a serious scientific challenge.

What is the probability (return time) of the 2003 Europe heatwave ?



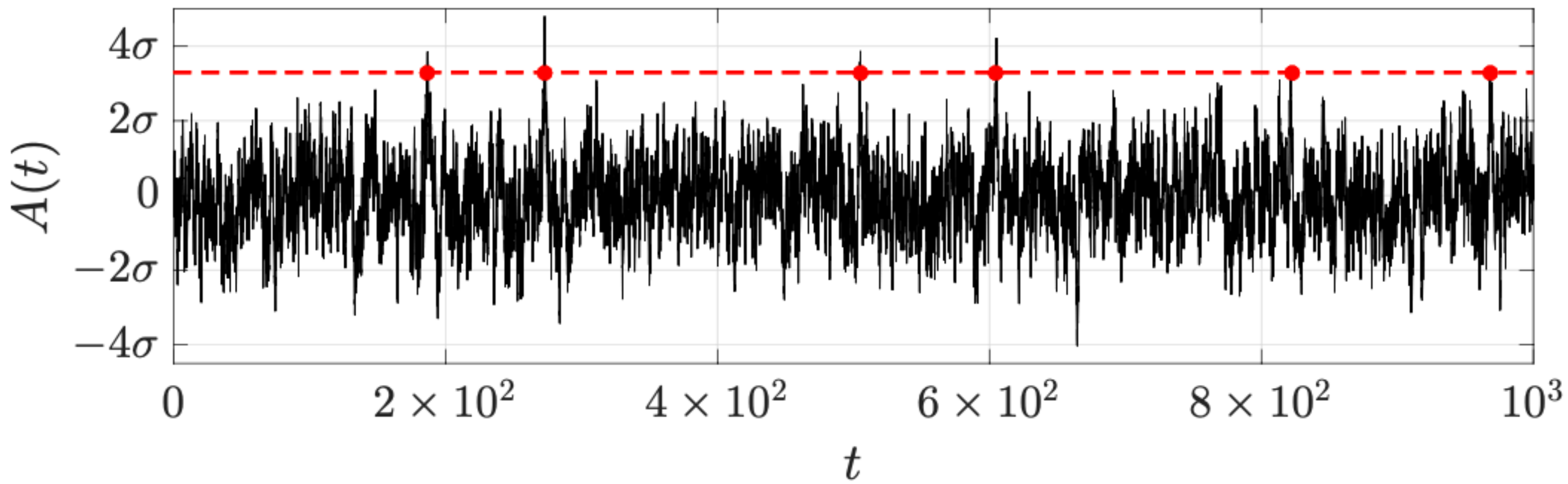
July 20 2003-August 20 2003 land surface temperature minus the average for the same period for years 2001, 2002 and 2004 (TERRA MODIS).

Why are return times so hard to estimate?

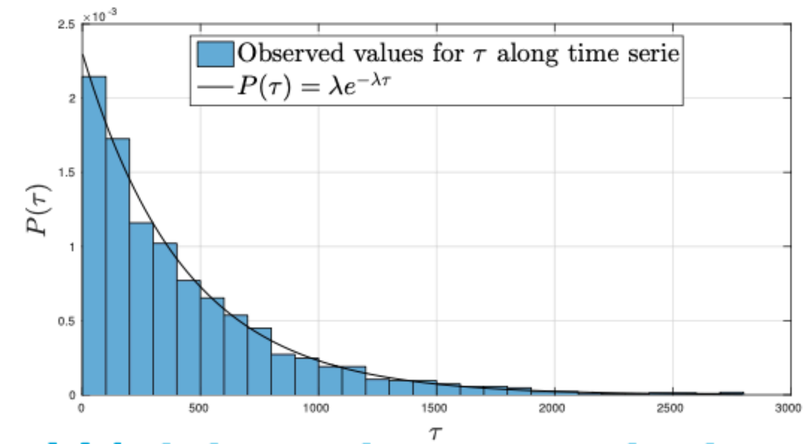
i) lack of observation data, ii) model biases,

iii) because of rareness, gathering good model statistics is too costly.

Extreme Events, Poisson Statistics, and Return Times



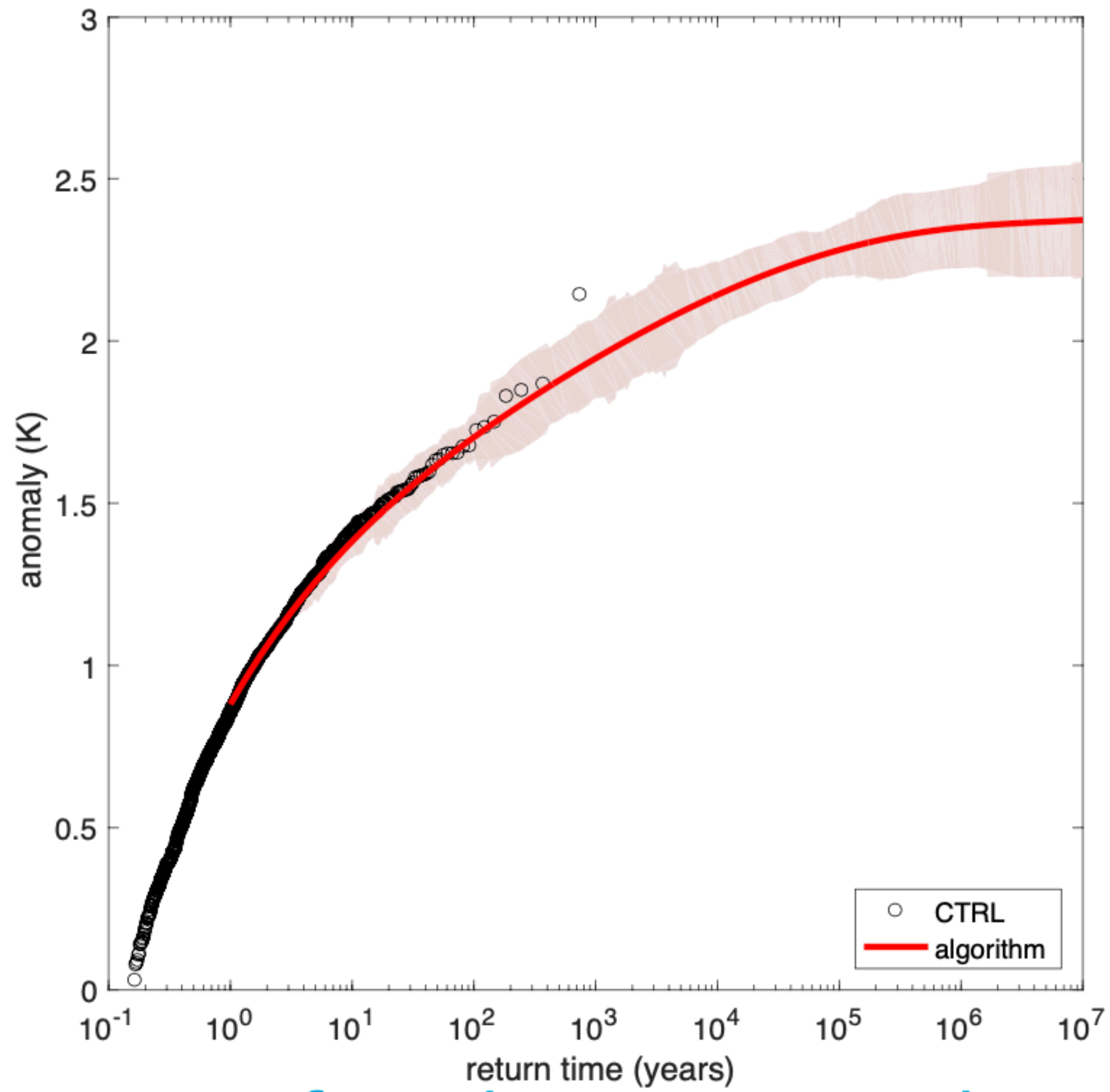
Ornstein–Uhlenbeck extremes



Waiting time statistics

For systems with a single state, rare enough events are uncorrelated and have a Poisson statistics

The Return Time of Extreme Heat Waves

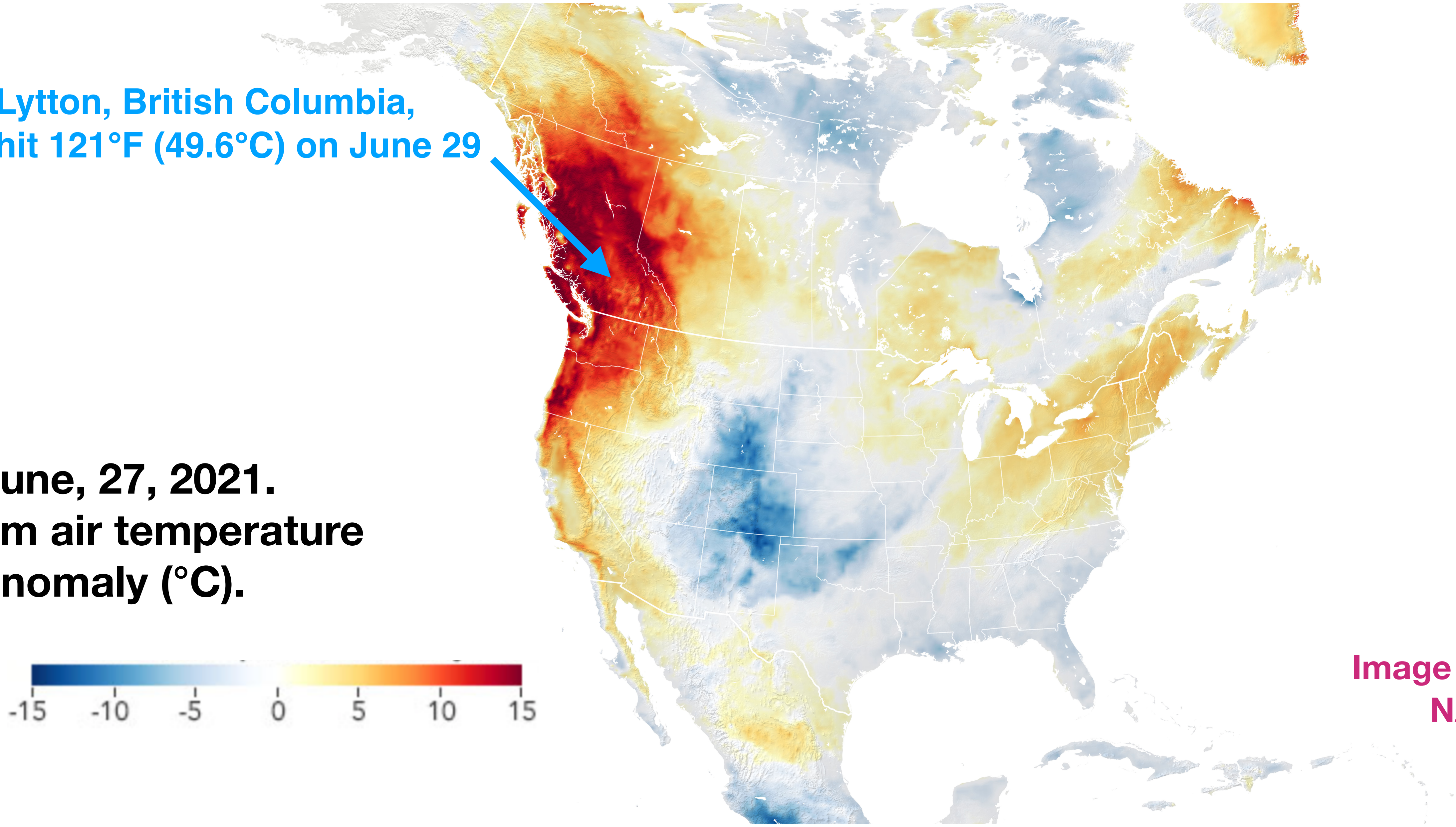


Return time of 90 day European heat waves

F. Ragone, J. Wouters, and F. Bouchet, PNAS, 2018

The 2021 northwest America heatwave - Unprecedented

Lytton, British Columbia, hit 121°F (49.6°C) on June 29



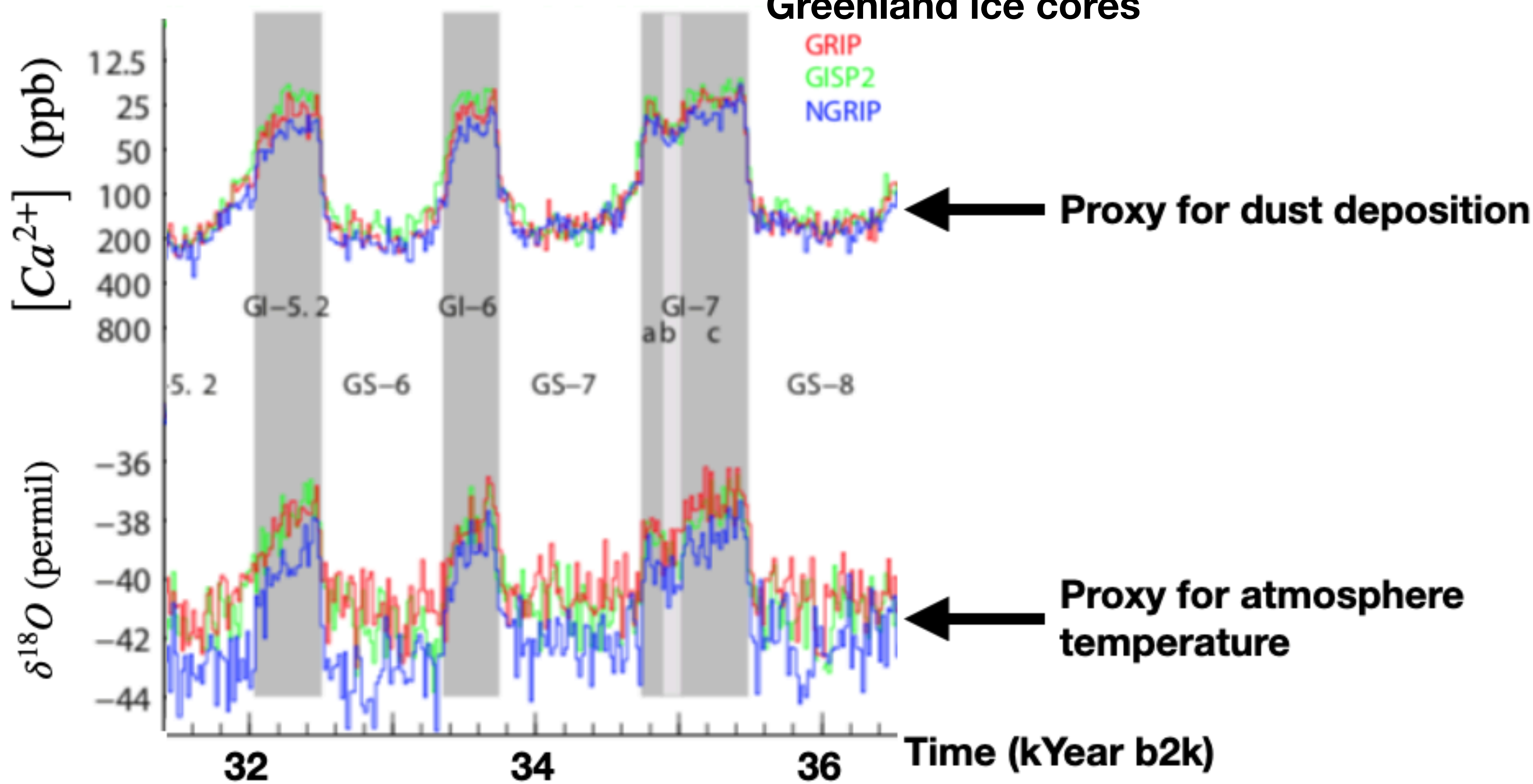
June, 27, 2021.
2m air temperature anomaly (°C).

Image Joshua Stevens
NASA, GEOS

How often shall we expect this to happen?

We do not know (see WWA study).

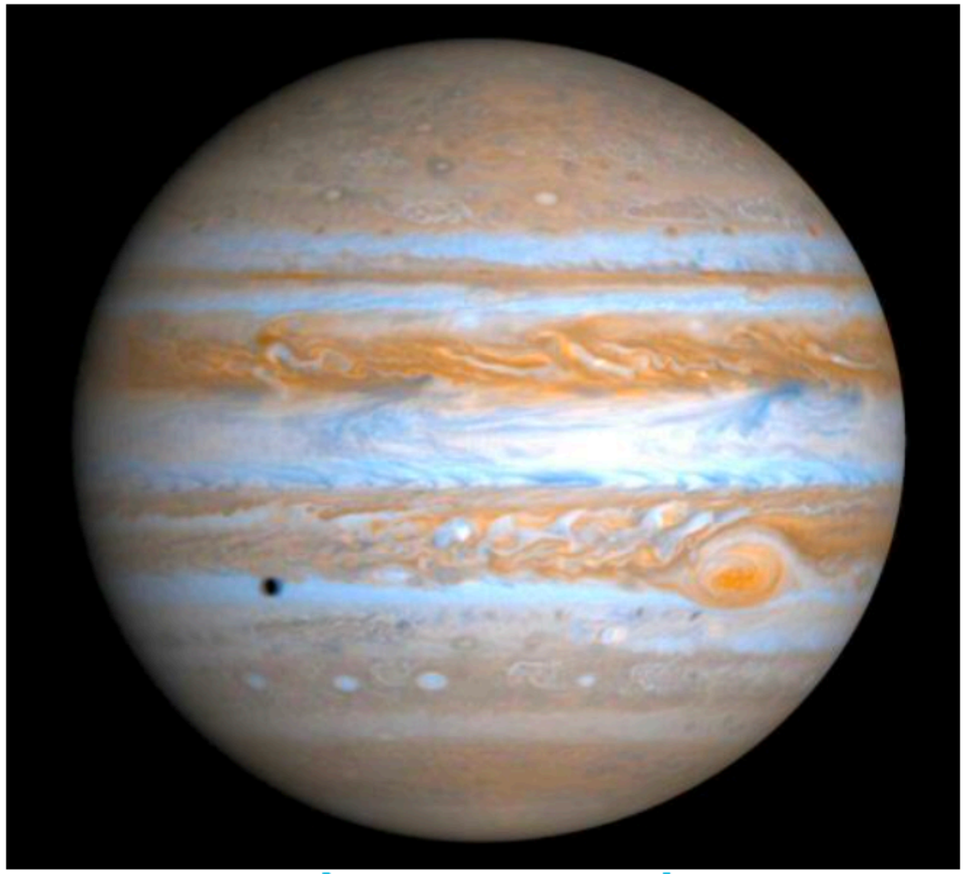
Rare events matter
2 - Spontaneous abrupt climate changes



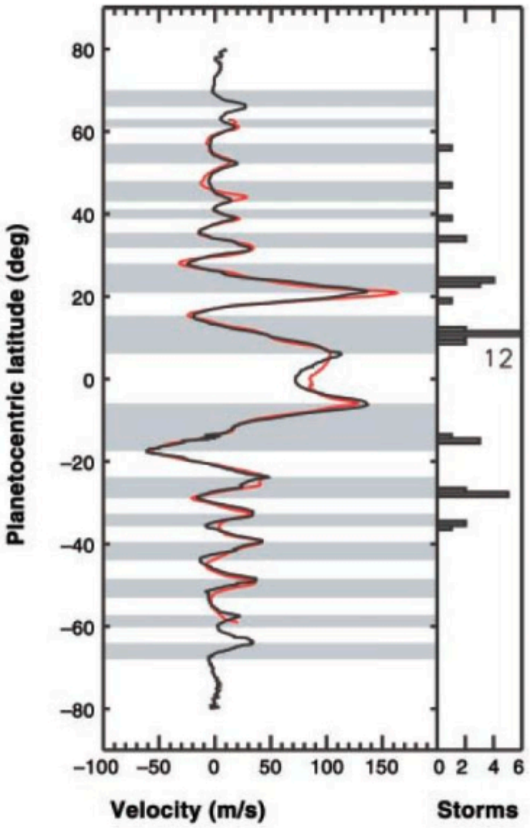
Climate abrupt transitions during the last glacial period.
(S. Rasmussen et al, 2014)

Jupiter's Zonal Jets

We look for a theoretical description of zonal jets



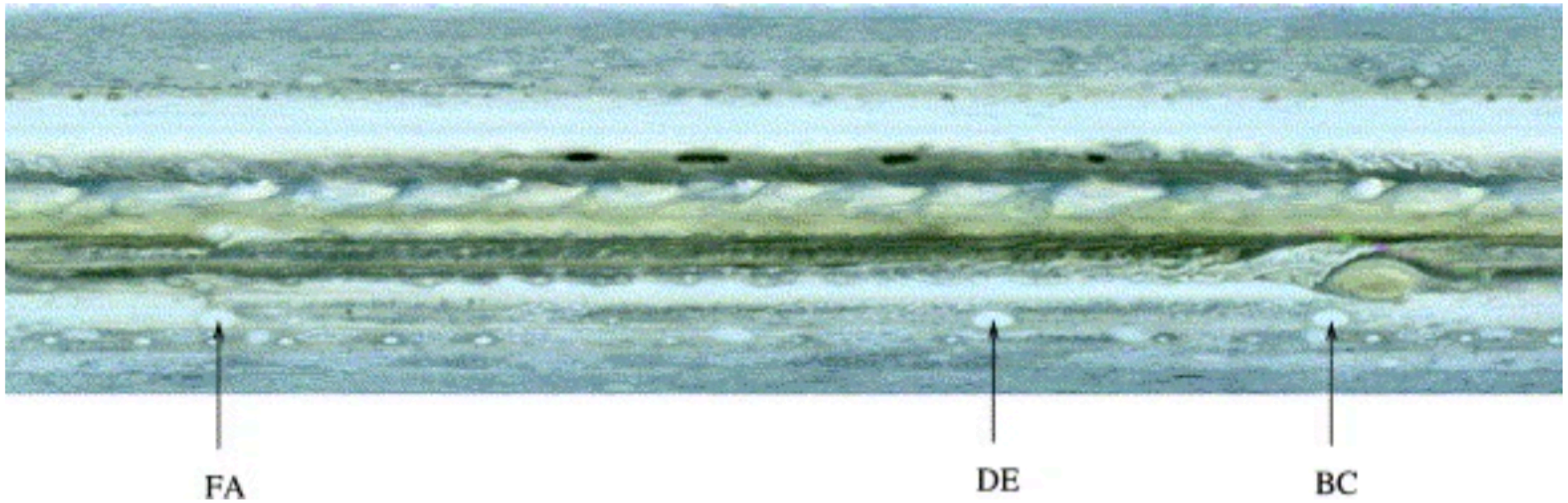
Jupiter's troposphere

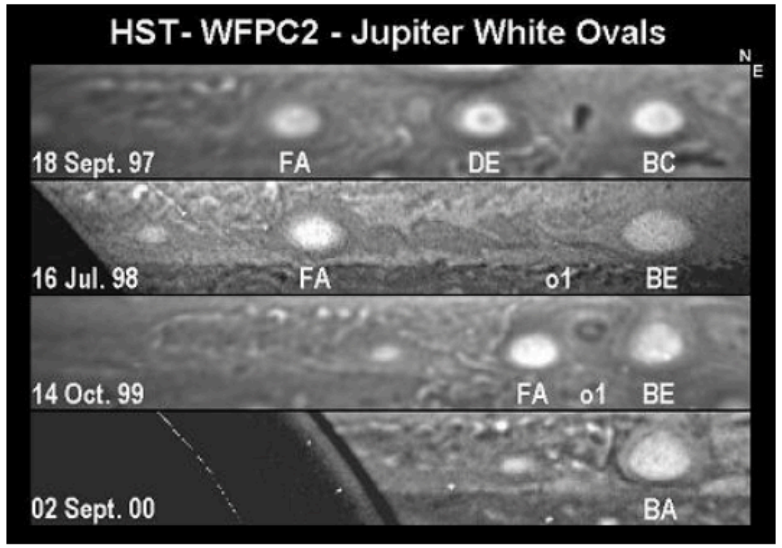


Jupiter's zonal winds (Voyager and Cassini, from Porco et al 2003)

Jupiter's Abrupt Climate Change

Have we lost one of Jupiter's jets ?





