Nonequilibrium Markov processes conditioned on large deviations

Hugo Touchette

National Institute for Theoretical Physics (NITheP)
Stellenbosch, South Africa

Workshop on Computational Methods for Statistical Mechanics
International Centre for Mathematical Sciences
Edinburgh, Scotland
2-6 June 2014

- Work with Raphaël Chetrite, Université de Nice, France

Problem

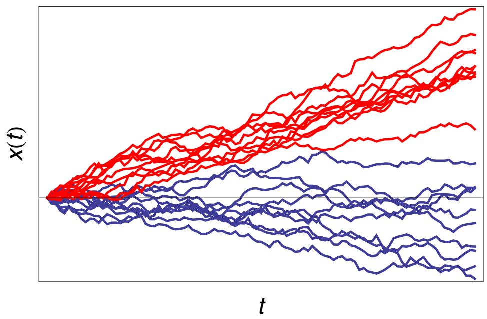
- Markov process: $\{X_t\}_{t=0}^T$
- Observable (rv): A_T
- Conditioned process: $X_t | A_T = a$

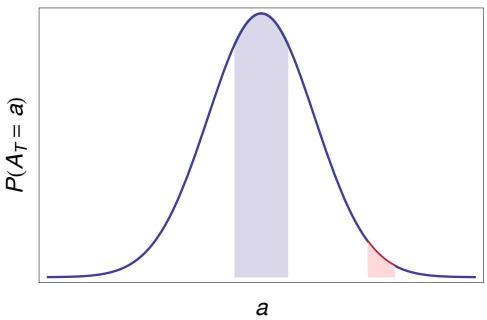
Questions

1. Conditional process Markov?
2. Generator?
3. Relation with X_t?

Connections

- Markov conditioning (Doob)
- Rare event simulations
- Nonequilibrium systems
- Quasi-stationary distributions
- Optimal control (Fleming)





Markov conditioning

- State conditioning (Doob 1957)

$$X_t \,|\, X_T \in \mathcal{A} \qquad \text{target point or set}$$

- Schrödinger bridge (Schrödinger 1931)

$$X_t \,|\, p(x,T) = q(x) \qquad \text{target distribution}$$

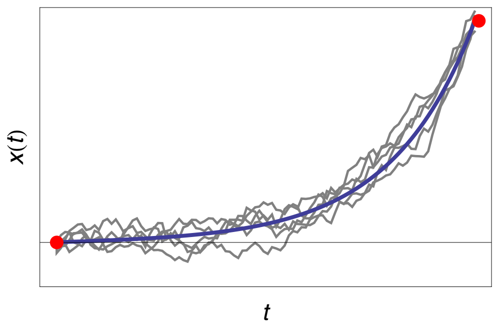
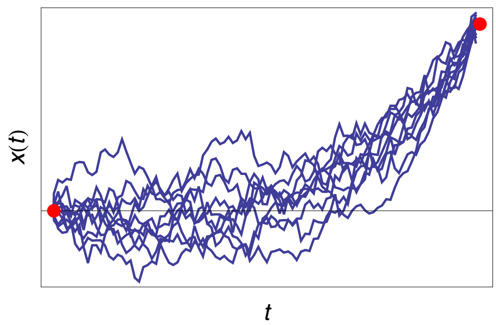
- Quasi-stationary distributions

$$\underbrace{X_t}_{\text{absorbing}} \,|\, \text{not reaching absorbing state} \equiv \underbrace{Y_t}_{\text{non-absorbing}}$$

Here

- $X_t|\mathcal{A}_T$ with $\mathcal{A}_T$ defined on $[0,T]$
- Requires generalization of Doob's transform
- Asymptotic equivalence

Comparison with optimal paths

Low noise limit	Arbitrary noise
	

Low noise limit

- Concentration in path space
- Prob dominated by single path
- Dominant path, instanton

Fluctuation path

Arbitrary noise

- No concentration
- Prob coming from many paths
- No dominating path

Fluctuating dynamics

Process

- Markov process: $X_t \in \mathcal{E}$
- State space: $\mathcal{E}$
- Time interval (horizon): $t \in [0, T]$
- Generator:

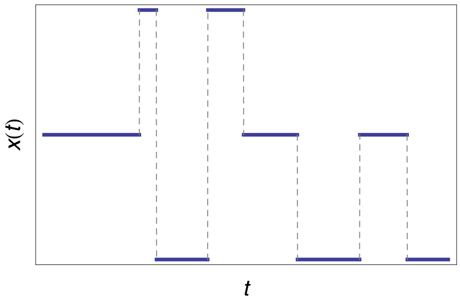
$$\partial_t E_x[f(X_t)] = E_x[Lf(X_t)]$$

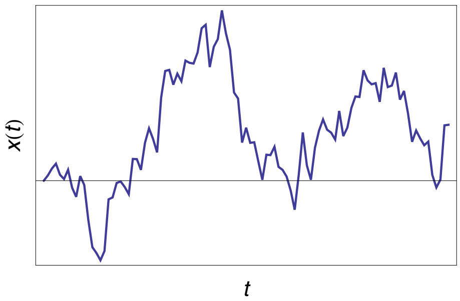
- Master (Fokker-Planck) equation:

$$\partial_t p(x,t) = L^\dagger p(x,t)$$

- Path measure:

$$P[x] = P(\{x_t\}_{t=0}^T)$$





Examples of Markov processes

Pure jump process

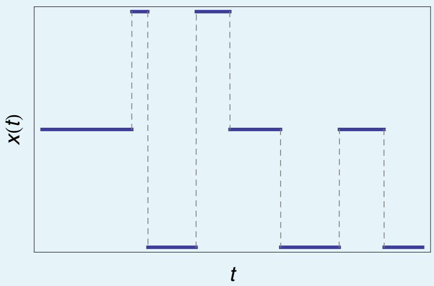
- Transition rates:

$$W(x,y) = P(x \to y \text{ in } dt)/dt$$

- Escape rates:

$$\lambda(x) = \sum_y W(x,y) = (W1)(x)$$

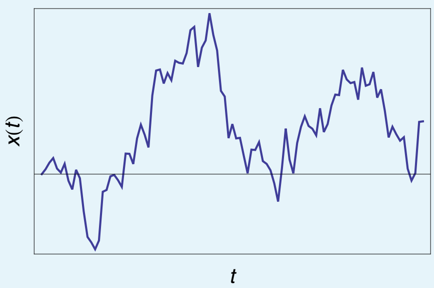
- Generator: $L = \underbrace{W}_{\text{off-diag}} - \underbrace{\lambda}_{\text{diag}}$



Pure diffusion

- SDE: $dX_t = F(X_t)dt + \sigma dW_t$
- Generator:

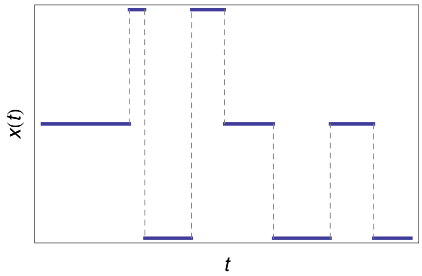
$$L = F \cdot \nabla + \frac{D}{2}\nabla^2, \qquad D = \sigma\sigma^T$$



Conditioning observable

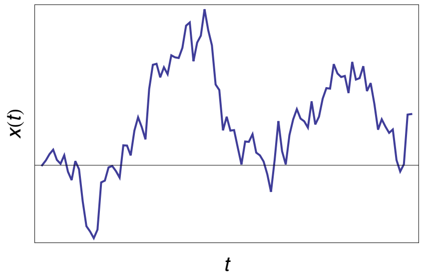
- Observable: $A_T[x]$
- Jump processes:

$$A_T = \frac{1}{T}\int_0^T f(X_t)\,dt + \frac{1}{T}\sum_{\Delta X_t \neq 0} g(X_{t^-}, X_{t^+})$$



- Diffusions:

$$A_T = \frac{1}{T}\int_0^T f(X_t)\,dt + \frac{1}{T}\int_0^T g(X_t)\circ dX_t$$



Examples

- Occupation time $X_t \in \Delta$
- Mean number jumps (activity), current
- Work, heat, entropy production,...

