Large deviation simulations: Equilibrium vs nonequilibrium systems

Hugo Touchette

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

Numerical Aspects of Nonequilibrium Dynamics
Institut Henri Poincaré, Paris
25 April 2017

Problems

Large deviation problem

$$P(A_n \in B) \approx e^{-nI}$$

Dual problem

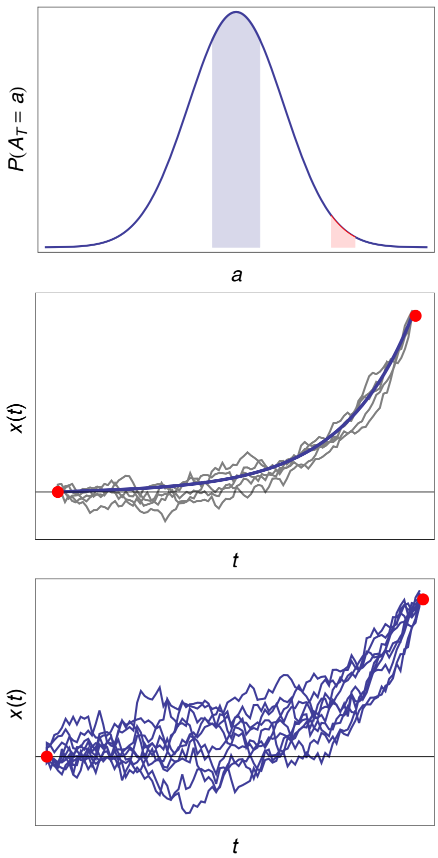
$$E[e^{kA_n}] \approx e^{n\lambda(k)}$$

- Generating function
- Spectral problem (Kac)

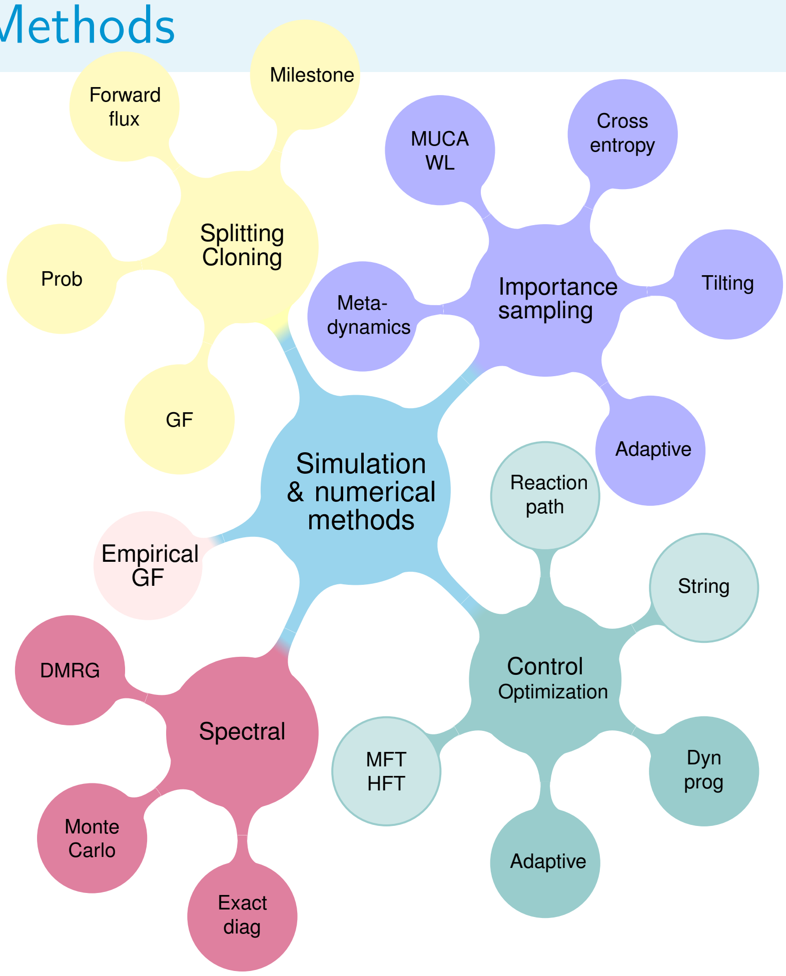
Prediction problem

How is the fluctuation created?

- Reaction or optimal path
- Fluctuation process
- Conditioning



Methods



Questions

- Equilibrium or nonequilibrium method?
- Equilibrium vs nonequilibrium fluctuations
- Nonequilibrium more difficult?