Jupiter's white ovals (see Youssef and Marcus 2005)

The white ovals appeared in 1939-1940 (Rogers 1995). Following an instability of one of the zonal jets?

Large deviation theory: an aesthetic journey to rare events

- **Large deviation theory** is a general framework to describe probability distribution in asymptotic limits:

$$P\left[X_\epsilon = x\right] \underset{\epsilon \ll 1}{\asymp} \mathrm{e}^{-\frac{\mathscr{F}[x]}{\epsilon}}.$$

- Large deviation theory is **the modern language of statistical mechanics.**

- It provides the mathematical and conceptual framework to study and understand rare events.

Large deviation theory: an aesthetic journey to rare events

- **Large deviation theory** is a general framework to describe probability distribution in asymptotic limits:

$$P\left[X_\epsilon = x\right] \underset{\epsilon \ll 1}{\asymp} \mathrm{e}^{-\frac{\mathscr{F}[x]}{\epsilon}}.$$

- Large deviation theory is **the modern language of statistical mechanics.**

- It provides the mathematical and conceptual framework to study and understand rare events.

Not today

Three key problems in the study of climate extreme events

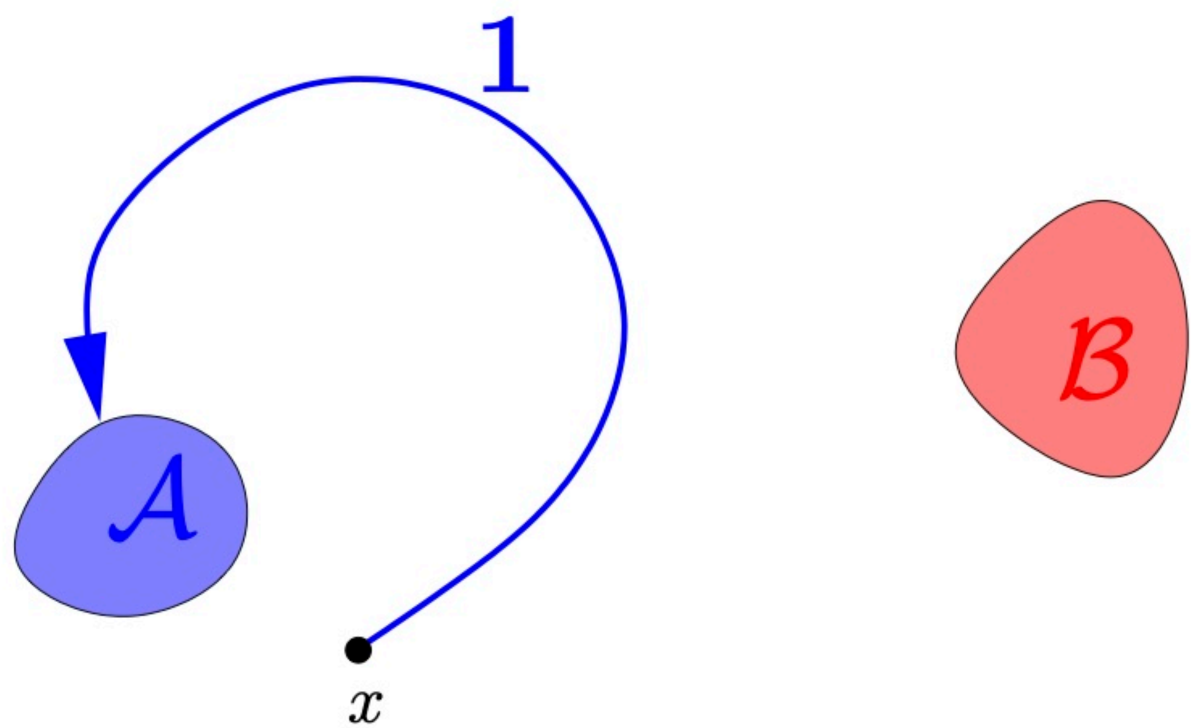
- The historical records are way too short to make any meaningful predictions for the rarest events (those that matter the most).
- Climate models are wonderful tools, but they have biases. The more precise they are, the more computationally costly they are.
- Because they are too rare, the most extreme events cannot be computed using direct numerical simulations (the needed computing times are often unfeasible).

The practical questions: How to sample the probability and dynamics of rare events in complex models? How to build effective models which are relevant for estimating the probability of rare events?

How to study a 10 000 year heat wave with a 200 year simulation ?

- **Because they are too rare, extreme events cannot be computed using direct numerical simulations** (the needed computing times are often unfeasible).
- **Rare event algorithms:** Kahn and Harris (1953).
- **Statistical mechanics:** diffusion Monte-Carlo, Wang Landau algorithms, go with the winners, …
- **Applied Mathematics:** Chandler, Vanden-Eijnden, Schuss, Del Moral, Dupuis, Lelièvre, Guyader, …
- **For turbulence and climate applications:** J. Weare and D. Abbot, R. Grauer and T. Grafke, E. Vanden-Eijnden, Lyon group, …

How to compute extremely rare trajectories ?

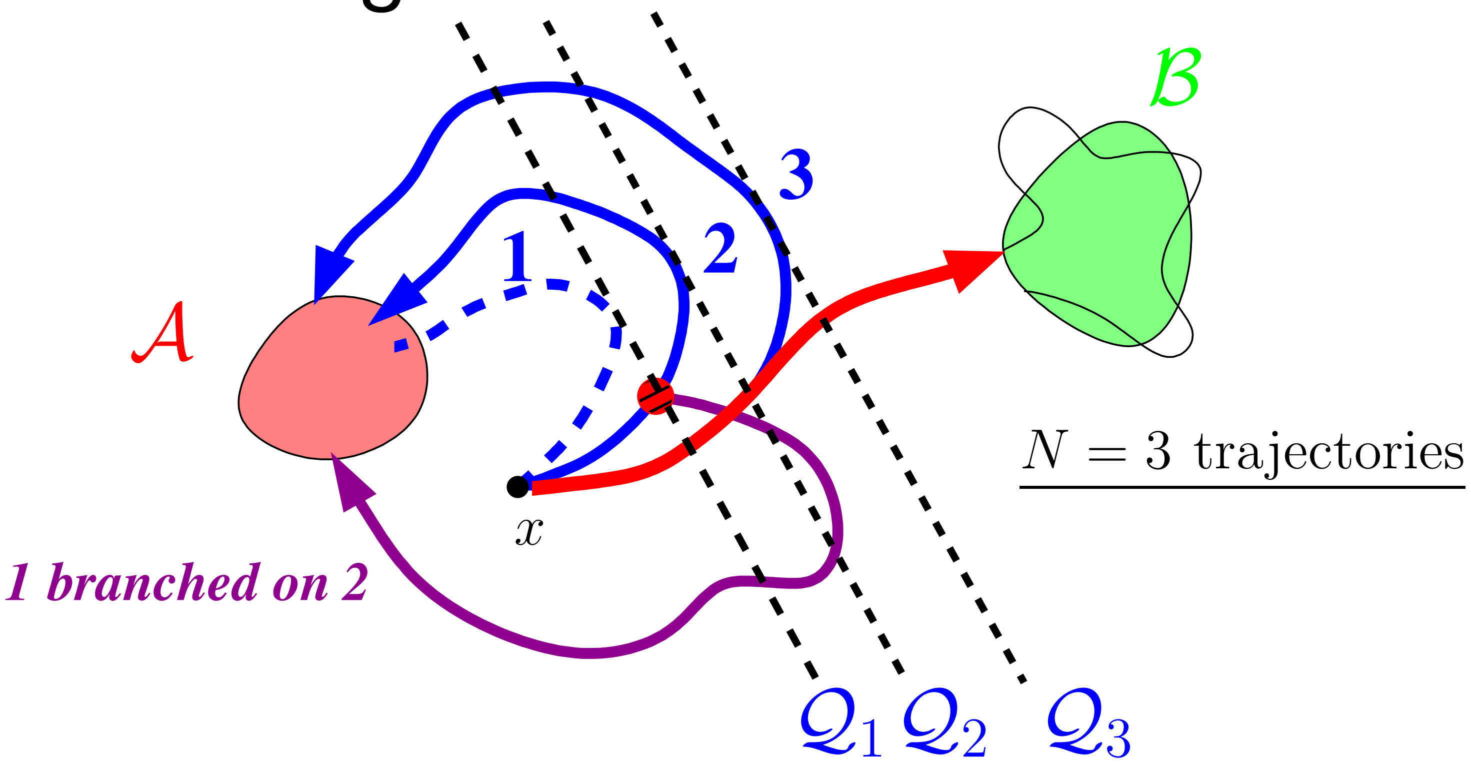


Aim: compute **extremely rare trajectories** from the point x to the rare event set $\mathscr{B}$.

Most of the time, trajectories that start from x end in $\mathscr{A}$.
The probability to reach $\mathscr{B}$ may be 10^{-3} or 10^{-20}.

The Adaptive Multilevel Splitting (AMS) rare event algorithm

Strategy: ensemble computation, selection, pruning and cloning.



Probability estimate:

$\hat{p} = (1 - 1/N)^K$,

where N is the clone number and K is the iteration number.

Cérou, Guyader (2007). Cérou, Guyader, Lelièvre, and Pommier (2011).
PDEs: Rolland, Bouchet et Simonnet (2016) - TAMS: Lestang et al (2018)
Atmosphere turbulent jets: Rolland, Bouchet et Simonnet (2019 and 2021).

Rare event algorithms for climate dynamics
Outline

I) Introduction: rare events do matter - rare event algorithms

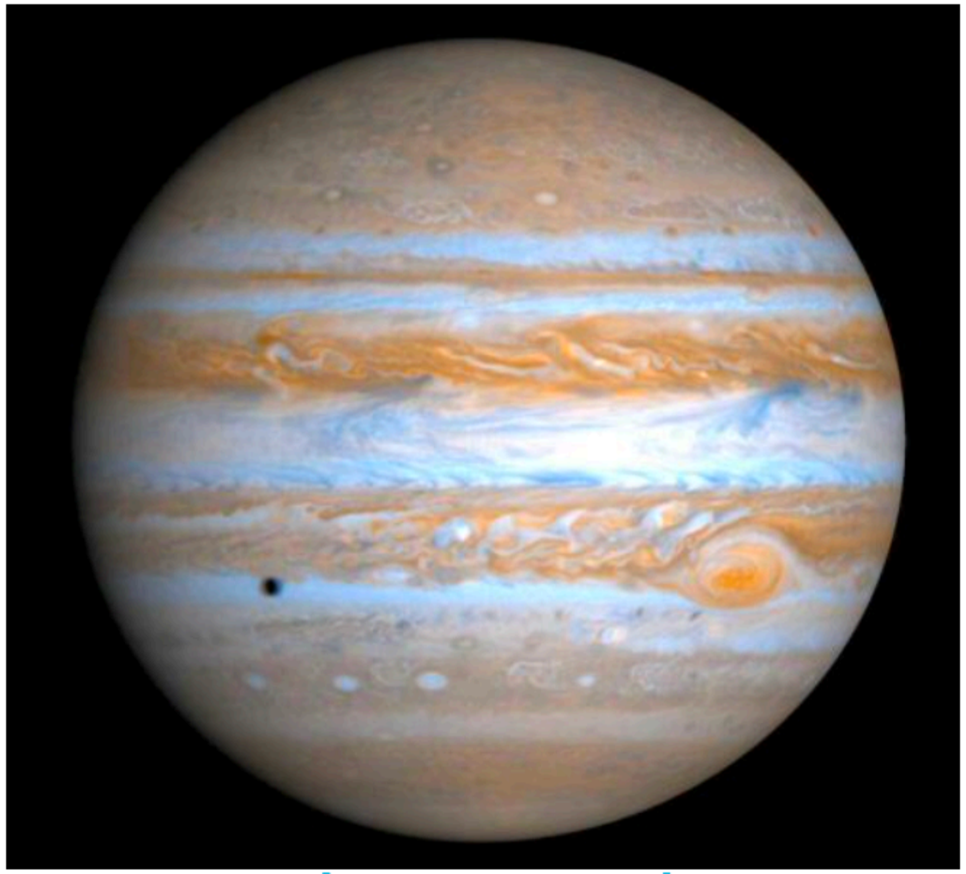
II) Rare events algorithms for predicting Jupiter's abrupt climate change

III) Rare events algorithms for predicting extreme heat waves

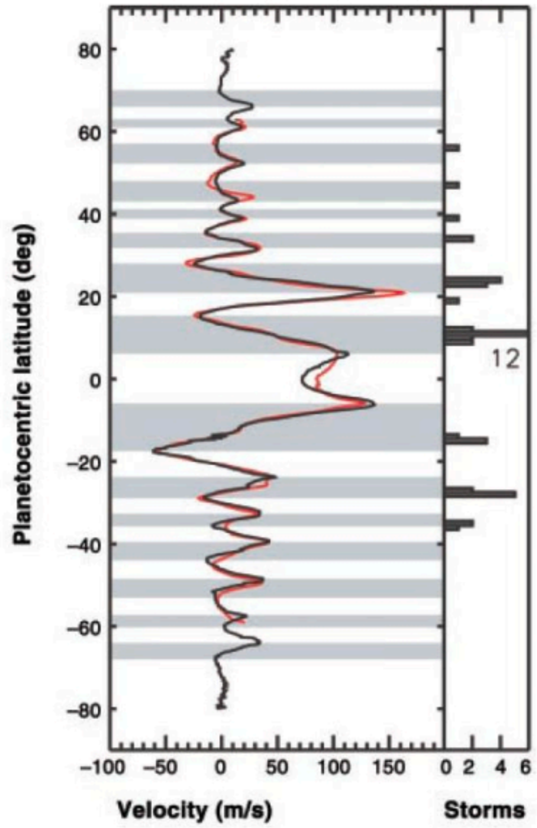
IV) Coupling rare event algorithms with machine learning

V) Predicting extreme heat waves and committor functions using deep neural networks

II) Rare events algorithms for predicting Jupiter's abrupt climate change



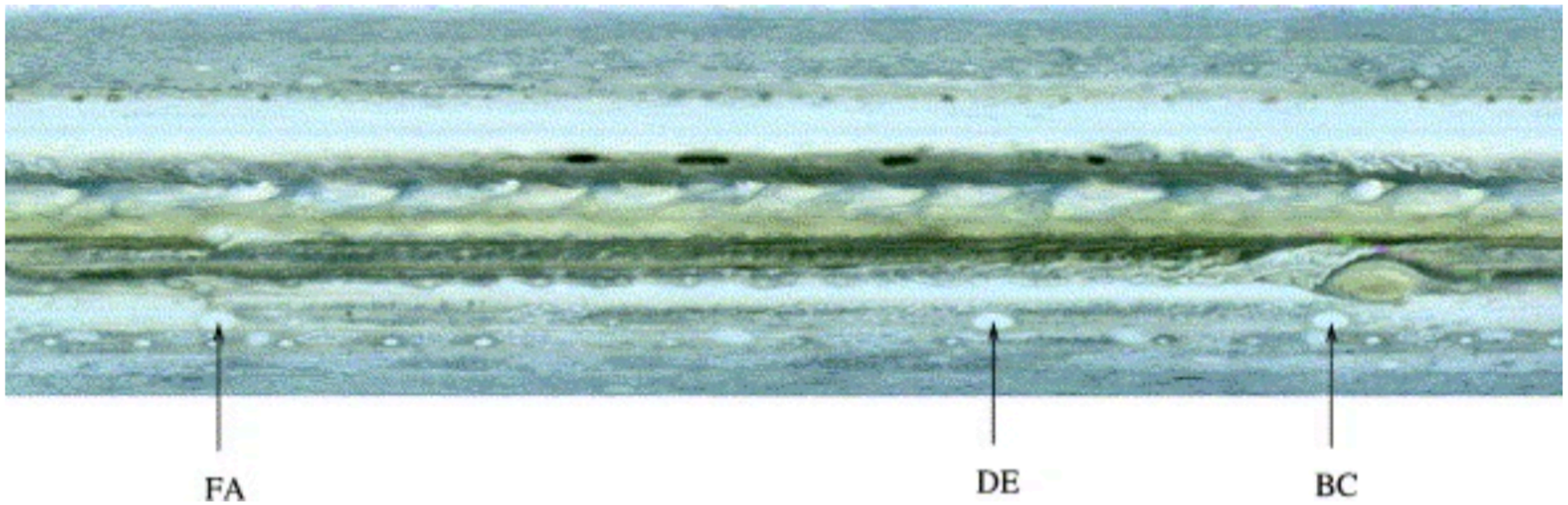
Jupiter's troposphere

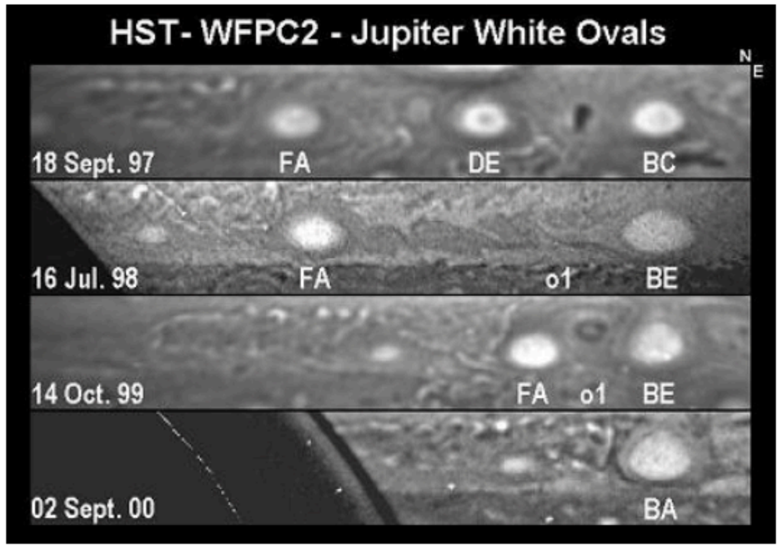


Jupiter's zonal winds (Voyager and Cassini, from Porco et al 2003)

Jupiter's Abrupt Climate Change

Have we lost one of Jupiter's jets ?





Jupiter's white ovals (see Youssef and Marcus 2005)

The white ovals appeared in 1939-1940 (Rogers 1995). Following an instability of one of the zonal jets?

The Barotropic Quasi-Geostrophic Equations

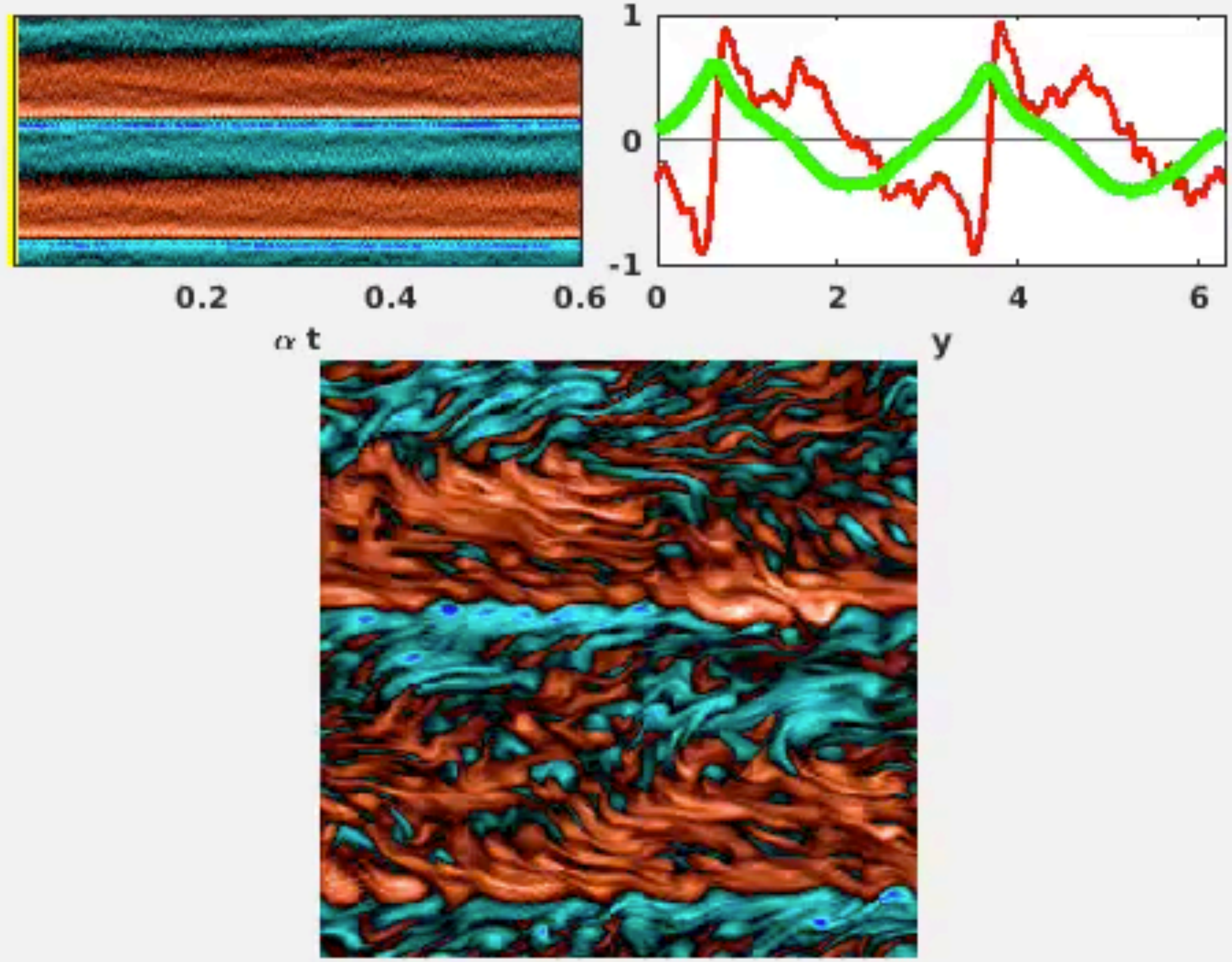
- The simplest model for geostrophic turbulence.
- Quasi-Geostrophic equations with random forces

$$\frac{\partial q}{\partial t} + \mathbf{v}.\nabla q = \nu \Delta \omega - \alpha \omega + \sqrt{2\alpha} f_s,$$

where $\omega = (\nabla \wedge \mathbf{v}).\mathbf{e}_z$ is the vorticity, $q = \omega + \beta y$ is the Potential Vorticity (PV), f_s is a random Gaussian field with correlation $\langle f_S(\mathbf{x},t) f_S(\mathbf{x}',t') \rangle = C(\mathbf{x}-\mathbf{x}')\delta(t-t')$, ε is the average energy input rate, λ is the Rayleigh friction coefficient.