Rare event conditioning

Large deviation principle

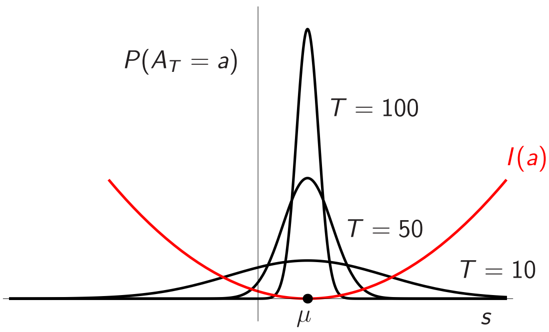
$$P(A_T = a) \asymp e^{-TI(a)}$$

- Meaning of $\asymp$:

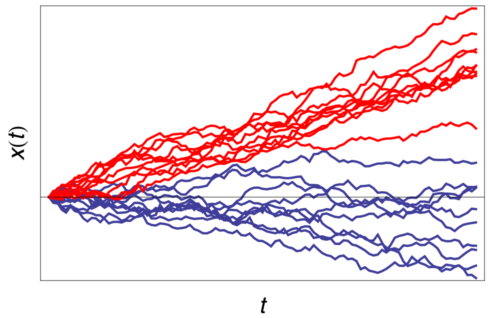
$$\lim_{T\to\infty} -\frac{1}{T}\ln P(A_T = a) = I(a), \qquad P(A_T = a) = e^{-TI(a)+o(T)}$$

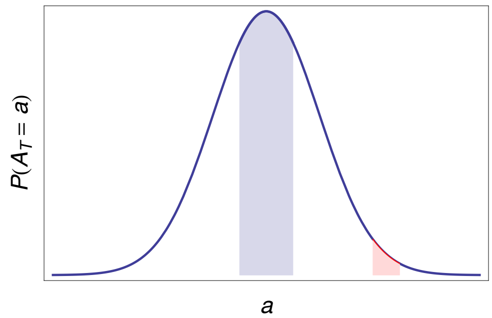
- Rate function: $I(a)$

- Zero of I = Law of Large Numbers
- Concentration point(s): $I(a^*) = 0$
- Small fluctuations = Central Limit Theorem



Conditioned process





- Conditioned process: $X_t | A_T = a$
- Path measure:

$$P^a[x] = P[x|A_T = a] = \frac{P[x, A_T = a]}{P(A_T = a)} = P[x] \frac{\delta(A_T[x] - a)}{P(A_T = a)}$$

- Path microcanonical ensemble

- Not Markov for $T < \infty$
- Becomes equivalent to Markov process as $T \to \infty$
- Non-conditioned process realizing conditioning

Spectral elements

Scaled cumulant function

$$\Lambda_k = \lim_{T\to\infty} \frac{1}{T} \ln E[e^{TkA_T}]$$

- $k \in \mathbb{R}$

Gärtner-Ellis Theorem

Λ_k differentiable, then

1. LDP for A_T
2. $I(a) = \sup_k \{ka - \Lambda_k\}$

Feynman-Kac-Perron-Frobenius

$$\mathcal{L}_k r_k = \Lambda_k r_k$$

- Tilted (twisted) operator: $\mathcal{L}_k$
- Dominant eigenvalue: Λ_k
- Dominant eigenfunction: r_k

Jump processes

$$\mathcal{L}_k = We^{kg} - \lambda + kf$$

Diffusions

$$\mathcal{L}_k = F \cdot (\nabla + kg) + \frac{D}{2}(\nabla + kg)^2 + kf$$

Driven process

Definition

- Process Y_t
- Generator:

$$L_k = r_k^{-1} \mathcal{L}_k r_k - r_k^{-1}(\mathcal{L}_k r_k)$$

- Generalized Doob transform
- Positive, Markov operator: $(L_k 1) = 0$
- Path measure:

$$\frac{P_k^{\text{driven}}[x]}{P[x]} = r_k^{-1}(X_0)\, e^{T(kA_T - \Lambda_k)}\, r_k(X_T)$$

- Radon-Nikodym derivative

Main result

Hypotheses

- A_T satisfies LDP
- Rate function $I(a)$ convex
- Other properties of spectral elements (gap, regular r_k)

Result

Conditioned		Driven	
$X_t \vert A_T = a$	$\overset{T\to\infty}{\cong}$	Y_t	$k(a) = I'(a)$
$P^a[x]$	$\asymp$	$P_{k(a)}^{\text{driven}}[x]$	almost everywhere
$A_T = a$		$A_T \to a$	in probability

- Driven process realizes conditioning $A_T = a$
- Same typical states for other observables