Methods covered

- Spectral
- Optimization
- Importance sampling

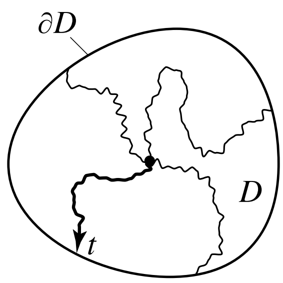
Low-noise large deviations

- Process:

$$dX_t^\varepsilon = F(X_t^\varepsilon)dt + \sqrt{\varepsilon}\, dW_t, \qquad X_0 \in O$$

- Transition rare event:

$$P(X_\tau^\varepsilon \in B | X_0 \in O) \approx e^{-V/\varepsilon}, \qquad \varepsilon \to 0$$



Freidlin-Wentzell-Graham

$$V = \inf_{x\in B} \underbrace{V(x)}_{\text{quasi-potential}} = \inf_{x\in B} \inf_{x_0\in O, x_t = x} \underbrace{\frac{1}{2}\int_0^t (\dot{x}_s - F(x_s))^2 ds}_{\text{FW action, Lagrangian } J[x]}$$

- Deterministic control problem
- Optimal path $\{x_t^*\}$
- $V(x)$ solves Hamilton-Jacobi-Bellman equation (1st order)

Low-noise large deviations (cont'd)

Fluctuation created by optimal path

Conditioning

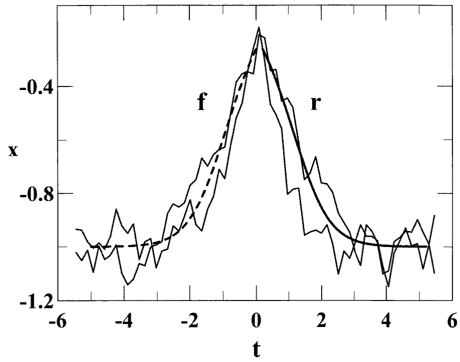
- $P[x] \approx e^{-J[x]/\varepsilon}$ max
- $P[x|\text{escape event}]$ concentrates on optimal path

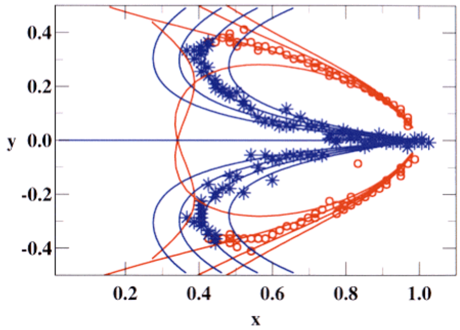
Equilibrium

- Optimal path = relaxation path^R
- $V(x) = J[x^R_{\text{relax}}]$

Nonequilibrium

- Optimal path $\neq$ relaxation path^R
- Current loops
- Transversal decomposition [Graham 80's]





[Luchinsky et al 1997]

Applications

SDEs

$$dX_t = F(X_t)dt + \sqrt{\varepsilon}\,\text{noise}$$

- Escape, transition paths, escape time
  [Onsager-Machlup 1953] [Freidlin-Wentzell 70s] [Graham 80-90s] ...
- Experiments [Luchinsky and McClintock 90s]
- Metastability [Olivieri-Vares 2005]

SPDEs

$$d\rho(x,t) = F[\rho]dt + \sqrt{\varepsilon}\,\text{noise}$$

- Heat equation [Faris, Jona-Lasinio 1982]
- Ginzburg equation [Graham 1990s]
- Reaction-diffusion [Vanden-Eijnden 2000s]
- 2D fluid equations [Laurie-Bouchet 2014]
- MFT/HFT [Bertini et al 2000s]

Long-time large deviations

- Process:

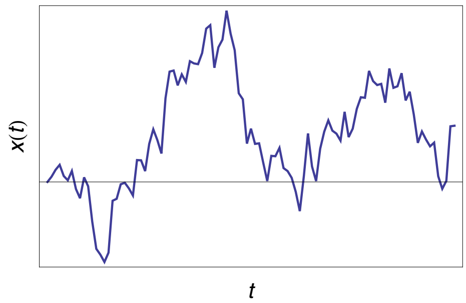
$$dX_t = F(X_t)dt + \sigma dW_t$$

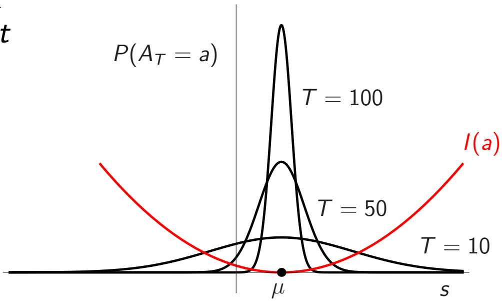
- Observable:

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

- Large deviation principle (LDP):

$$P(A_T = a) \approx e^{-TI(a)}, \quad T\to\infty$$





Examples

- Occupation time, empirical density
- Current, mean speed, activity
- Work, heat, entropy production (stochastic thermo)

Dual problem

Scaled cumulant function

$$\lambda(k) = \lim_{T\to\infty}\frac{1}{T}\ln E[e^{TkA_T}]$$

Gärtner-Ellis Theorem

$\lambda(k)$ 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: $\lambda(k)$
- Dominant eigenfunction: r_k

Jump processes

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

Diffusions

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

Spectral problem

Equilibrium

- X_t reversible, g gradient, f arbitrary
- $\mathcal{L}_k$ non-Hermitian but conjugated to Hermitian:

$$\mathcal{H}_k = \rho^{1/2}\mathcal{L}_k\rho^{-1/2}, \qquad \psi_k^2 = r_k l_k$$

- Real spectrum (quantum problem)

Nonequilibrium

- X_t nonreversible OR g non-gradient, f arbitrary
- $\mathcal{L}_k$ non-Hermitian, not conjugated to Hermitian
- Complex spectrum
- Full spectral problem:

$$\begin{aligned} \mathcal{L}_k r_k &= \lambda(k) r_k \\ \mathcal{L}_k^\dagger l_k &= \lambda(k) l_k \end{aligned} \qquad r_k(x) l_k(x) \stackrel{|x|\to\infty}{\longrightarrow} 0$$

Algorithms

Markov chains

- Non-symmetric positive matrices
- Direct diagonalization
- Power method
- DMRG [Gorissen-Vanderzande 2011]

Diffusions

- Equilibrium: (Quantum) spectral methods
- Nonequilibrium: Fourier decomposition, basis functions

Cloning/splitting

- Particle/trajectory simulation
- Yields SCGF
- No eigenvector

Other approach: Optimization and control

$$A_T = \tilde{A}(\rho_T, J_T) = \int f(x)\underbrace{\rho_T(x)}_{\text{dist.}}dx + \int g(x)\underbrace{J_T(x)}_{\text{current}}dx$$

Drift optimization

$$I(a) = \inf_{u:E_u[A_T]=a} \frac{1}{2\sigma^2}\int (u(x)-F(x))^2 \rho^u_{\text{inv}}(x)\,dx$$

- Requires stationary distribution ρ^u_{inv}

Stochastic optimal control

$$I(a) = \lim_{T\to\infty} \inf_{u:A^u_T\to a} \frac{1}{2\sigma^2 T}\int_0^T (u(X^u_t)-F(X^u_t))^2 dt$$

- Controlled process: X^u_t
- Constrained control (dual for SCGF is unconstrained)
- Solves Hamilton-Jacobi-Bellman equation (2nd order)

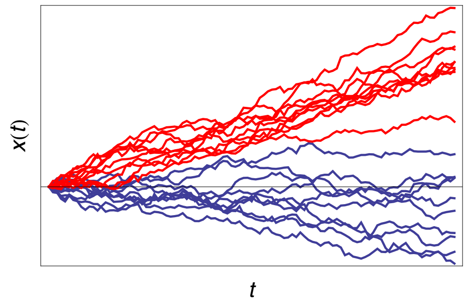
Fluctuation process

Optimal control process

$$d\hat{X}_t = F_k(\hat{X}_t)dt + \sigma dW_t$$

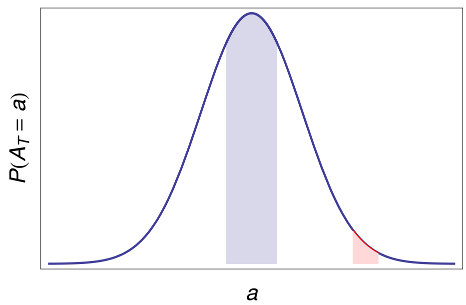
- Optimal drift:

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



Conditioning

$$\underbrace{X_t \mid A_T = a}_{\substack{\text{conditioned}\\ \text{microcanonical}}} \overset{T\to\infty}{\cong} \underbrace{\hat{X}_t}_{\substack{\text{driven}\\ \text{canonical}}}$$



- Effective process creating the fluctuation
- Process generalization of optimal path

[Chetrite HT, PRL 2013, AHP 2015, JSTAT 2015]

Equilibrium vs nonequilibrium

$$A_T = \tilde{A}(\rho_T, J_T) = \int f(x) \underbrace{\rho_T(x)}_{\text{dist.}} dx + \int g(x) \underbrace{J_T(x)}_{\text{current}} dx$$