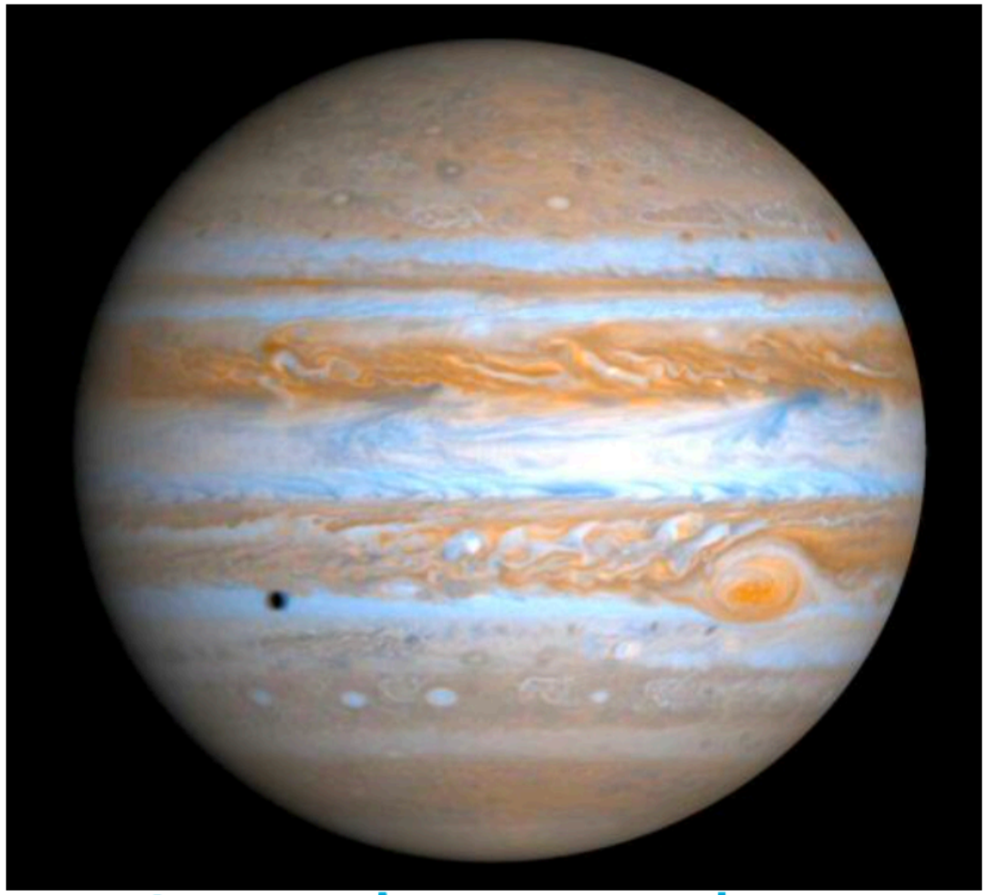
- Spin up or spin down time $= 1/\alpha \ll 1 =$ jet inertial time scale.
- A reasonable model for Jupiter's zonal jets.

Dynamics of the Barotropic Quasi-Geostrophic Equations

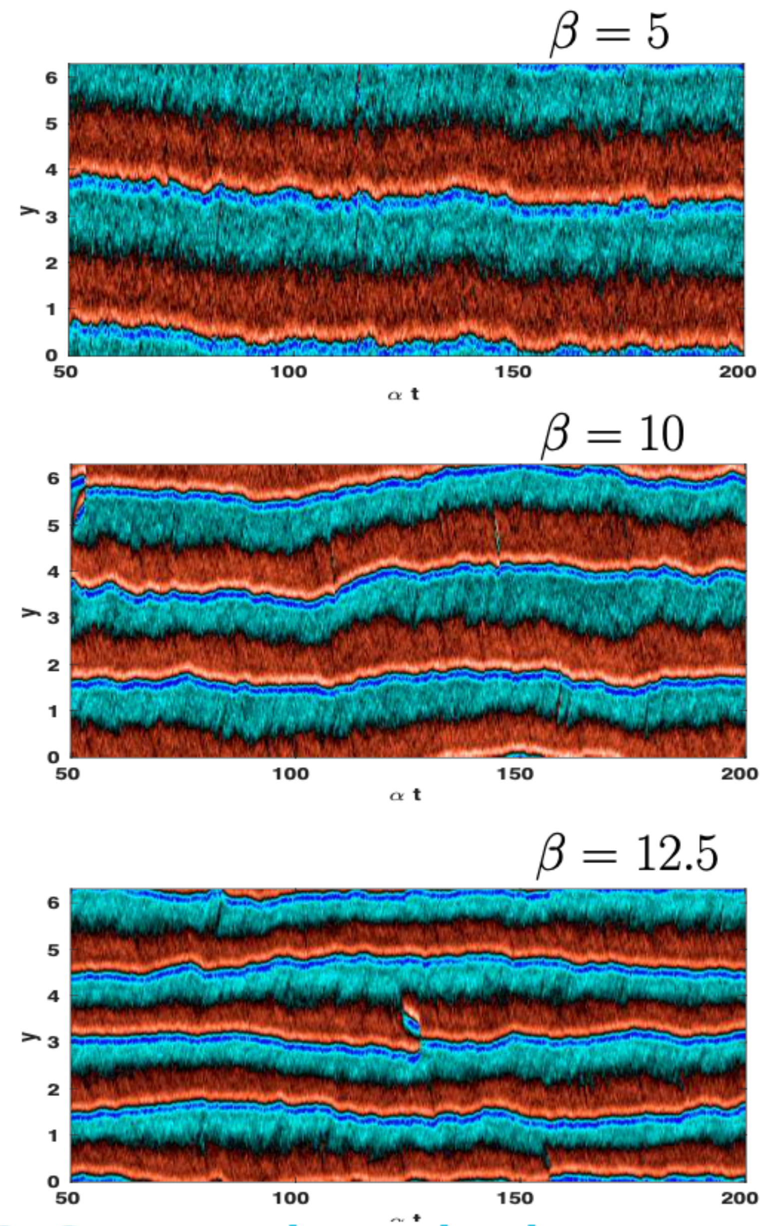


Top: Zonally averaged vorticity (Hovmöller diagram and red curve) and velocity (green). Bottom: vorticity field

Multistability for Quasi-Geostrophic Jets



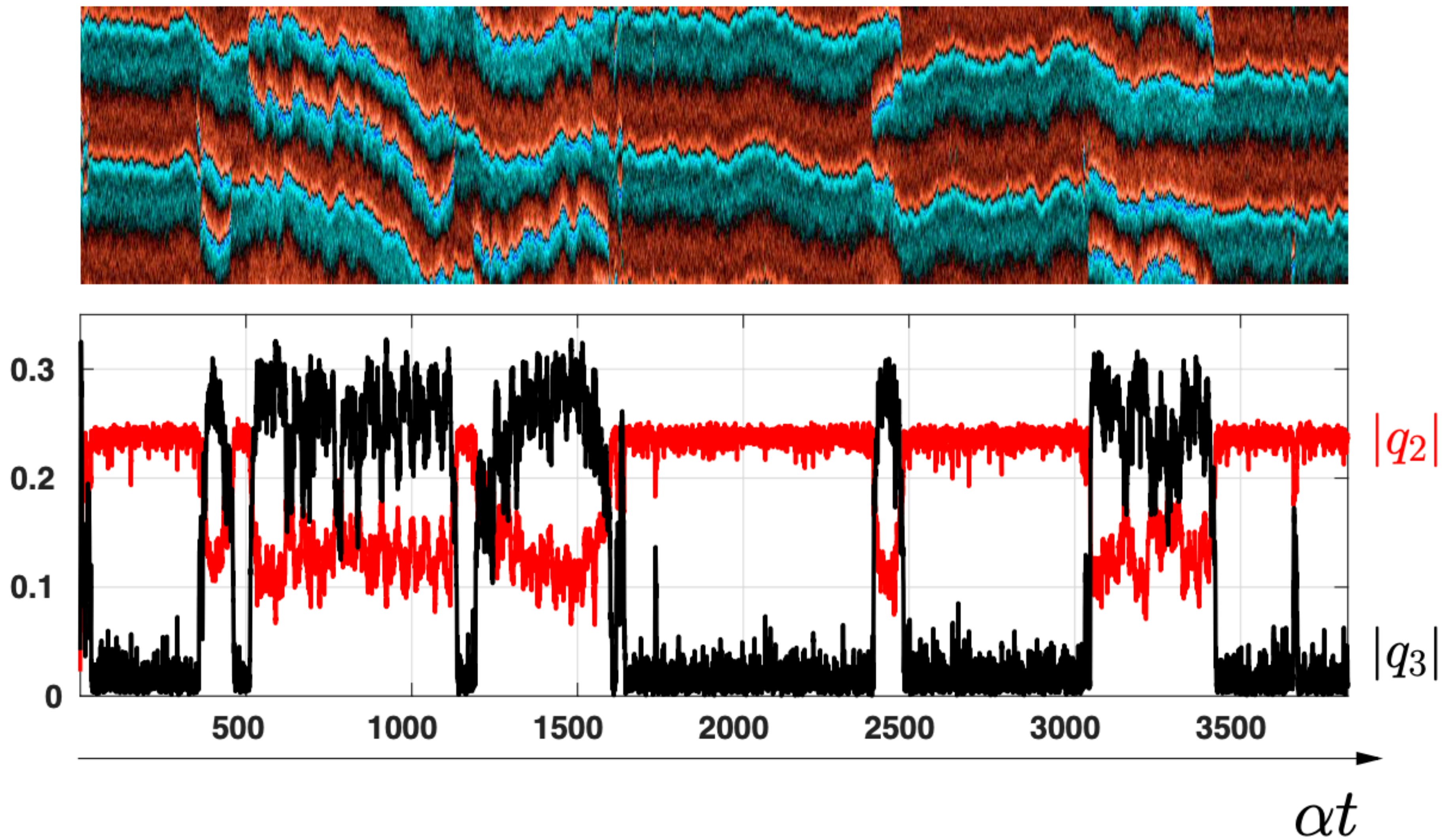
Jupiter's atmosphere



QG zonal turbulent jets

- Multiple attractors had been observed previously by B. Farrell and P. Ioannou.

Rare Transitions Between Quasigeostrophic Jets

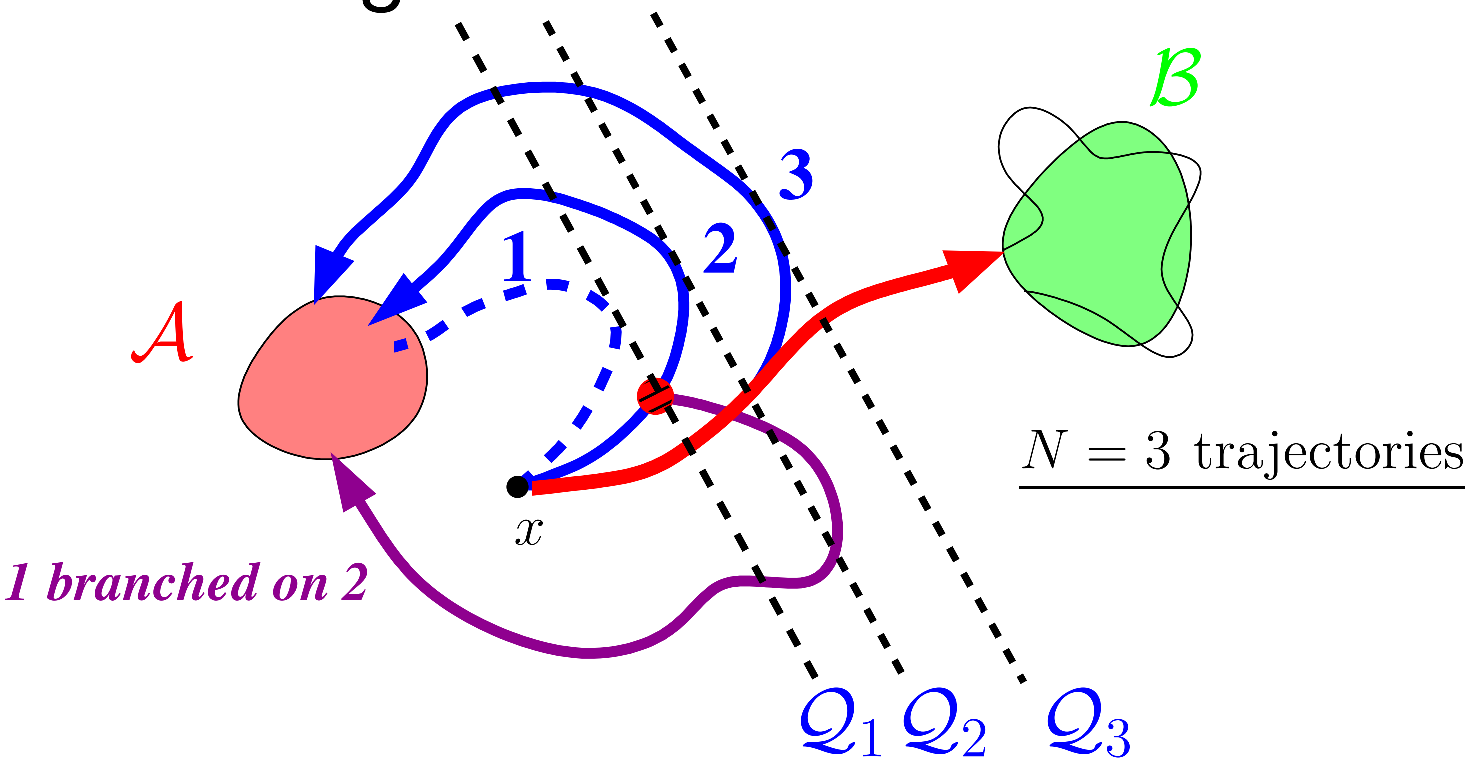


Rare transitions for quasigeostrophic jets (with E. Simonnet)

- This is the first observation of spontaneous transitions.
- How to predict those rare transitions? What is their probability? Which theoretical approach?

The Adaptive Multilevel Splitting (AMS) rare event algorithm

Strategy: ensemble computation, selection, pruning and cloning.



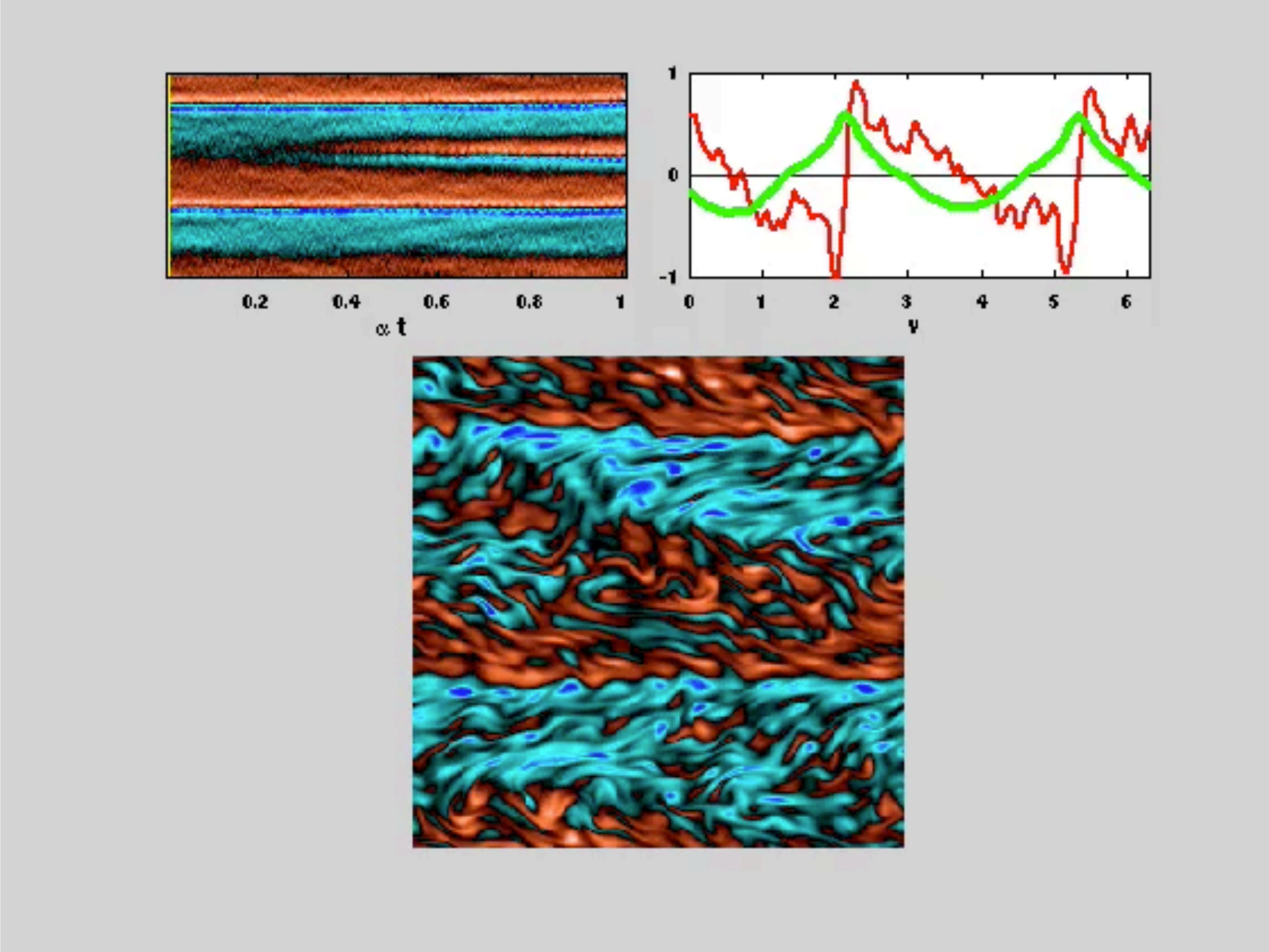
Probability estimate:

$\hat{p} = (1 - 1/N)^K,$

where N is the clone number and K is the iteration number.

Cérou, Guyader (2007). Cérou, Guyader, Lelièvre, and Pommier (2011).
PDEs: Rolland, Bouchet et Simonnet (2016) - **TAMS:** Lestang et al (2018)
Atmosphere turbulent jets: Rolland, Bouchet et Simonnet (2019 and 2021).

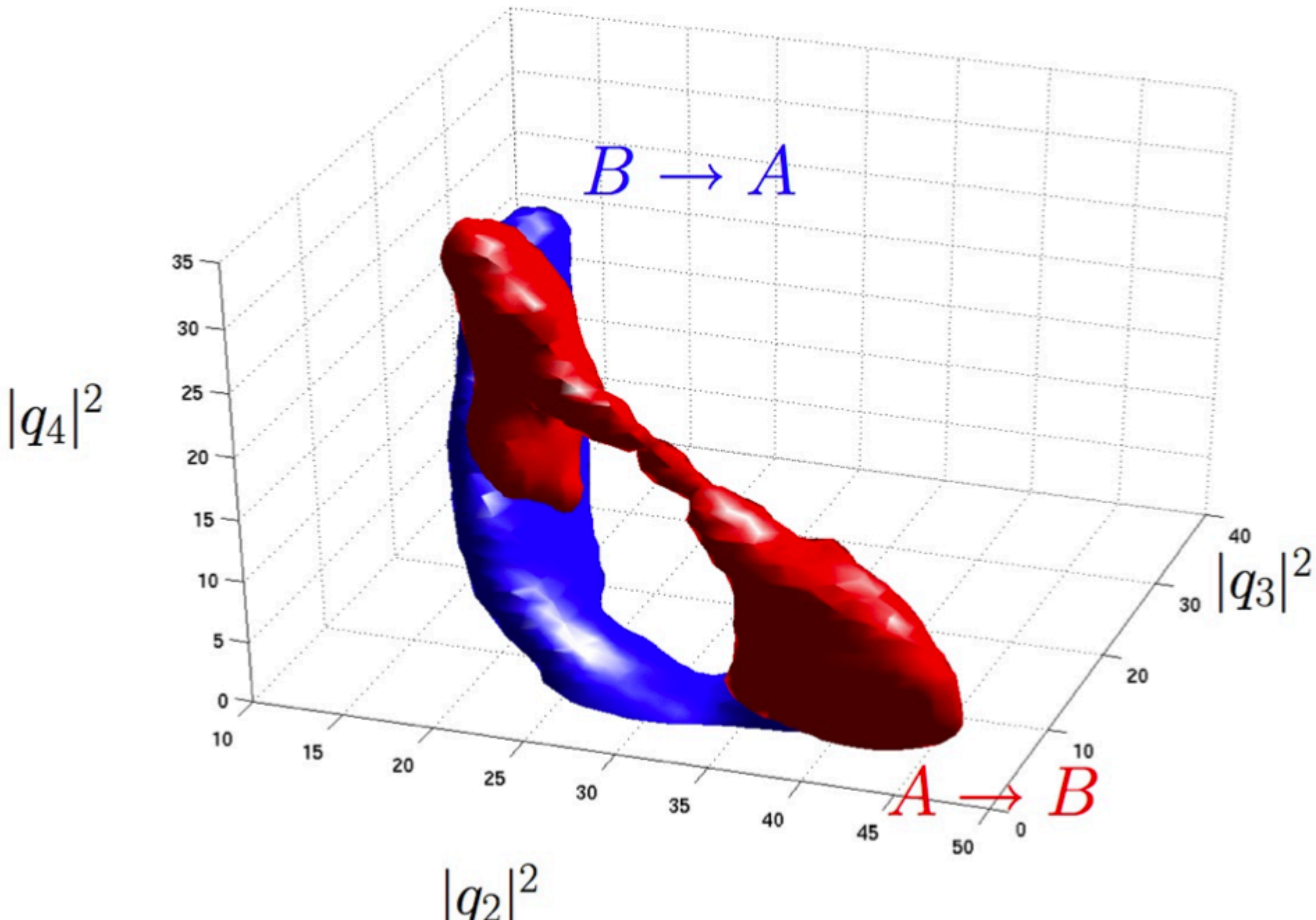
A Transition from 2 to 3 Jets



Top: Zonally averaged vorticity (Hovmöller diagram and red curve) and velocity (green). Bottom: vorticity field