Interpretation of equivalence

Measure equivalence

$$\lim_{T\to\infty} \frac{1}{T} \ln \frac{P^a[x]}{dP_k^{\text{driven}}[x]} = 0, \qquad P^a \text{ or } P_k^{\text{driven}}\text{-a.e.}$$

- $P^a[x] = P_k^{\text{driven}}[x]\, e^{o(T)}$
- Same path measures on log scale

Observable equivalence

- Observable (rv): B_T

Conditioned		Driven
$X_t \| A_T = a$		Y_t
$B_T \to b^*$	$\Leftrightarrow$	$B_T \to b^*$

- Same typical states
- Different fluctuations (LDPs) in general

Idea of the proof

Microcanonical	Canonical	Driven
$X_t \| A_T = a$		Y_t
	$P_k^{\text{cano}}[x] = \dfrac{e^{kTA_T[x]}}{E[e^{kTA_T}]} P[x]$	
$P^a[x] = P[x \| A_T = a]$		$P_k^{\text{driven}}[x]$

Driven $\longrightarrow$ canonical

- $P_k^{\text{driven}}[x] \asymp P_k^{\text{cano}}[x]$
- Same large deviations

Microcanonical $\longrightarrow$ canonical

- $P^a[x] \asymp P_k^{\text{cano}}[x]$ if $I(a)$ convex
- Same typical states
- General result about conditioning vs tilting

Driven process: Explicit form

Jump process

- Original process: $W(x,y)$
- Driven process:

$$W_k(x,y) = r_k^{-1}(x)\, W(x,y)\, e^{kg(x,y)}\, r_k(y), \qquad k = I'(a)$$

- Evans PRL 2004, Jack and Sollich PTPS 2010

Diffusion

- Reference SDE:

$$dX_t = F(X_t)dt + \sigma dW_t$$

- Driven SDE:

$$dY_t = F_k(Y_t)dt + \sigma dW_t$$

- Modified drift:

$$F_k(y) = F(y) + D(kg + \nabla \ln r_k), \qquad k = I'(a)$$

Application: Langevin equation

$$dX_t = -\gamma X_t dt + \sigma dW_t \qquad \longrightarrow \qquad X_t | A_T = a$$

Area under path

$$A_T = \frac{1}{T}\int_0^T X_t dt$$

- $f(x) = x$, $g = 0$
- Rate function: $I(a) = \frac{\gamma^2 a^2}{2\sigma^2}$
- Eigenfunction: $r_k(x) = e^{kx/\gamma}$
- Modified drift:

$$F_{k(a)}(x) = -\gamma x + \frac{a}{\gamma}$$

- $k(a) = I'(a)$

Empirical variance

$$A_T = \frac{1}{T}\int_0^T X_t^2 dt$$

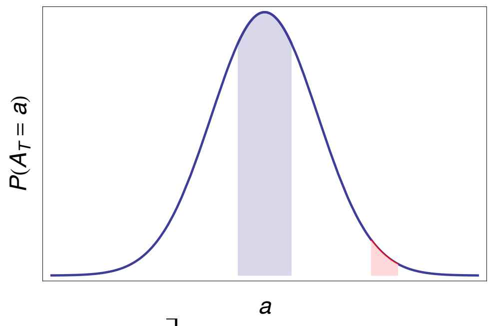
- $f(x) = x^2$, $g = 0$
- Modified drift:

$$F_{k(a)} = -\frac{\sigma^2}{2a}x$$

- Modified friction

Large deviation simulations

- $A_T = a$ exponentially rare
- Direct sampling: sample size $\sim e^T$
- Importance sampling (reweighting)
  - Change process
  - Make $A_T = a$ typical



$$P(A_T = a) = E_X[\delta(A_T - a)] = E_Y\left[\frac{dP_X}{dP_Y}\,\delta(A_T - a)\right]$$

Driven process Y_t

- Makes $A_T = a$ typical
- Good (optimal) change of process
- Problem: Y_t based on r_k, Λ_k and $I(a)$

Learning algorithm [Borkar 2008]

1. Direct sampling + feedback $\rightarrow$ iterative estimation of Λ_k, r_k
2. Control leading to driven process

Conclusions

$$\underbrace{X_t \,|\, A_T = a}_{\text{conditioned}} \overset{T\to\infty}{\cong} \underbrace{Y_t}_{\text{driven}}$$