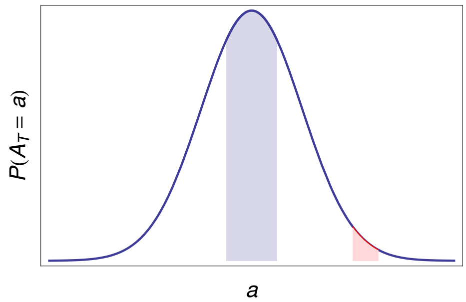
X_t	g	$\hat{X}_t$
Reversible	0	Reversible Same spectrum Rayleigh-Ritz variational principle
	Gradient	Reversible Rayleigh-Ritz variational principle
	Non-gradient	Non-reversible
Non-reversible	0	Non-reversible Donsker-Varadhan principle
	Other	Non-reversible

- Process closest to X_t that creates fluctuation
- Distance: $\|u - F\|_{\rho^u_{\text{inv}}}$ or $\frac{1}{T} S(P_{\hat{X}} \| P_X)$

Simulations: Importance sampling

- $P(A_T = a) \approx e^{-TI(a)}$
- Direct sampling:

$$\text{sample size } = L \sim e^T$$



Importance sampling

- Change process: $X_t \to \hat{X}_t$
- Make $A_T = a$ typical
- Change of measure:

$$P(A_T = a) = E_X[\mathbb{1}_a(A_T)] = E_{\hat{X}}\left[\frac{dP_X}{dP_{\hat{X}}} \mathbb{1}_a(A_T)\right]$$

- Estimator:

$$\hat{P}_L(a) = \frac{1}{L} \sum_{j=1}^{L} \mathbb{1}_a(A_T^{(j)}) R$$

Efficiency

Zero-variance process

- Conditioned process: $X_t | A_T = a$
- Estimator variance:

$$\mathrm{var}_{\hat{X}}(\hat{P}_L) = \frac{E_{\hat{X}}[R^2 \mathbb{1}_a(A_T)] - p^2}{L} = 0$$

Effective process

- Not zero variance
- Asymptotic optimality:

$$\lim_{T\to\infty} -\frac{1}{T} \ln E_{\hat{X}}[R^2 \mathbb{1}_a(A_T)] = 2I(a)$$

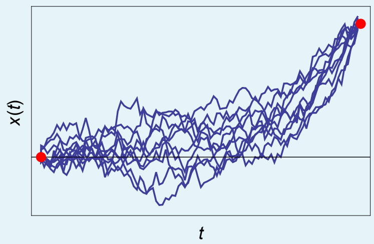
- Variance goes to 0 with largest rate
- Exponential tilting:

$$P_{\text{driven}}[x] \approx P_k[x] = \frac{e^{TkA_T[x]} P[x]}{E[e^{TkA_T}]}$$

Connections

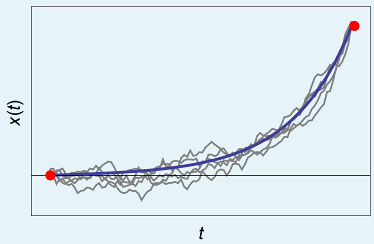
$\sigma > 0$

- Stochastic optimal control
- Quadratic cost function (log RND)
- Nonlinear HJB equation (2nd order)
- Similar to cross-entropy minimization



$\sigma \to 0$

- Deterministic optimal control
- Quadratic cost function (log RND)
- Nonlinear HJB (1st order, viscosity)



Optimal controller		Value function		
Dyn programming	$\longleftrightarrow$	HJB equations	$\longleftrightarrow$	Exponential tilting
$\hat{X}_t$		LD functions		

Conclusions

LD ≡ control/optimization ≡ spectral
Type of fluctuations determines class of control designs

Equilibrium fluctuations

Reversible control
Hermitian spectral problem

Nonequilibrium fluctuations

Non-reversible control
Non-hermitian spectral problem

Future work

- Adapt methods from quantum mechanics
- Methods for non-hermitian operators
- HJB-based methods
- Adaptive control methods
- Error bars
- Benchmarking

References

H. Touchette
Introduction to large deviations: Theory, applications, simulations
2011 Oldenburg School Lecture Notes, arxiv:1106.4146

R. Chetrite, H. Touchette
Variational and optimal control representations
of conditioned and driven processes
JSTAT P12001, 2015

J. Bucklew
Introduction to Rare Event Simulation
Springer, 2004

More pointers at
www.physics.sun.ac.za/~htouchette/ldt