Atmosphere Jet "Instantons" Computed using the AMS

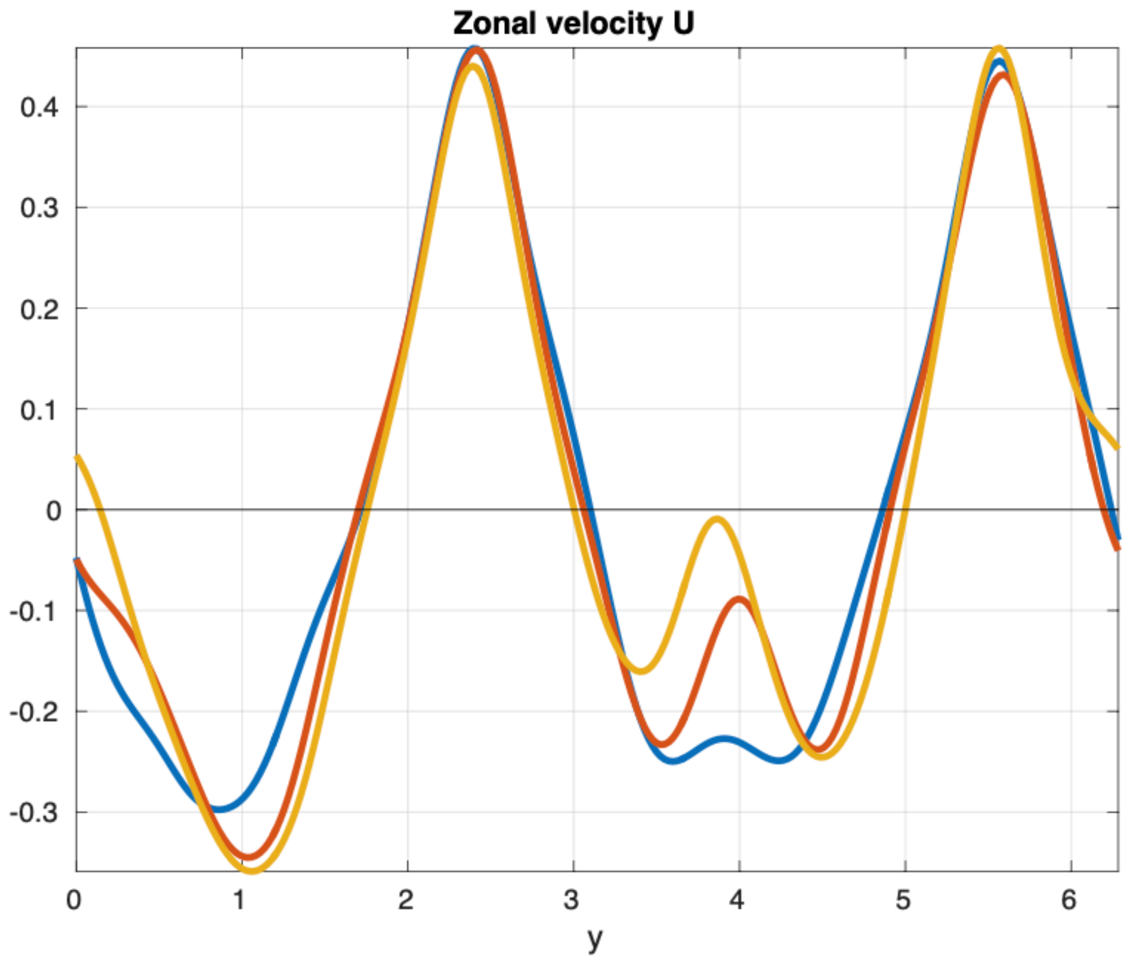
AMS: an algorithm to compute rare events, for instance rare reactive trajectories

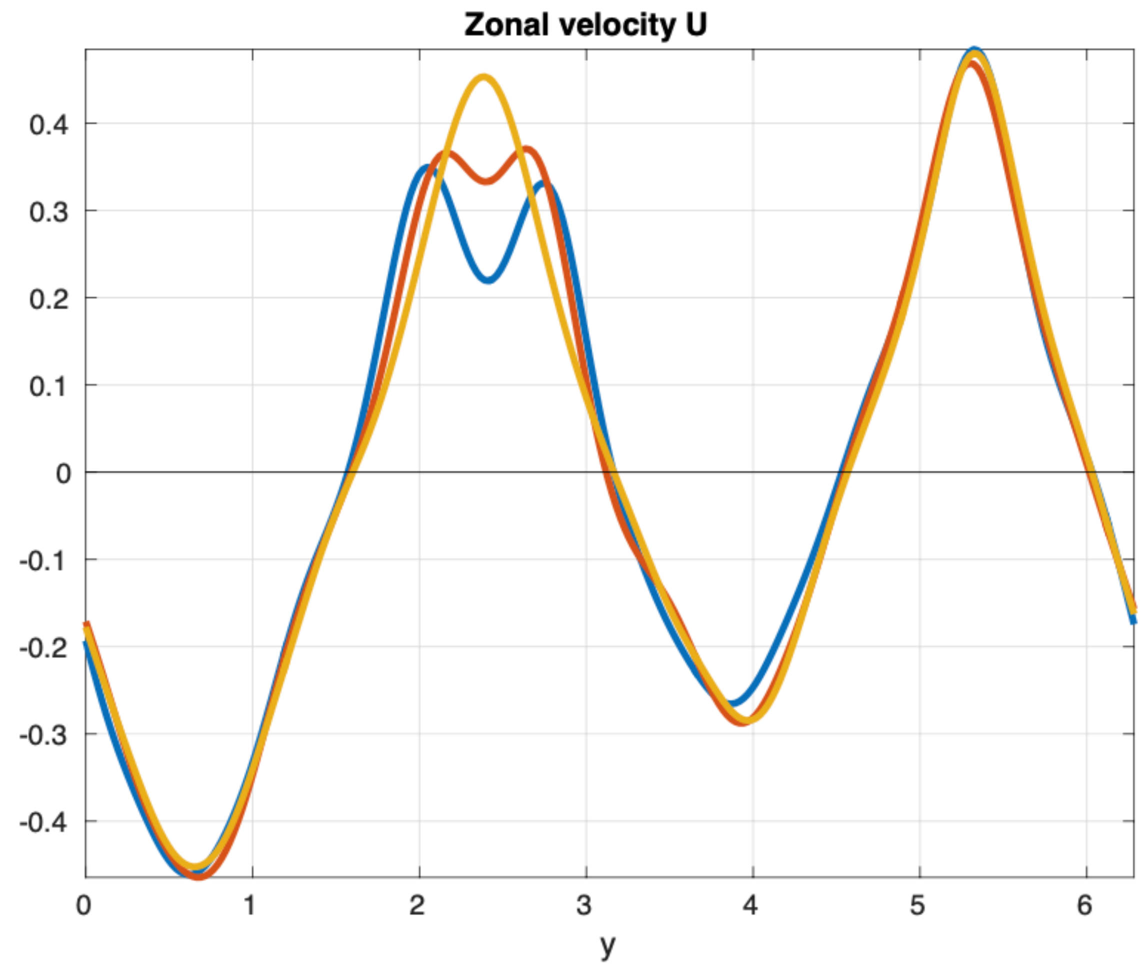


Transition trajectories between 2 and 3 jet states

- The dynamics of turbulent transitions is predictable.
- Asymmetry between forward and backward transitions.

Evolution of Velocity Fields During the Transition





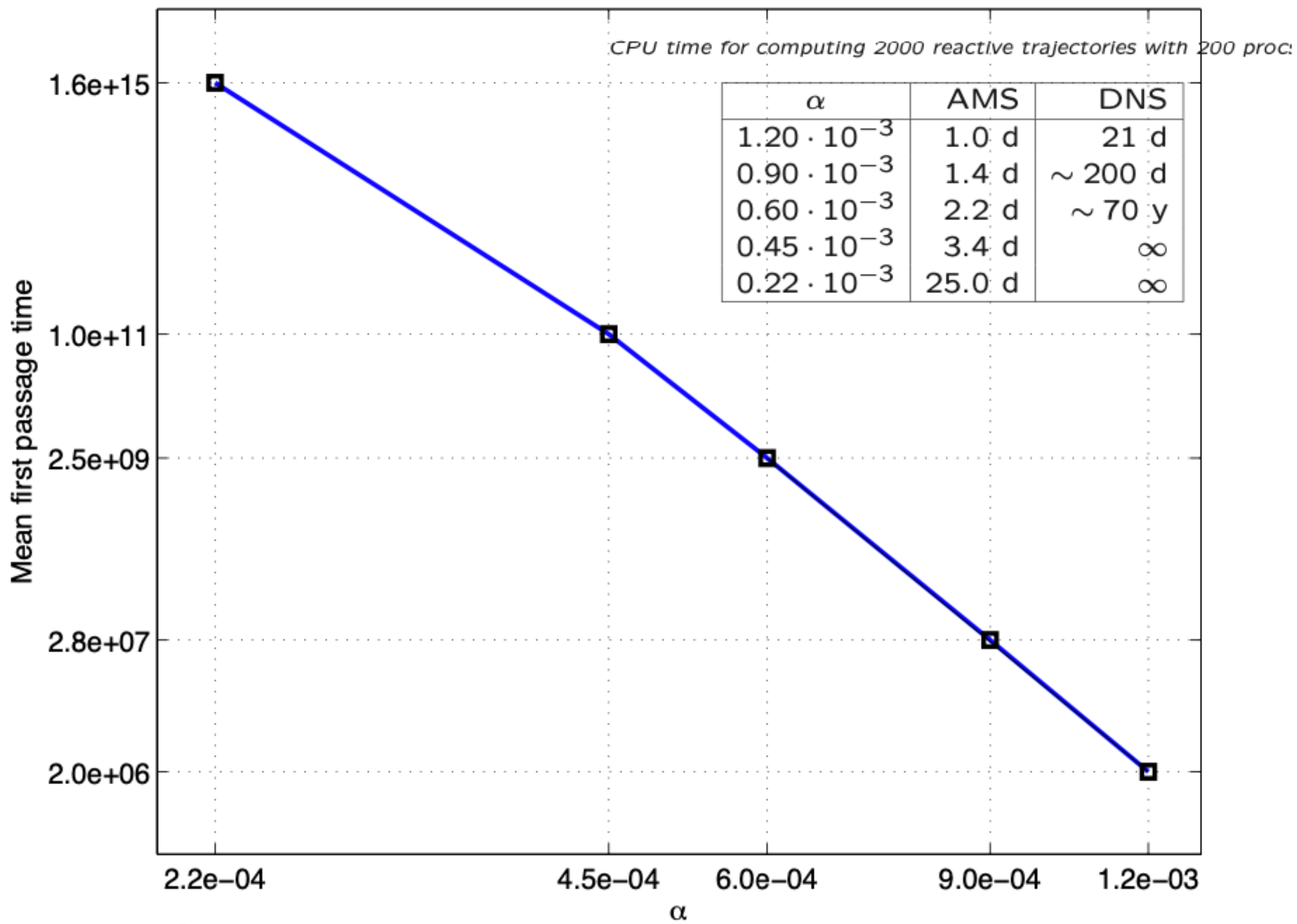
Nucleation of a new jet

Merging of two jets

- Asymmetry between forward and backward transitions.

Transition Rates for Unreachable Regimes Through DNS

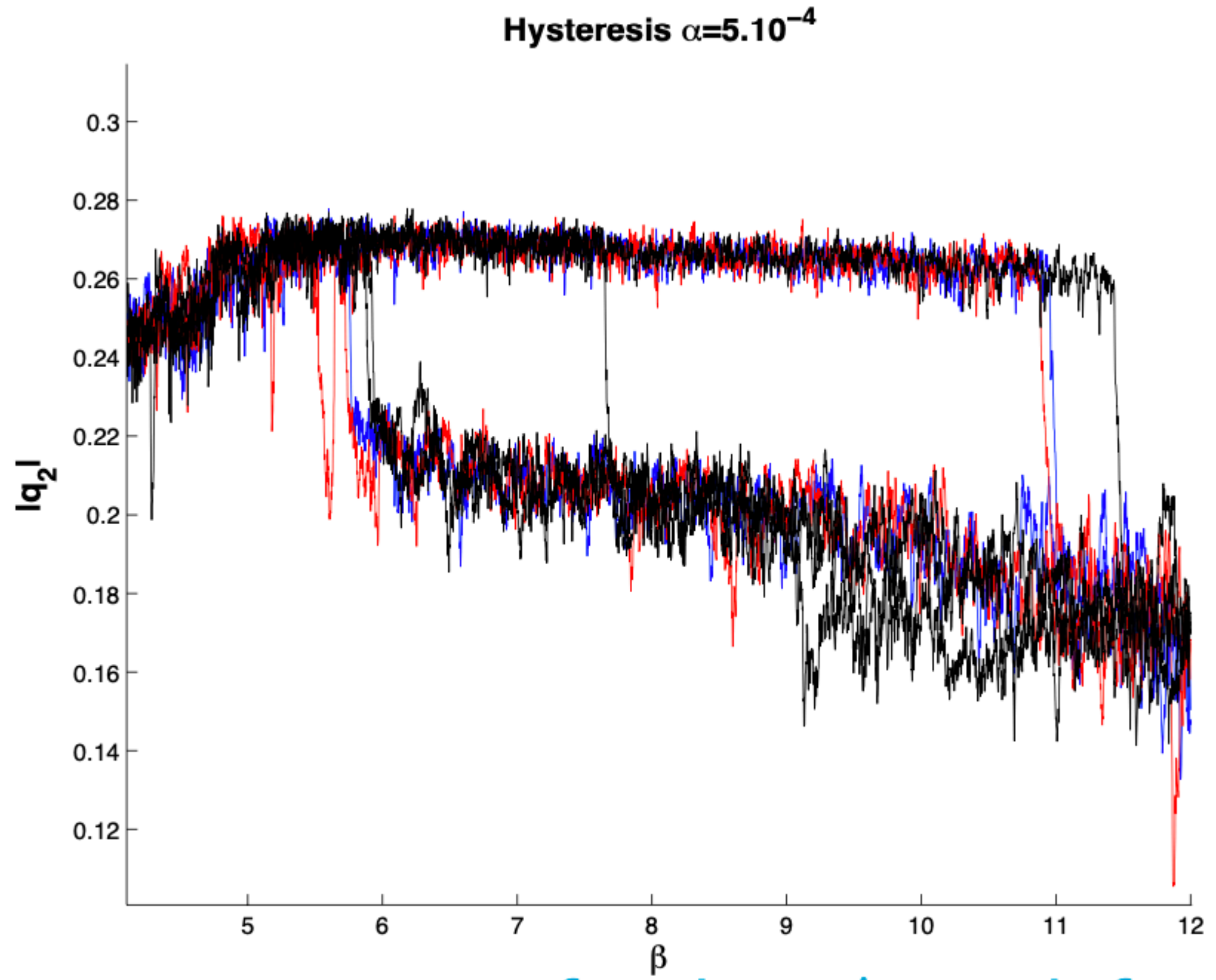
With the AMS we can estimate huge average transition times



Average transition time versus α

- With the AMS algorithm, we study transitions that would require an astronomical computation time using direct numerical simulations.

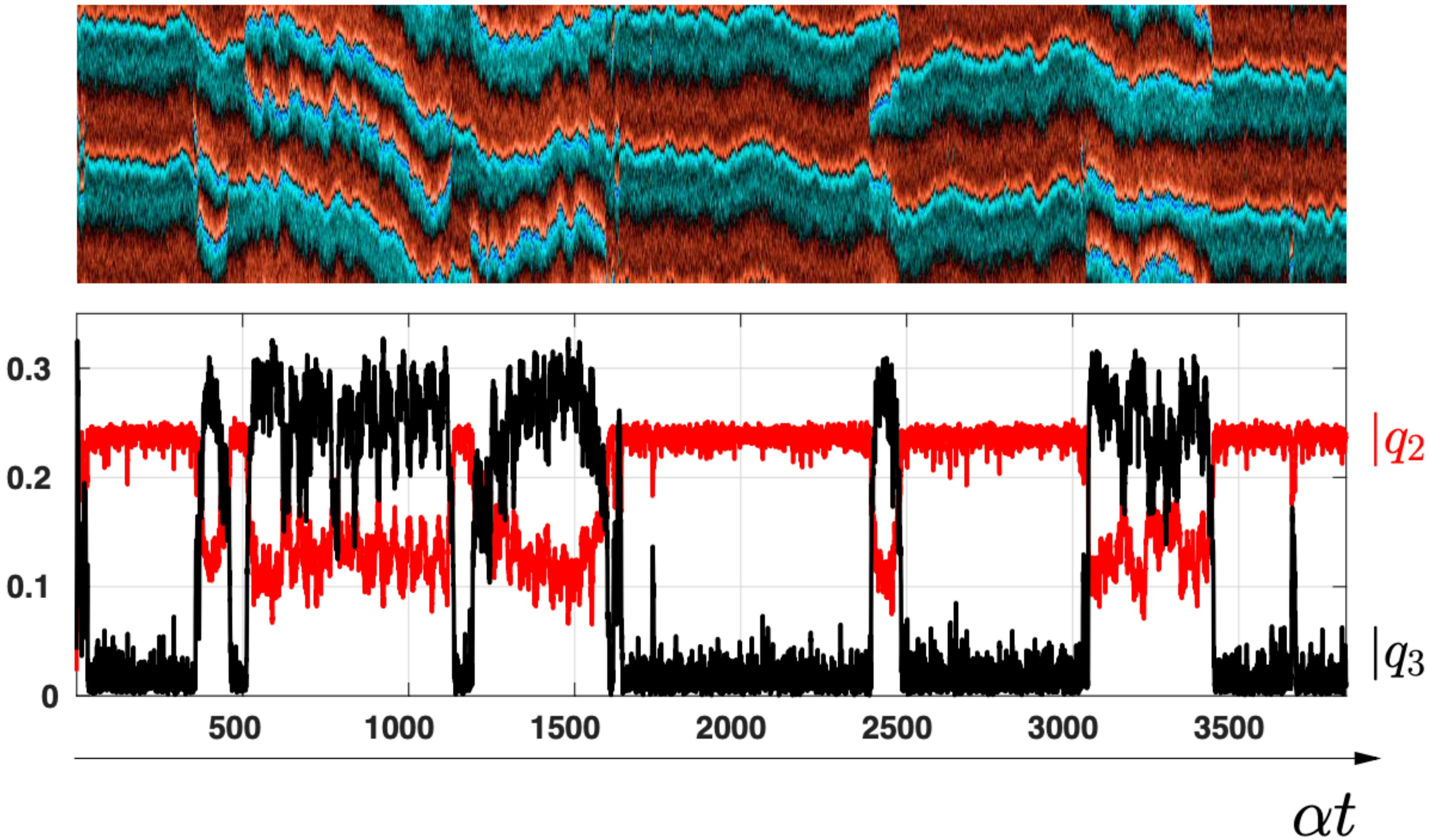
A Complex Internal Dynamics for the 3-Jet States



Hysteresis experiment for the 2/3 jet bifurcations

- The 3 jet states have larger fluctuations than the 2 jet states.

Rare Transitions Between Quasigeostrophic Jets

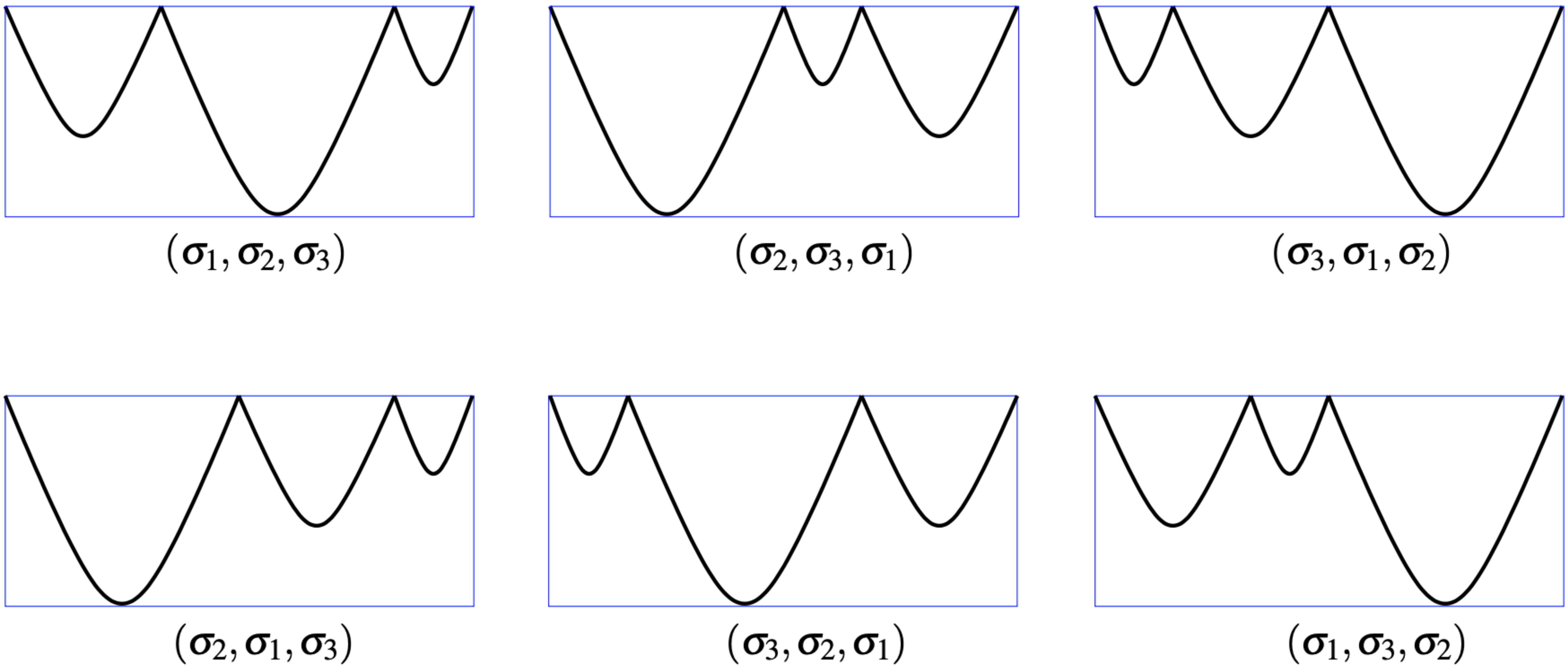


Rare transitions for quasigeostrophic jets

- It seems that the 3 jet states might have different structures.