- Effective Markov dynamics for rare events
- Explicit interpretation of asymptotic equivalence
- Similar to equivalence of equilibrium ensembles
- Generalization of Markov conditioning and bridges
- Links: QSD, stochastic control, conditional limit theorems

Future work

- Large deviation simulations
- Consequences for nonequilibrium systems

References

R. Chetrite, H. Touchette
Nonequilibrium Markov processes conditioned on large deviations
arxiv:1405.5157 (submitted to Ann. IHP)

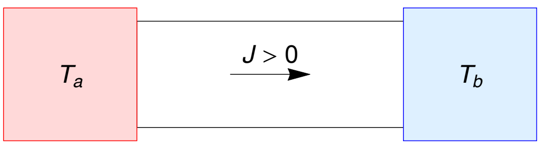
R. Chetrite, H. Touchette
Nonequilibrium microcanonical and canonical ensembles
and their equivalence
Phys. Rev. Lett. **111**, 120601, 2013

H. Touchette
The large deviation approach to statistical mechanics
Phys. Reports **478**, 1-69, 2009

http://www.maths.qmul.ac.uk/~ht

Nonequilibrium systems

Nonequilibrium

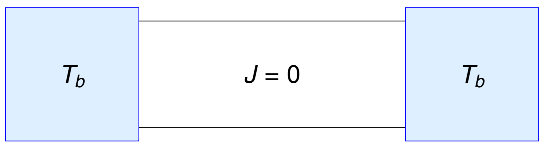


- Microscopic dynamics:

$$W^{\text{noneq}}(x \to y)?$$

- Many models possible

Equilibrium



- Microscopic dynamics known
- Detailed balance:

$$\frac{W^{\text{eq}}(x \to y)}{W^{\text{eq}}(y \to x)} = e^{\beta \Delta E}$$

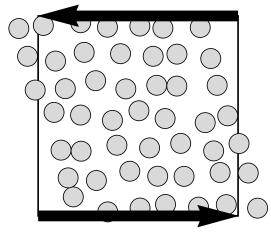
Evans's hypothesis

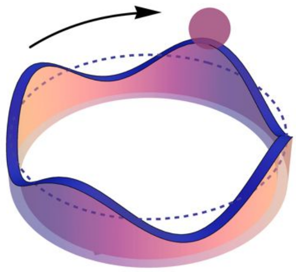
$$W^{\text{noneq}}(x \to y) = W^{\text{eq}}(x \to y | J)$$

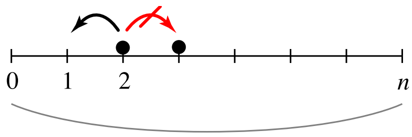
- Nonequilibrium = conditioning of equilibrium
- True? Approximation?

Other applications

- Sheared fluids
  - R. M. L. Evans PRL 2004; JPA 2005
  - Baule & Evans PRL 2008; PRE 2008
- Diffusion on circle
  - Conditioning on current
  - Chetrite & HT PRL 2013
  - Nemoto & Sasa PRE 2011, PRL 2014
- Interacting particles on lattices
  - Conditioning on current
  - TASEP: Schütz *et al.* JSTAT 2010; JSP 2011
  - Zero-range: Harris *et al.* 2013
  - Glauber-Ising: Jack & Sollich PTPS 2010
  - East model: Jack & Sollich JPA 2014
  - Rotators: Knezevic & Evans PRE 2014







Conditioning typically induces long-range interaction

Other connections

Conditional limit theorems

- Sequence of rvs: $X_1, X_2, \ldots, X_n, \quad X_i \sim P(x)$
- Sample mean: $S_n = \frac{1}{n}\sum_{i=1}^{n} f(X_i)$
- Conditional marginal:

$$\lim_{n\to\infty} P(X_i = x | S_n = s) = \frac{e^{kf(x)}}{E[e^{kf(X)}]} P(x)$$

Control representations of PDEs

PDE $\phi(x,t)$ $\partial_t \phi = L\phi$	$\overset{I=-\ln\phi}{\longrightarrow}$	Hamilton-Jacobi equation	(Hopf-Cole)
		$\downarrow$	
		Dynamic programming	
		$\downarrow$	
		Optimal stochastic control $=$	Doob transform

- Fleming, Sheu, Dupuis 1980's, 1990's