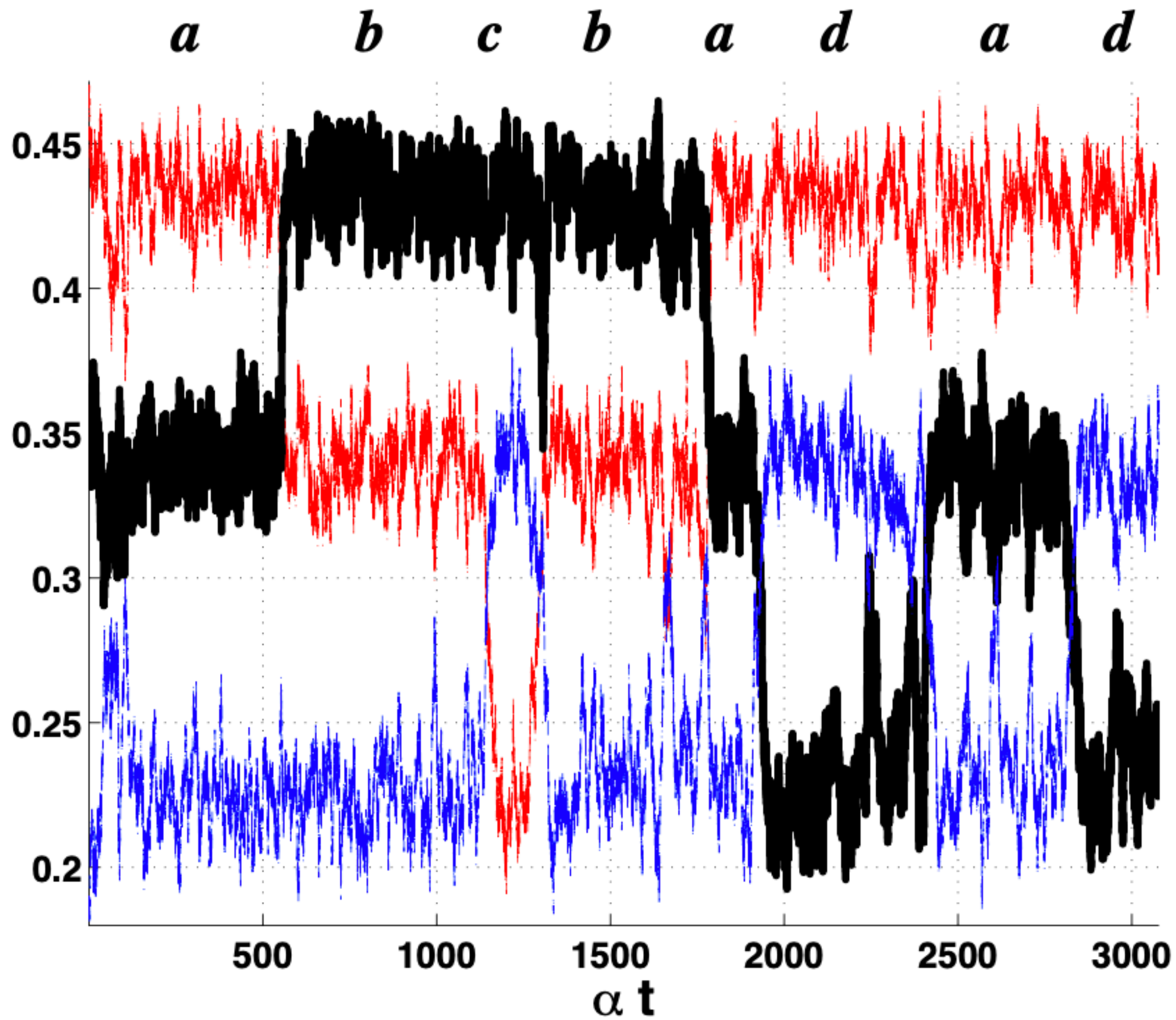
A Family of Different 3-Jet Attractors

Symmetry breaking within the set of 3-jet attractors



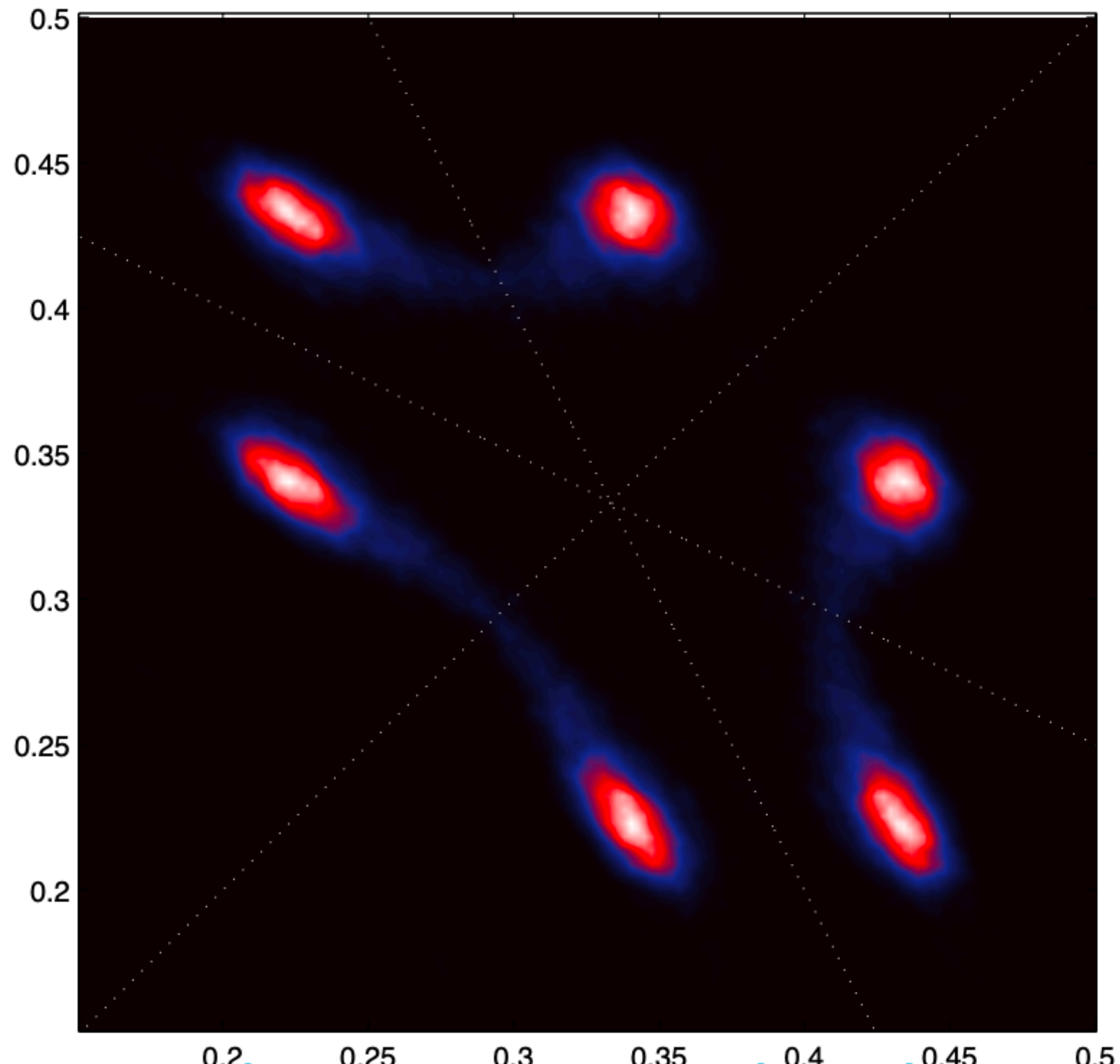
Schematic zonal velocity fields $U(y)$ for the 3-jet attractors

Internal Multistability for the 3-Jet Attractors



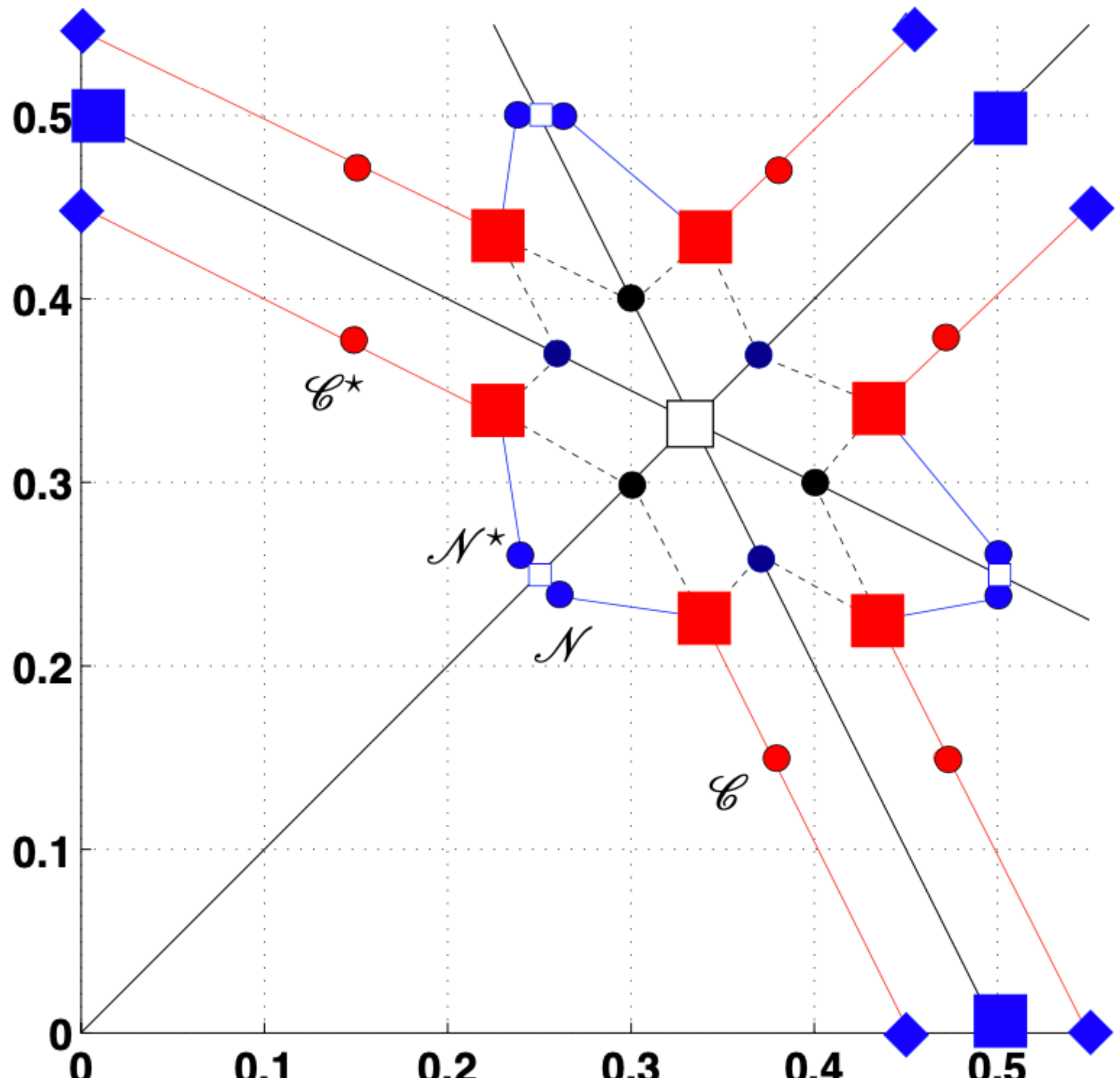
Timeseries for the distance between jets within the 3-jet attractors

Internal Multistability for the 3-Jet Attractors



PDF of distances between jets within the 3-jet attractors.

Bifurcation Diagram for the 3-Jet Attractors



Attractors and saddles for the 3 jet states

Each axe represent one of the 3 distances between the 3 jets.

Conclusions

- We have computed rare transitions between zonal jets, similar to Jupiter's abrupt climate changes, that can not be computed using direct numerical simulations (with E. S.).
- We have partial results for the justification of averaging (ergodicity, etc ...), (with C.N., and T.T.).
- For small scale forces, the average Reynolds stress can be computed explicitly and is universal. We have a good qualitative agreement with Jupiter's jets. (with E.W.).
- The rare transitions involve non-Gaussian fluctuations of the Reynolds stress. (with T.G., T.T., and E. V-E).

http://perso.ens-lyon.fr/freddy.bouchet/

Rare event algorithms for climate dynamics

Outline

I) Introduction: rare events do matter - rare event algorithms

II) Rare events algorithms for predicting Jupiter's abrupt climate change

III) Rare events algorithms for predicting extreme heat waves

IV) Coupling rare event algorithms with machine learning

V) Predicting extreme heat waves and committor functions using deep neural networks

III) Rare event algorithms to study extreme heat waves with climate models

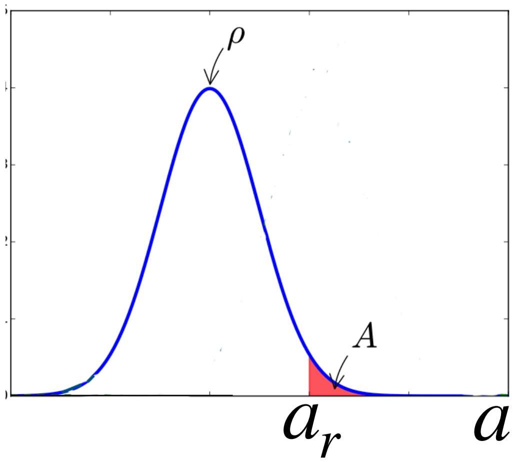
Francesco Ragone
RMI, Bruxelles, Belgium

Jeroen Wouters
University of Reading, UK

Long lasting summer heat waves

We will study extremes of the time averaged temperature:

$$a = \frac{1}{T}\int_0^T \mathrm{d}t \frac{1}{|\mathscr{A}|} \int_{\mathscr{A}} \mathrm{d}\mathbf{r}\ \mathrm{T_S}(\mathbf{r}, t)$$



- $\mathscr{A} =$ Scandinavia, Europe, France, Alberta, Russia, …
- $T =$ one week, a few weeks, a month, or a season.
- **Climate models** (CESM or PLASIM) or reanalysis datasets.

The Giardina—Kurchan (Del-Moral—Garnier) rare event algorithm

- With $A[X](t) = \dfrac{1}{|\mathscr{A}|}\displaystyle\int_{\mathscr{A}} \mathrm{d}\mathbf{r}\ \mathrm{T_S}(\mathbf{r}, t)$, we sample the tilted path-distribution

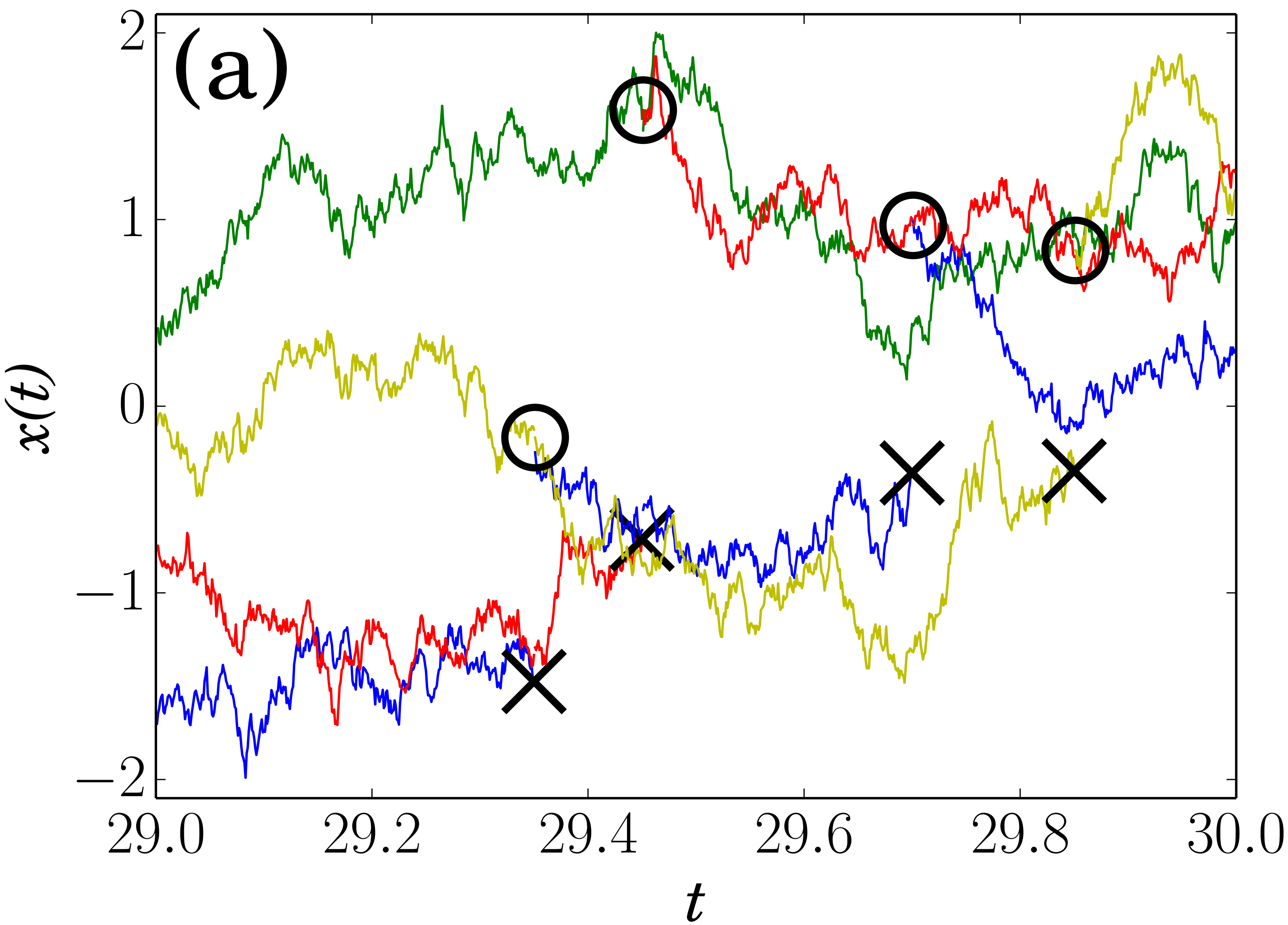
$$\tilde{P}_k\left(\{X(t)\}_{0\le t\le T}\right) = \frac{1}{\exp(T\lambda(k))} P_0\left(\{X(t)\}_{0\le t\le T}\right) \exp\left[k\int_0^T A[X](t)\,\mathrm{d}t\right].$$

- We simulate an ensemble of N trajectories $x_n(t)$. At each time step $t_i = i\tau$, each trajectory can be killed or cloned according to the weights

$$\frac{1}{W_i(k)}\exp\left(k\int_{t_{i-1}}^{t_i} A[x_n](t)\,\mathrm{d}t\right) \quad \text{with} \quad W_i(k) = \sum_{n=1}^{N} \exp\left(k\int_{t_{i-1}}^{t_i} A[x_n](t)\,\mathrm{d}t\right).$$

- Algorithm: Giardina et al. 2006. Mathematical aspects: Del Moral's book (2004).

Genealogical algorithm: selecting, killing and cloning trajectories

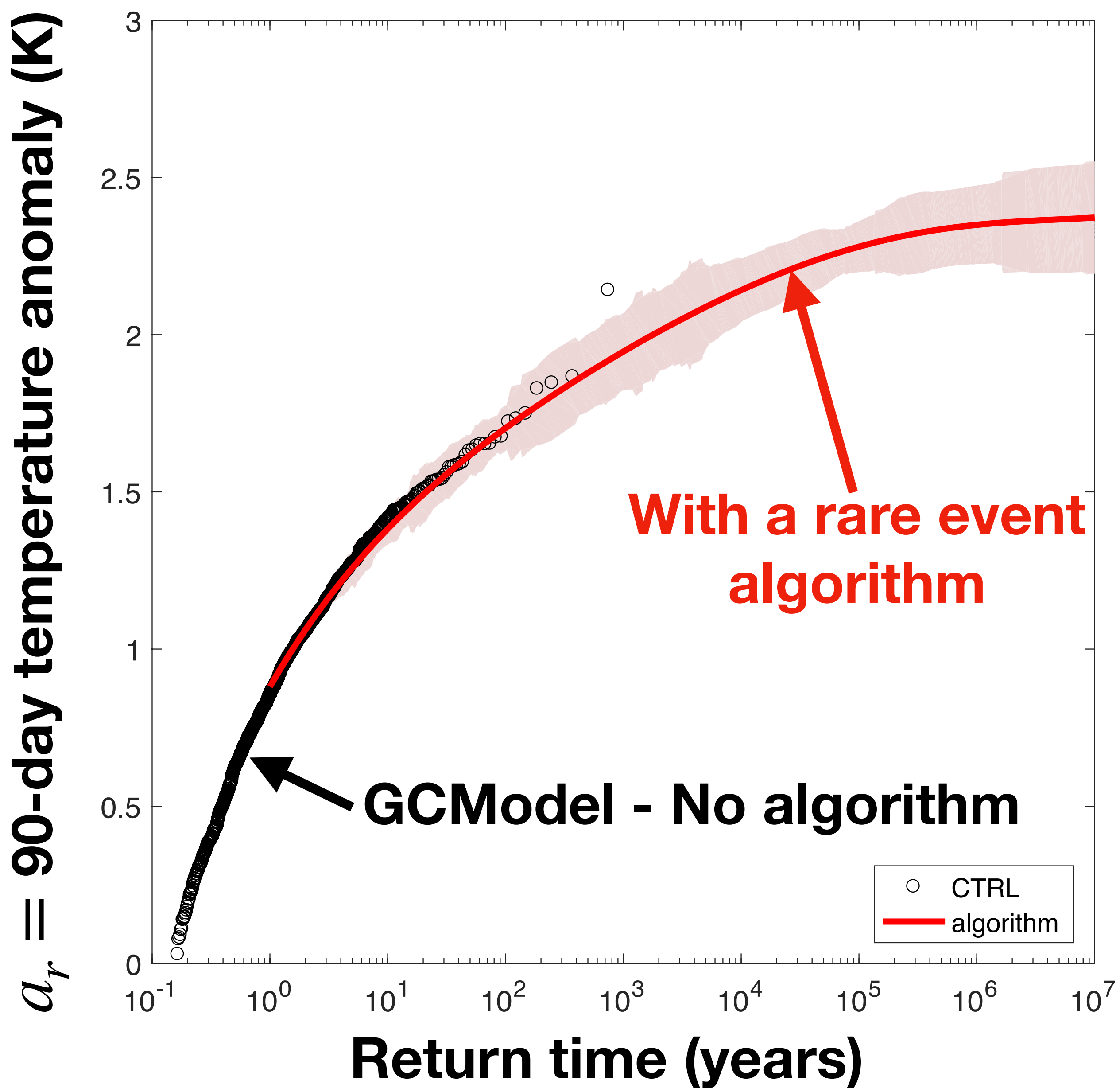


The trajectory statistics is tilted towards the events of interest.

Sample paths of the Giardina Kurchan algorithm

(from Bouchet, Jack, Lecomte, Nemoto, 2016)

Return time plot computed using a rare event algorithm (PLASIM)



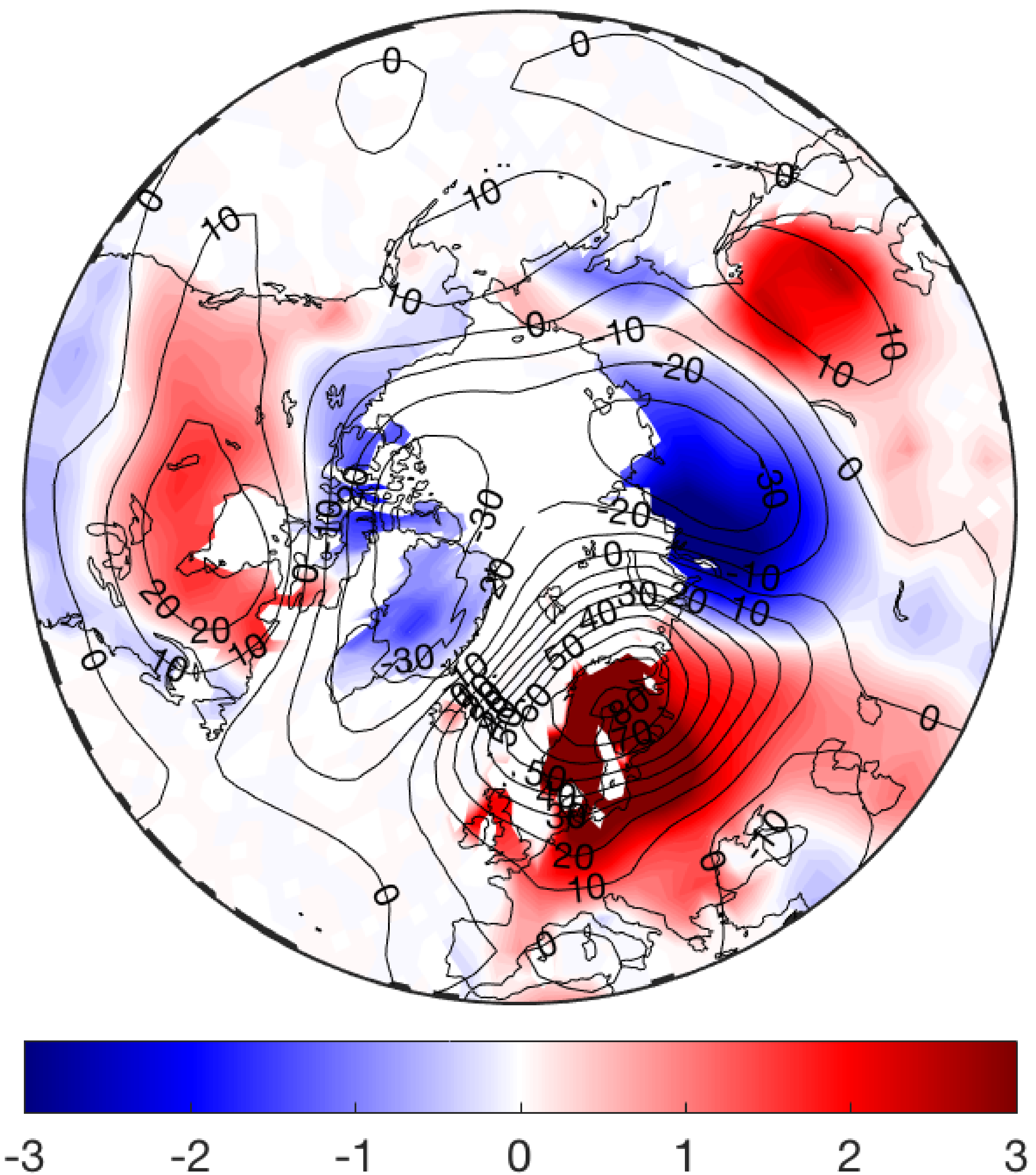
Extremes of 90-day Europe heat waves

PLASIM model.
No seasonal cycle.
Del-Moral—Garnier (or Giardina—Kurchan) algorithm.

F. Ragone, J. Wouters, and F. Bouchet, PNAS, 2018

At a fixed numerical cost, we can study events which are several orders of magnitude rarer.

Extreme teleconnection pattern



Extreme teleconnection patterns = conditional averages with

$$\frac{1}{T}\int_0^T \mathrm{d}t \frac{1}{|\mathscr{A}|}\int_{\mathscr{A}} \mathrm{d}\mathbf{r}\ \mathrm{T_S}(\mathbf{r},t) > 2\ \mathrm{K}$$

and $T = 40$ days.

Plasim model.
Summer Scandinavian heat waves.

F. Ragone, J. Wouters, and F. Bouchet, PNAS, 2018

500 hPa geopotential height and temperature anomalies

Extreme teleconnection patterns differ from teleconnections for typical fluctuations and are not characterized by a single wavenumber but are much constrained by geography.

Rare event algorithms for climate dynamics
Outline

I) Introduction: rare events do matter - rare event algorithm

II) Rare events algorithms for predicting Jupiter's abrupt climate change

III) Rare events algorithms for predicting extreme heat waves

IV) Coupling rare event algorithms with machine learning

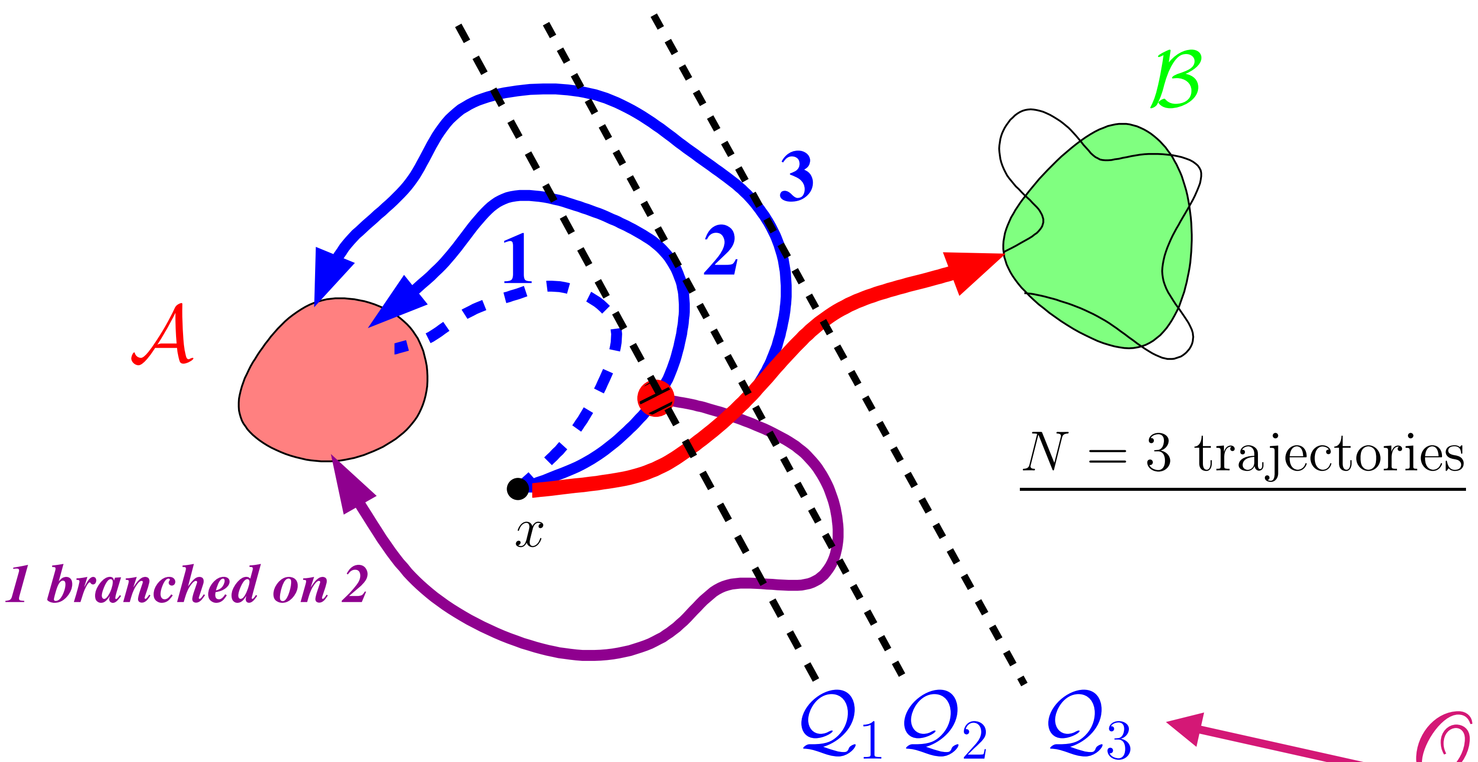
V) Predicting extreme heat waves and committor functions using deep neural networks

IV) Coupling Rare Event Algorithms with Machine Learning

IV.a) The challenge to have efficient rare event algorithms: estimating good score functions for complex dynamics

The Adaptive Multilevel Splitting (AMS) rare event algorithm

Strategy: selection, pruning and cloning.



Probability estimate:

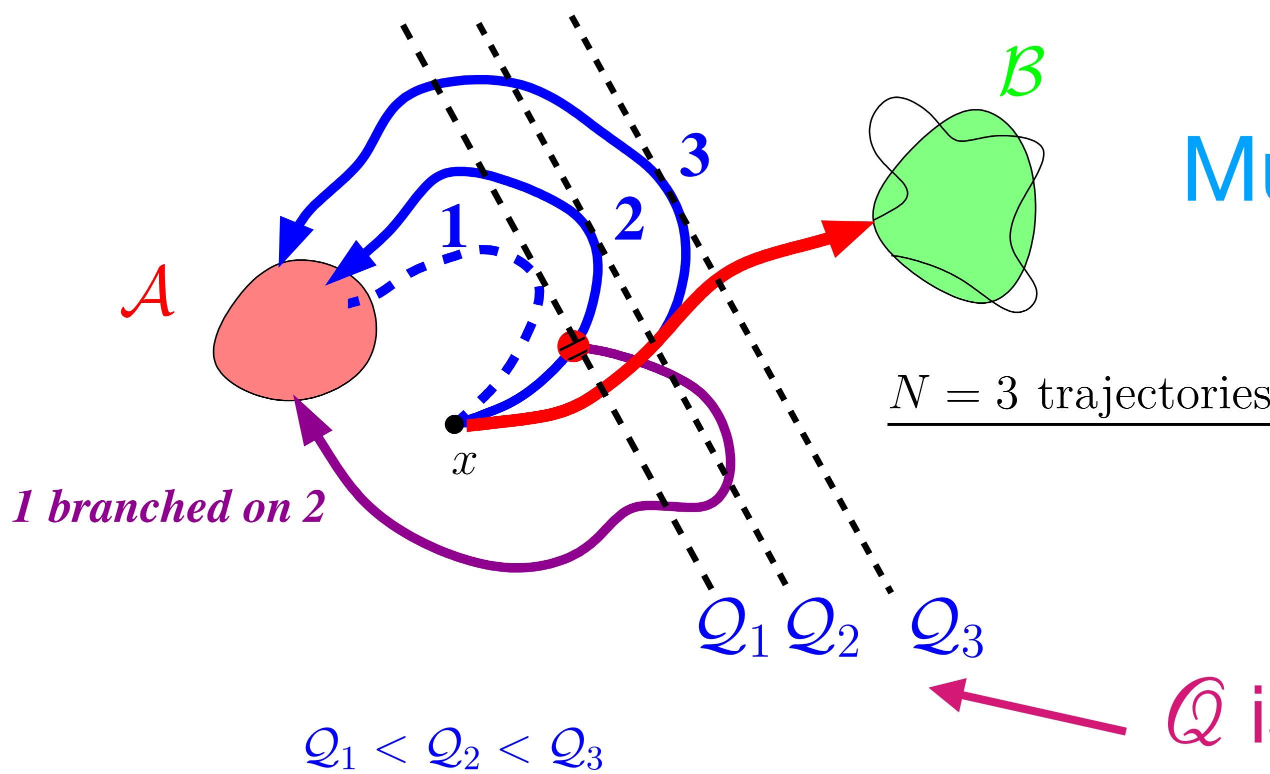
$$\hat{p} = (1 - 1/N)^K,$$

where N is the clone number and K is the iteration number.

$\mathcal{Q}$ is the score (selection) function

Cérou, Guyader (2007). Cérou, Guyader, Lelièvre, and Pommier (2011).
PDEs: Rolland, Bouchet et Simonnet (2016) - TAMS: Lestang et al (2018)
Atmosphere turbulent jets: Rolland, Bouchet et Simonnet (2019 and 2021).

Committor functions are optimal score functions for rare event algorithms



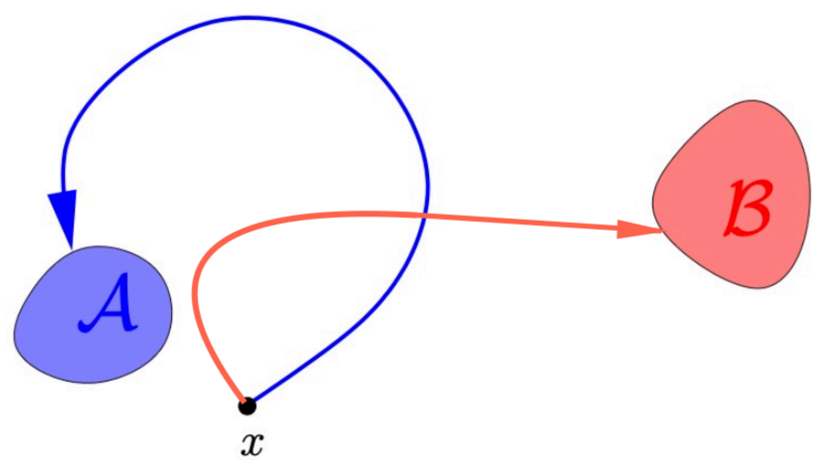
The Adaptive Multilevel Splitting algorithm

$\mathcal{Q}$ is the score (selection) function

The efficiency of the algorithm depends on the choice of the score function.

The optimal score function is the committor function.

Committor function



- $\{X(t)\}_{-\infty \leq t < +\infty}$ is a Markov process. A, B are subsets of the phase space.

- For a given sample path $\{X(t)\}_{-\infty \leq t < +\infty}$, the first hitting time τ_A is $\tau_A = \inf\{t \,|\, X(t) \in A\}$.

- The committor function *q(x)* of the sets A and B is defined as the probability that a trajectory starting at the point *x* reaches the set *B* before the set *A*

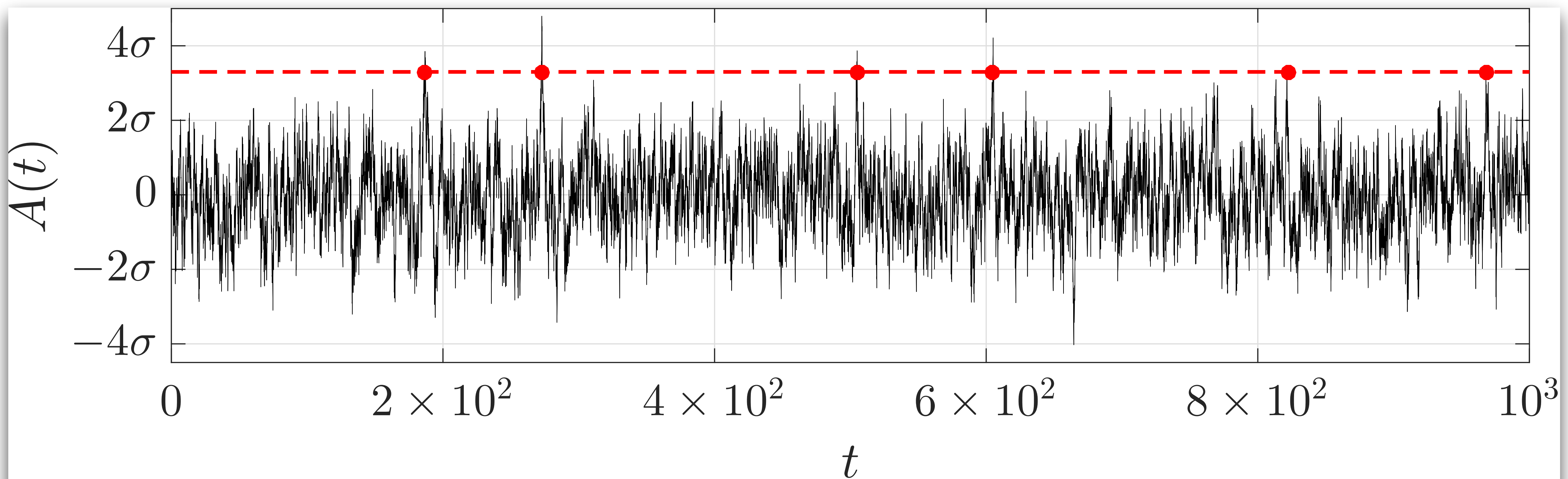
$$q(x) = \mathbb{P}_x\left(\tau_B < \tau_A\right).$$

- How to estimate the committor function? With a rare event algorithm!

Committor function for extreme events

- We define the committor function $q(x, \tau)$ as the probability that a given observable $A(X(t))$ becomes greater than a threshold a at a certain lag time τ given that $X(0) = x$:

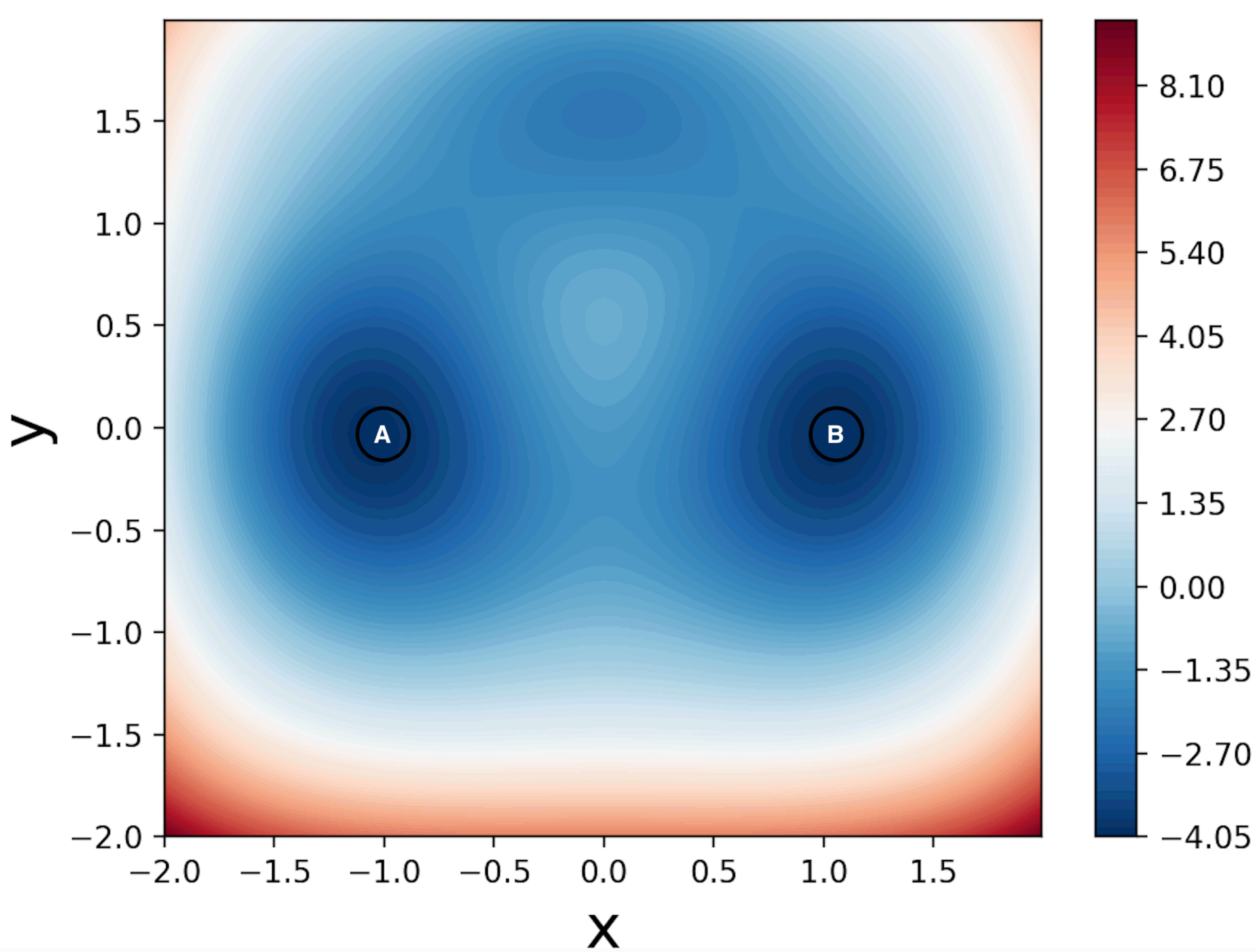
$$q(x, \tau) = \mathbb{P}(A(X(\tau)) > a \,|\, X(0) = x)\,.$$



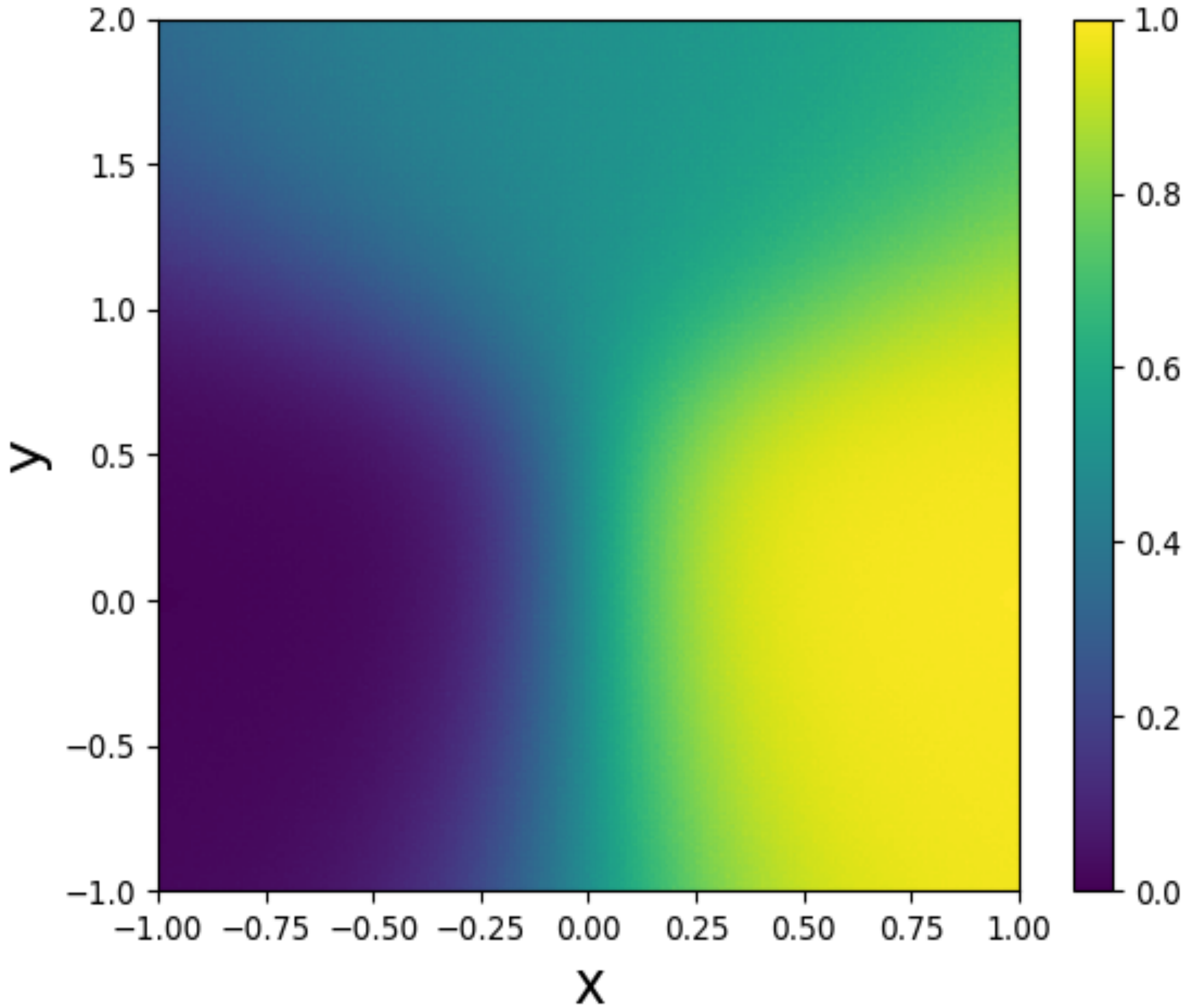
Committor function for a simple gradient dynamics

$$\mathrm{d}X_t = -\nabla V(X_t)\mathrm{d}t + \sqrt{2\epsilon}\,\mathrm{d}W_t \quad \text{and} \quad X = (x, y)$$

Potential V



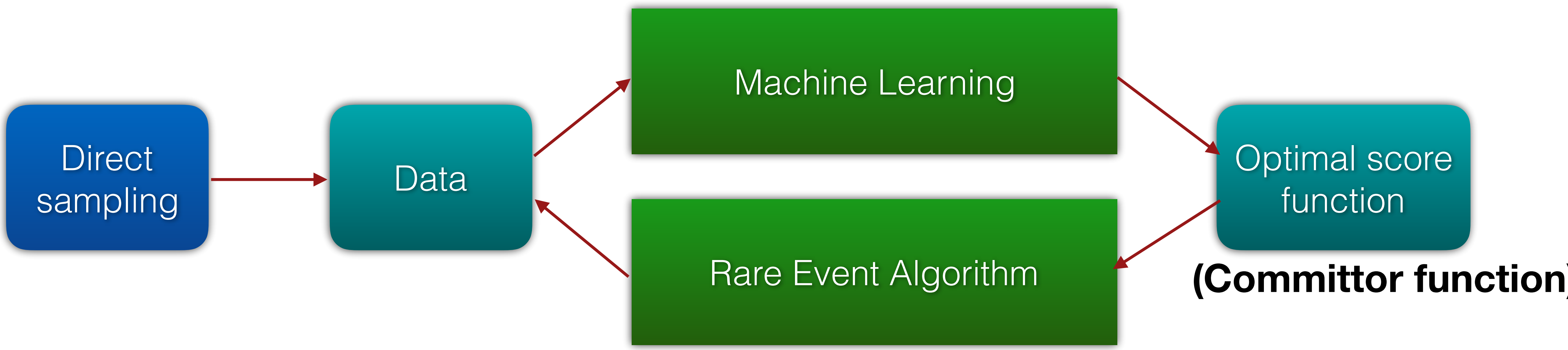
Committor function q



The score function is the key practical problem

- With a poor score function, rare event algorithms are useless.
- How to build good score functions?
- Running a rare event algorithm !!
- One needs one algorithm to improve the algorithm efficiency: use an adaptive strategy.
- Examples: Wang Landau algorithm, multi canonical methods, adaptive importance sampling, etc.

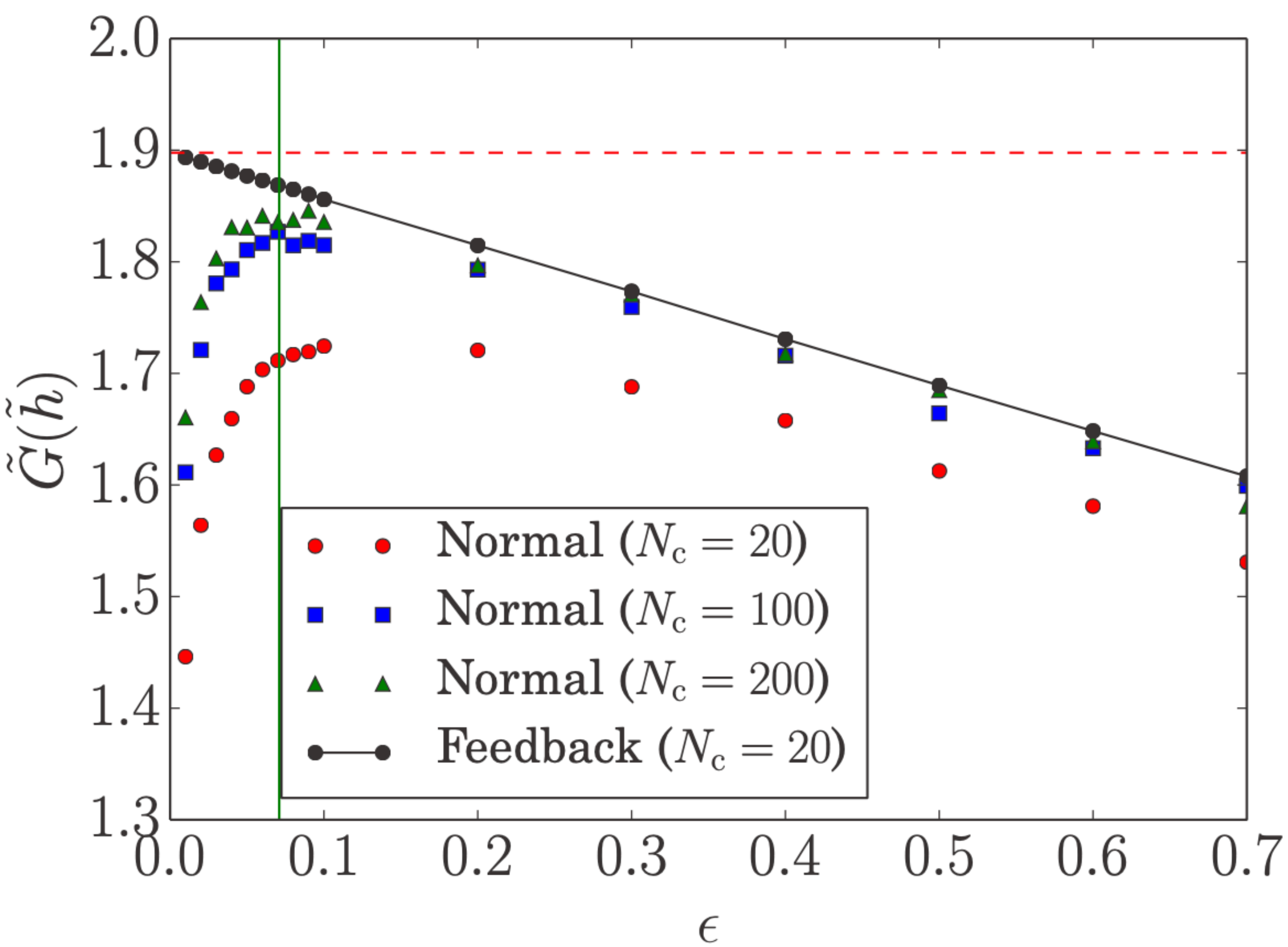
Coupling rare event algorithms with machine learning of committor functions

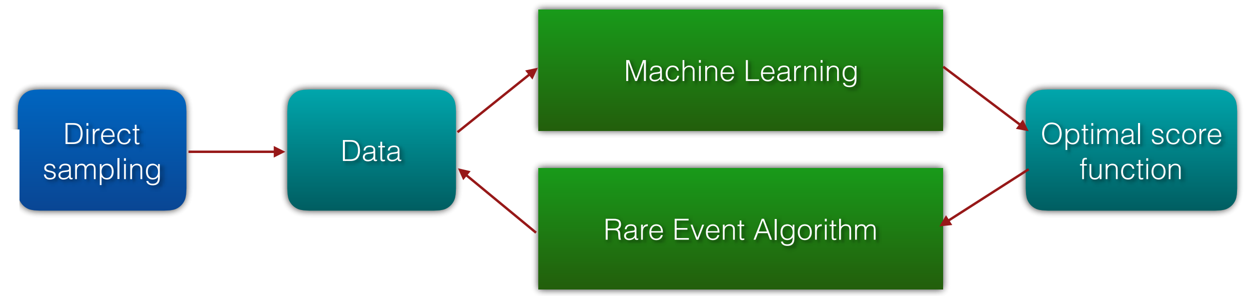


One example: Bouchet, Jack, Lecomte, Nemoto, PRE, 2016

(For X in dimension 1)

Coupling rare event algorithms with data-based learning of committor functions





$$\mathrm{d}X_t = -X_t^3\mathrm{d}t + \sqrt{2\epsilon}\,\mathrm{d}W_t,$$

with X_t in dimension 1.

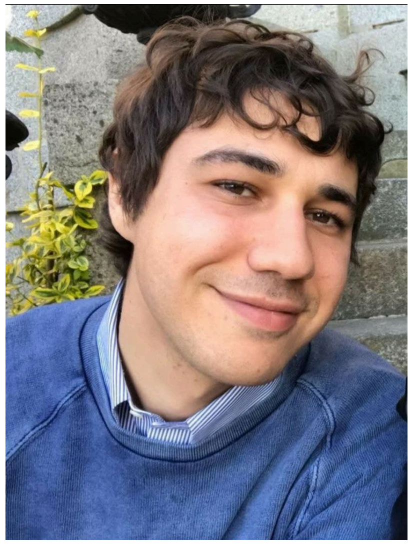
$$G(h) = \lim_{T\to+\infty} \frac{1}{T} \log \mathbb{E}\left[\exp\left(\int_0^T \lambda(X_t)\mathrm{d}t\right)\right]$$

with $\lambda(x) = x(x+1)$

Bouchet, Jack, Lecomte, Nemoto, PRE, 2016

IV-b) Predictions at the predictability margin and committor functions for a model of ENSO (El Niño)

With Dario Lucente, Corentin Herbert, Stefan Duffner and Joran Rolland



Dario Lucente

Conclusions: Predictions at the Predictability Margin and Committor Functions

- At the predictability margin, predictions should be probabilistic. **The committor functions are the proper mathematical objects.**

- For a simple dynamics of El Niño, which is a small stochastic perturbation of a chaotic deterministic system, we have computed a committor function for a transition to occur.

- The committor functions shows **areas of the phase space with hard, respectively easy, probabilistic predictability potential** and quantifies the probability of the event.

- This informs us on what to expect for a predictability problem at the predictability margin, **in a perfect information context.**

D. Lucente, S. Duffner, C. Herbert, J. Rolland, and F. Bouchet, proceeding of Climate Informatics 2019.

D. Lucente, C. Herbert, and F. Bouchet, to be published by Journal of the Atmospheric Sciences, 2022.

IV.c) Coupling rare event simulations and machine learning for dynamics with few degrees of freedom

Conclusion: Coupling machine learning with rare event algorithms

- We can learn committor functions from dynamical datasets either using the definition, or first learning an approximate Markov dynamics.

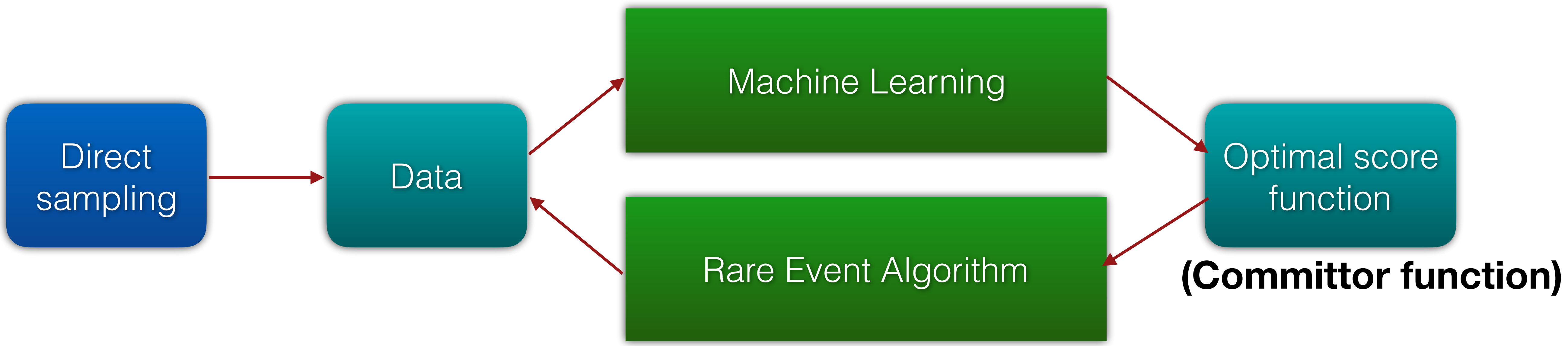
- The analogue Markov chain does not require an impossible discretization of the phase space, and can use any kind of dynamical data, including short trajectories.

- Using learned committor functions is much more efficient than using user-defined score functions with the AMS rare event algorithm.

- The range of applicability of this approach, in terms of system dimension and complexity, is a key question for the future.

D. Lucente, J. Rolland, C. Herbert and F. Bouchet, ArXiv, submitted to. JSTAT

Coupling rare event algorithms with machine learning of committor functions



Work in progress for climate models!

Rare event algorithms for climate dynamics
Outline

I) Introduction: rare events do matter - rare event algorithms

II) Rare events algorithms for predicting Jupiter's abrupt climate change

III) Rare events algorithms for predicting extreme heat waves

IV) Coupling rare event algorithms with machine learning

V) Predicting extreme heat waves and committor functions using deep neural networks

III) Predicting extreme heat waves and committor functions using deep neural networks

With P. Abry, P. Borgnat, V. Jacques-Dumas, G. Miloshevich, and F. Ragone

Valerian Jacques-Dumas

George Miloshevich

Machine learning, climate, and weather forecast models

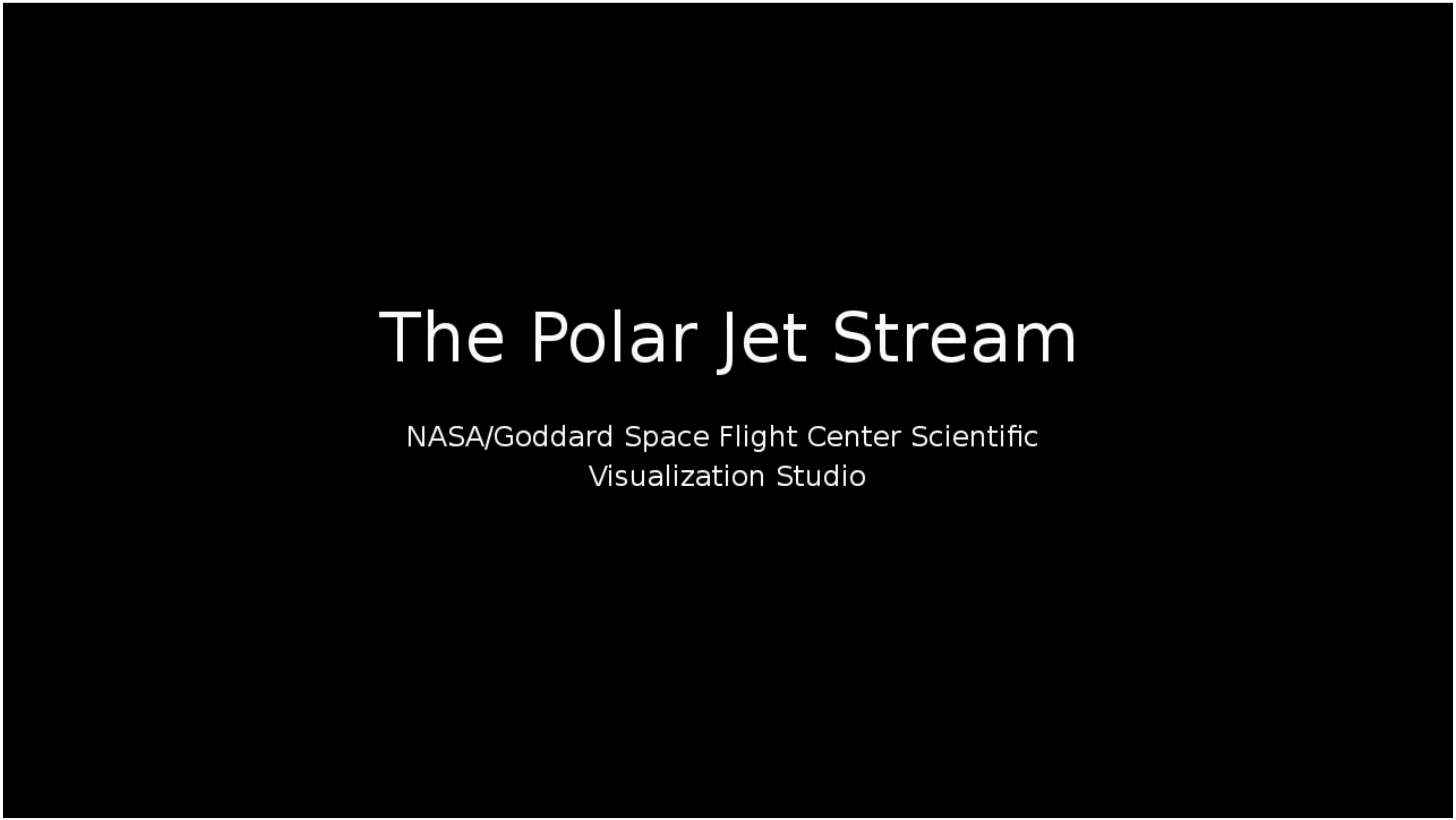
- The Earth (atmosphere, ocean, land, etc.) is the most observed system with an exponentially growing dataset.

- Those observations are coupled to physical models through data assimilation techniques (a very old and very smart machine learning scheme for physically based data integration).

- Machine learning and deep neural networks enter in many different ways for both weather forecast and climate dynamics.

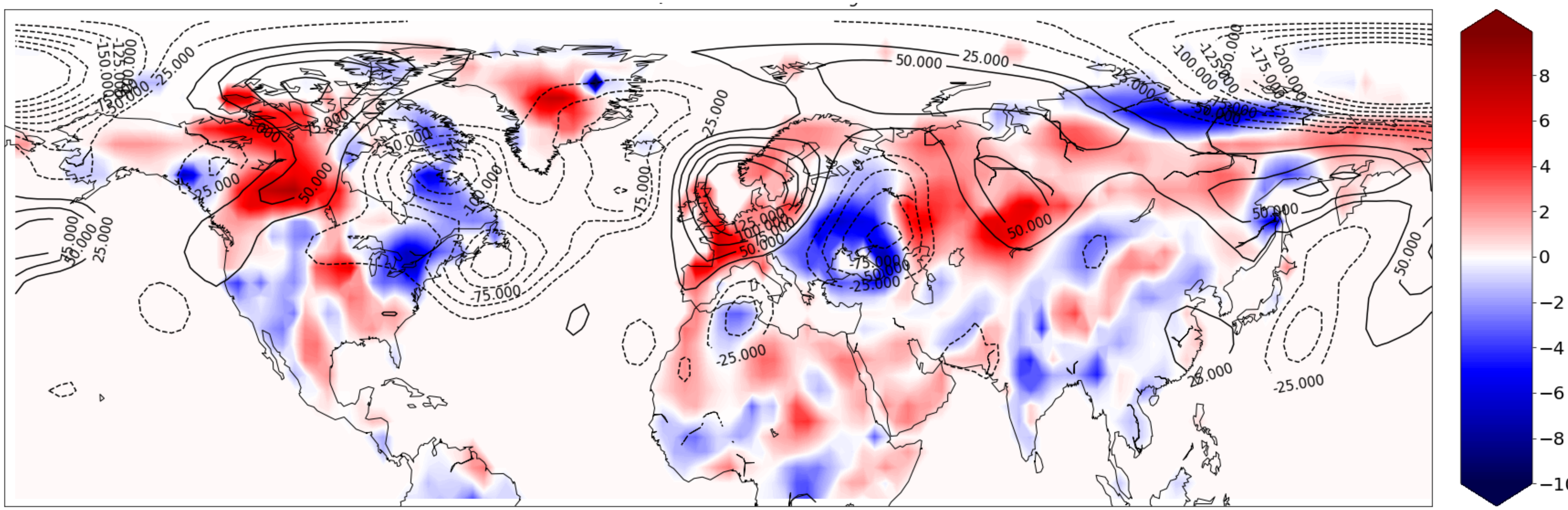
- For many (not all) of these problems, machine learning should be performed in a regime of lack of data. This is key for understanding the challenge for machine learning.

Jet stream dynamics



Higher troposphere wind speed. (NASA/Goddard Space Flight Center Scientific Visualization Studio, MERRA reanalysis dataset)

Predicting heat waves with a deep neural network - 1) Data



Surface temperature (T_s, colors) and 500 hPa geopotential height (Z_g, lines) anomalies

- **Plasim and CESM climate models.**
- We use summer (JJA) data: 8 maps/day, 90 days/year, 1000 year = 720 000 maps.
- For Plasim data, each field has a resolution 64×128, restricted to 25×128 above 30^o North.

Heat wave definition

- $X(t) = T_s$ field at time t, or $X(t) = (T_s, Z_g)$ fields at time t.

- $Y(t)$: **time and space averaged surface temperature anomaly within τ days:**

$$Y(t) = \frac{1}{T}\int_{t+\tau}^{t+\tau+T} \frac{1}{|\mathcal{A}|}\int_{\mathcal{A}} T_s(\vec{r}, u)\ \mathrm{d}\vec{r}\ \mathrm{d}u,$$
$$\text{and } Z(t) = 1 \text{ if } Y(t) > a, \text{ and } Z(t) = 0 \text{ otherwise}$$

- $Z(t) \in \{0,1\}$. A heat wave occurs if $Z = 1$.

- We have a classification problem for the data (X, Z). We want to learn the probability $q(x)$ that $Z = 1$ given that $X = x$ (committor function).

- 5% most extreme events: $a = a_5 = 3.08\ \mathrm{K}$. 2.5% most extreme events: $a = a_{2.5} = 3.7\ \mathrm{K}$. 1.25% most extreme events: $a = a_{1.25} = 4.23\ \mathrm{K}$.

Predicting heat waves with a deep neural network

Observing the temperature and geopotential height at 500 hPa today, what is the probability to observe a T-day heat wave starting τ days from now?

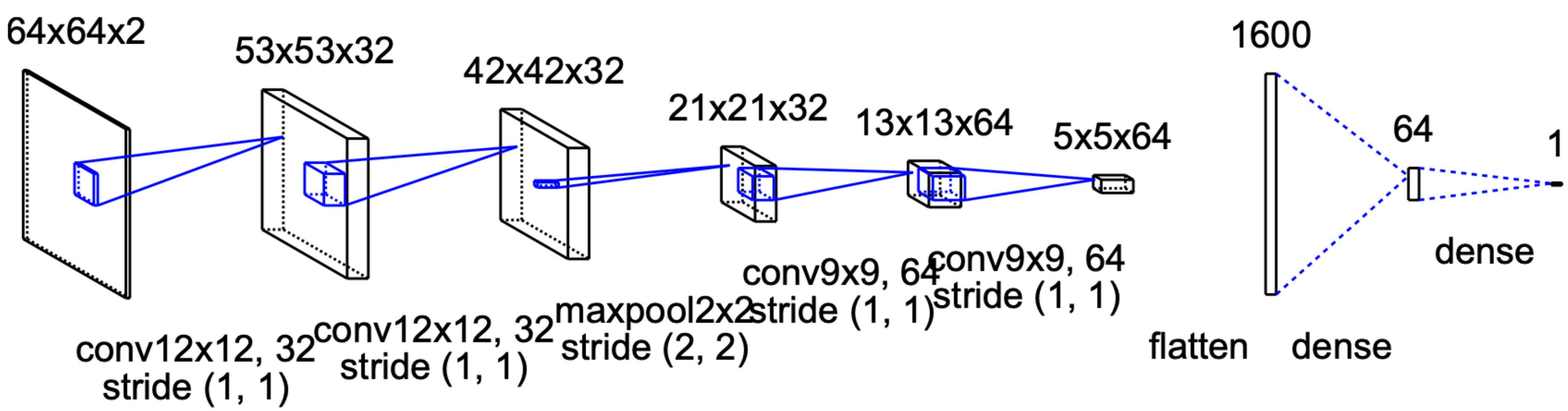
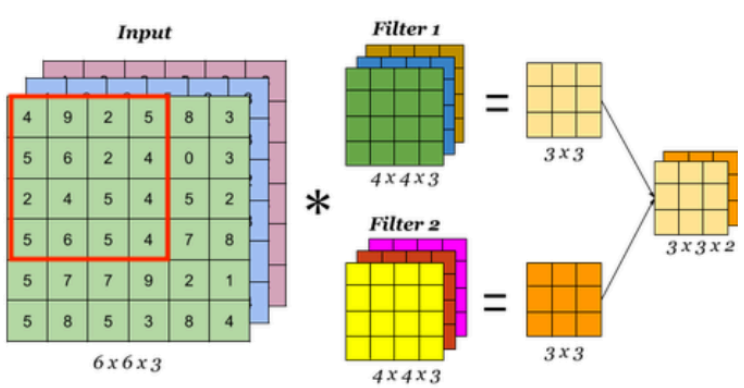
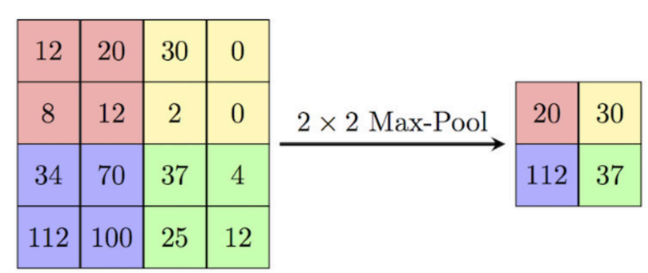


Figure 2: Architecture of the CNN used to forecast extreme heatwaves.



(a) A Convolution layer

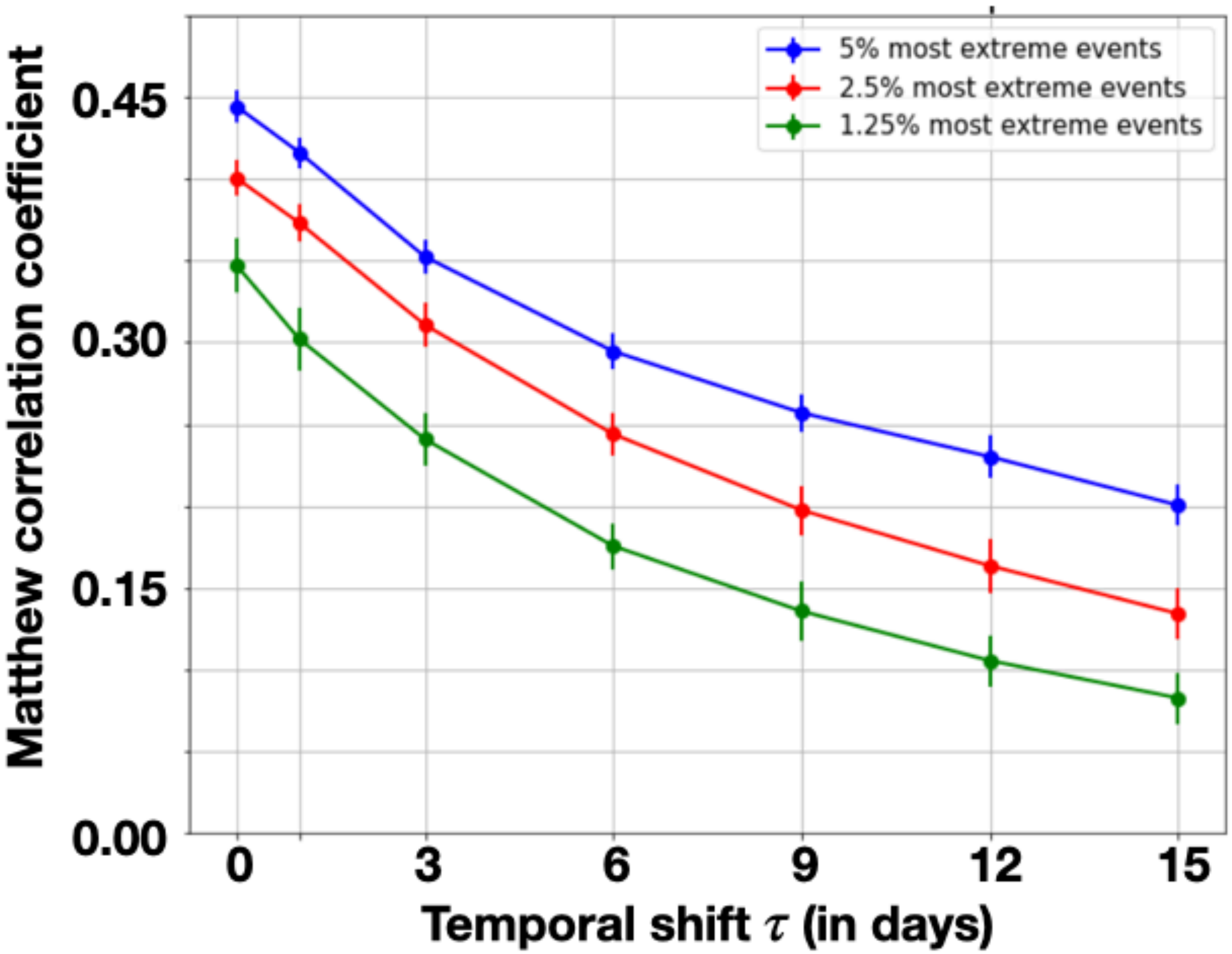


(b) A MaxPool layer

Machine learning for extreme heat waves

- **Supervised learning** from 1,000 years of climate model data (720 000 couples (X, Z)).
- We use **undersampling** to deal with class imbalance.
- We use **transfer learning** between return levels a, first training a deep neural network for less rare events, and then transferring to learn rarer events with less data.

Predicting heat waves



Heat waves over France
$T = 14$

Predictability, τ day ahead, for a 14-day heatwave from the temperature and GPH fields

We have very interesting prediction capabilities up to 15 days ahead of time for $T = 14$-day heatwaves

V. Jacques-Dumas, F. Ragone, F. Bouchet, P. Borgnat and P. Abry, 2021, sub. to IEEE TPAMI + ArXiv

Probabilistic versus deterministic classification

- Recognize cats from dogs with a neural network is an example of a probabilistic prediction for a deterministic classification. We predict the class, in a probabilistic way measuring the level of confidence.

- We use scores like the Mathew correlation coefficient to test the class prediction.

- Predict the occurence of heat waves τ days ahead with a neural network is an example of a probabilistic prediction for a probabilistic classification. We predict the class probability.

- We use scores like the logarithmic or the Brier score, to test the prediction of the class probability.

The Normalized and Positively oriented logarithmic score

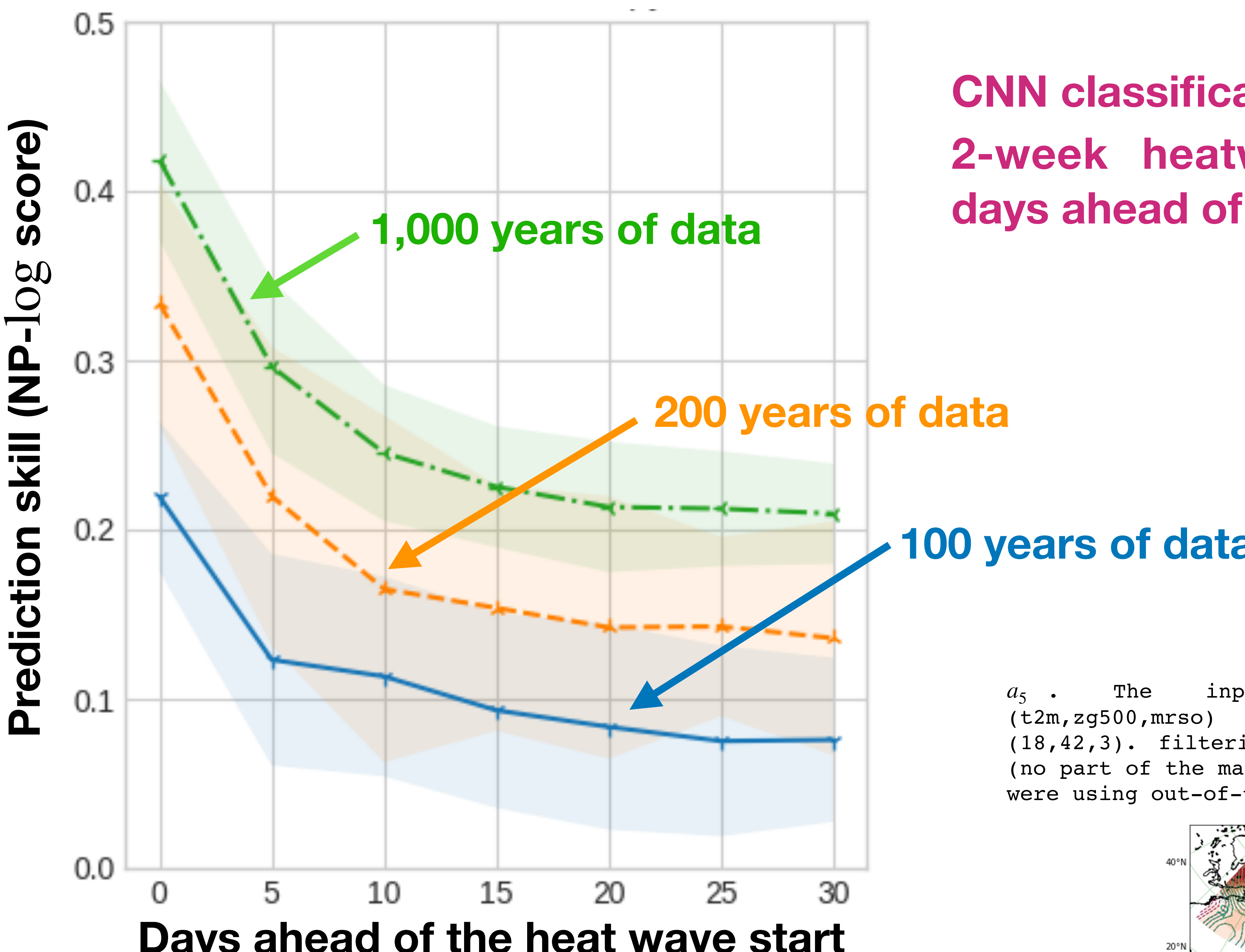
- In order to test the efficiency of the probabilistic prediction of the probabilistic classification, we use the logarithmic score $\mathbb{E}\left\{\log\left[p_{Y_n}\left(X_n\right)\right]\right\}$.

- We define a normalized and positively oriented logarithmic score

$$NP\log = a\mathbb{E}\left\{\log\left[p_{Y_n}\left(X_n\right)\right]\right\} + b,$$

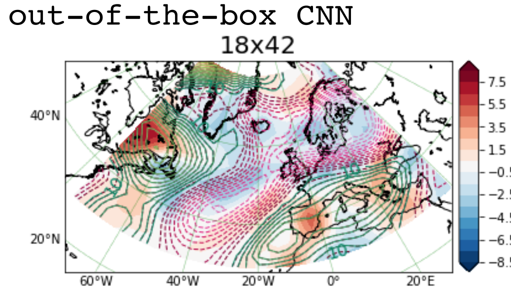
where a and b are such that $NP\log = 0$ for the prediction according to the climatology (prediction using no information on the state X) and $NP\log = 1$ for perfect prediction.

Machine learning for climate applications: a lack of data regime



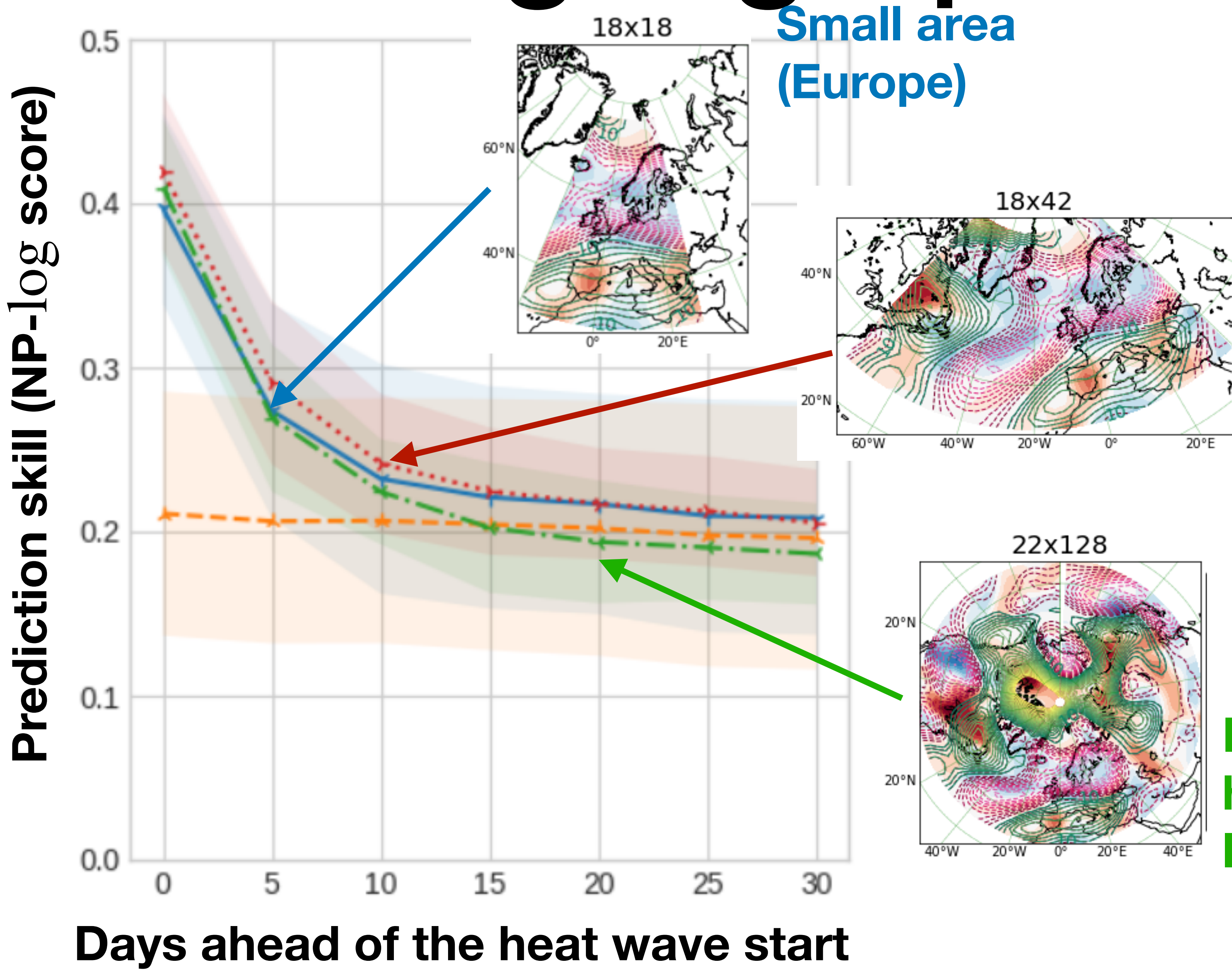
CNN classification of 2-week heatwaves τ days ahead of time

a_5 . The input consists of (t2m,zg500,mrso) with resolution (18,42,3). filtering applied to mrso (no part of the map is set to zero. We were using out-of-the-box CNN



The observation dataset is way too small for good machine learning prediction

Which is the optimal dataset geographical area?



CNN classification of 2-week heatwaves τ days ahead of time

Intermediate size area (Europe+North Atlantic)

Large area (North hemisphere mid-latitudes)

The best performance is obtained for an area of an intermediate size.
This also points to a regime of lack of data for optimal learning.

Conclusions: predicting heat waves with deep neural networks

- Prediction of heat waves is an example of a probabilistic classification problem.
- We use off-the-shelf CNN algorithms, adapted to this situation (probabilistic scores, undersampling, transfer learning).
- Two-week heat waves can be efficiently predicted up to 15 days ahead.
- We are clearly in a regime of lack of data for an optimal prediction.

Abrupt climate changes

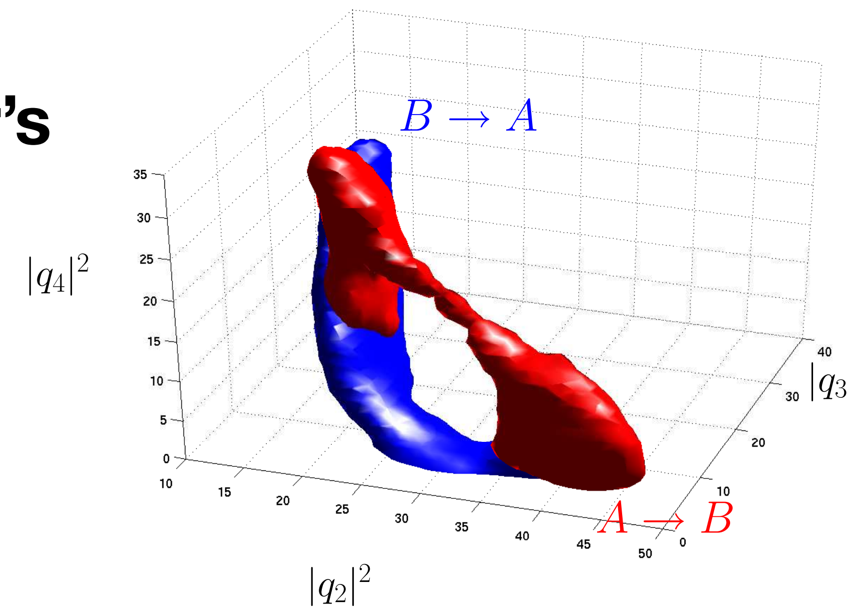
- **II a) Instantons and Arrhenius law for Jupiter's troposphere abrupt transitions.**

F. Bouchet, J. Rolland, and E. Simonnet, PRL, 2019: instantons and Arrhenius law.

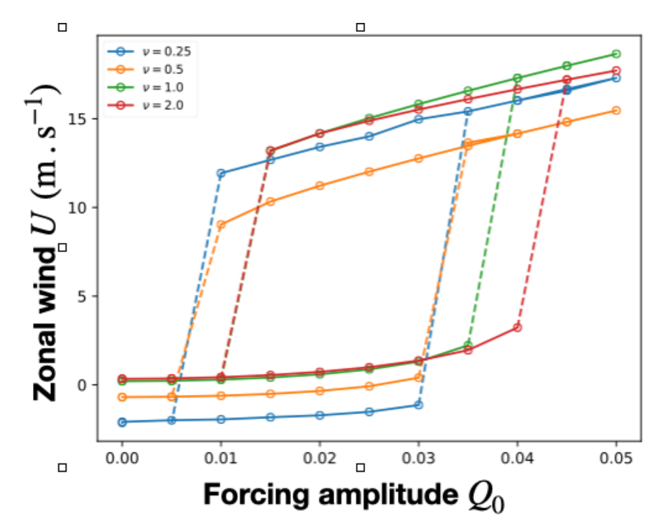
F. Bouchet, J. Rolland, and E. Simonnet, JAS, 2021: multistability with symmetry breaking and instantons.

- **II b) Discontinuous transitions to superrotation on planetary atmospheres.**

C. Herbert, R. Caballero and F. Bouchet, JAS, 2019: study of abrupt transitions, negative feedbacks and the robustness of the bistability range.



Atmosphere jet instantons



Superrotating atmosphere hysteresis

Large deviation theory and other applications (more theoretical works)

- **III a) Path large deviations for kinetic theories.**

F. Bouchet, J. Stat. Phys., 2020: path large deviations for the Boltzmann equation and the irreversibility paradox.

O. Feliachi and F. Bouchet, J. Stat. Phys., 2020: path large deviations for the plasma and the Vlasov equation.

- **III b) Rare events for the Solar System (planet collisions).**

F. Bouchet and E. Woillez, PRL, 2020.

RESEARCH HIGHLIGHTS

Nature Reviews Physics | **The path to the Solar system's destabilization**

16 July 2020

GDR « Theoretical challenges for climate sciences »

- **Identify and work on key theoretical issues that need to be solved for improving the quantitative predictions in climate sciences.**

- **A multidisciplinary consortium: climate sciences, mathematics, physics, computer sciences, statistical physics, data sciences.**

- **Examples :** i) How to reduce the uncertainty about climate sensitivity? ii) How to reduce uncertainty when quantifying probabilities of climate extreme events? iii) How to integrate data and theoretical constraints, using machine learning, to build the next generation of climate models? iv) How to make quantitative the study of future and past climate? v) How to build effective coarse-grained descriptions of climate processes?

Institut des Mathématiques pour la Planète Terre

STRUCTURE ▼ APPELS À PROJET ▼ ANIMATION ▼ RESSOURCES

Conclusions: Studying rare and extreme climate events with rare event algorithms and machine learning

- We can use **rare event algorithms** to gather an amazing statistics for extreme heat waves with PLASIM (PNAS, 2018), and CESM (GRL, 2021) models.

- The dynamical mechanism is the birth of **quasi-stationary non zonal global patterns**, which are much affected by the orography and oceans (PNAS, 2018, GRL 2021).

- **With CNN machine learning, we predict the probability of extreme heat waves** up to 15 days ahead of time (Sub. to Frontiers in Climate, 2021).

- **The coupling of learned committor functions with rare-event algorithms is extremely efficient for toy models** (Sub. to JSTAT, 2021). Work in progress for climate models.

Studying rare events is extremely fascinating!

Opened post-doc positions