Capture Slow Convergence of Markov Processes Numerically

Yao Li[1] Hui Xu[2]

[1]University of Massachusetts Amherst [2]Amherst College

Outline

- Markov processes on general (uncountable) state spaces may have slow (sub-exponential or polynomial) convergence to steady-states. Examples are seen in many models in statistical physics.
- Rigorous proof towards sub-exponential convergence rate is difficult in general.
- We introduce a **hybrid method** to capture the slow-convergence phenomenon for Markov processes. The problem is reduced to the numerical computation of first-passage times.
- Two numerical examples for the microscopic heat conduction are exhibited.

General setting and main challenge

Setting

- Markov chain Ψ_n on a measurable state space $(X, \mathcal{B})$. Ψ_n has a transition kernel P. $P(x, A) = \mathbb{P}[\Psi_{n+1} \in A \,|\, \Psi_n = x]$.
- Invariant probability measure π:
$$\pi(A) = \int_X \pi(\mathrm{d}x) P(x, A) \quad \text{for any } A \in \mathcal{B}$$
- **Interested in : *Speed of convergence* in total variation norm**
$$\|P^n(x, \cdot) - \pi(\cdot)\|_{TV} := \phi(n)\,.$$
Exponential convergence: $\phi(n) \sim r^n$, $0 < r < 1$; Polynomial convergence: $\phi(n) \sim n^{-\beta}$, $\beta > 0$.
- *Decay of correlation* (or mixing)
$$C_{\eta,\xi}(n) = |\int_X (P^n\eta)(x)\xi(x)\pi(\mathrm{d}x) - \int_X \eta(x)\pi(\mathrm{d}x) \int_X \xi(x)\pi(\mathrm{d}x)|$$
is closely related to the speed of convergence, where η, ξ are integrable functions on X.

Difficulty

- The total variation norm on uncountable state spaces is difficult to compute numerically.
- Correlation decay is more computable through Monte Carlo simulation.
- Example: If $C_{\eta,\xi}(n) \sim n^{-2}$, the tolerance of relative error is ϵ, then sample size for $C_{\eta,\xi}(N)$ is $\sim N^4\epsilon^{-2}$. Usually large N is necessary to capture the asymptotic trend of $C_{\eta,\xi}(n)$. When $N = 10^2$ and $\epsilon = 0.01$. Computational cost $\sim 10^{14}$.
- However, large N is necessary to precisely capture the asymptotic trend of the convergence.

Known analytical results

Notation and definition

Let Ψ_n be a Markov chain with transition kernel $P(x, \cdot)$.

- Ψ_n is irreducible with respect to a reference measure λ on $\mathcal{B}$ if for any $x \in X$, $A \in \mathcal{B}$ with $\lambda(A) > 0$, there exists $n > 0$ such that $P^n(x, A) > 0$.
- A set C is called a *uniform reference set* if there exist $\delta > 0$ and probability measure $\nu(\cdot)$ such that
$$P(x, \cdot) \geq \delta\nu(\cdot) \text{ for all } x \in C\,.$$
- For $A \in \mathcal{B}$, $\tau_A = \inf_{n>0}\{\Psi_n \in A\}$ is the first passage time to A.

Results

Suppose Ψ_n is an aperiodic, λ-irreducible Markov chain that admits a uniform reference set C.

- If $\sup_{x\in C}[\tau_C] < \infty$, then there exists an invariant probability measure π. (Theorem 10.0.1 in [2])
- For any $\beta > 0$, if $\mathbb{E}_{x_0}[\tau_C^\beta] < \infty$ for some $x_0 \in X$, then $\mathbb{E}_x[\tau_C^\beta] < \infty$ for λ-almost every $x \in X$.
- If there exists $\beta > 0$ such that
$$\sup_{x\in C} \mathbb{E}_x[\tau_C^\beta] < \infty,$$
then for any probability measures μ, ν on X satisfying
$$\mathbb{E}_\mu[\tau_C^\beta] < \infty \quad \text{and} \quad \mathbb{E}_\nu[\tau_C^\beta] < \infty\,,$$
we have
$$\lim_{n\to\infty} n^\beta \|\mu\mathcal{P}^n - \nu\mathcal{P}^n\|_{TV} = 0\,,$$
where $\|\cdot\|_{TV}$ is the total variation norm. (Theorem 2.7 in [1])

Verify conditions for polynomial convergence

Easier to verify analytically

Some conditions can be verified analytically.

- Ψ_n is an aperiodic Markov chain.
- Ψ_n is irreducible with respect to a reference measure λ.
- Ψ_n admits a uniform reference set C.

Easier to verify numerically

Verifying some conditions rigorously may be extremely difficult .....

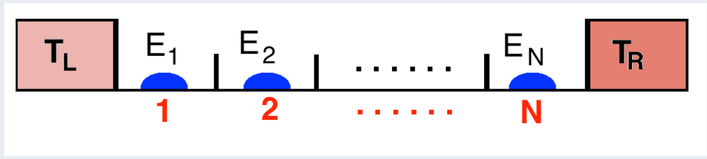
- For a given probability measure μ and the given uniform reference set, there exists a β such that
$$\mathbb{E}_\mu[\tau_C^\beta] < \infty\,.$$
Method: Simulate τ_C. Numerically compute the slope of $\mathbb{P}_\mu[\tau_C > n]$ on a log-log plot. $\mathbb{P}_\mu[\tau_C > n] \sim n^{-\beta}$ implies $\mathbb{E}_\mu[\tau_C^{\beta-\epsilon}] < \infty$ for any $\epsilon > 0$. Here μ could be δ_{x_0} for some x_0, any given probability measure, and the numerically obtained invariant measure π.
- The function
$$\gamma(x) := \sup_{n\geq 1} \frac{\mathbb{P}_x[\tau_C > n]}{n^{-\beta}}$$
is uniformly bounded for all x in the given set C and $\beta > 0$.
Method: Search a grid of lattice points in C, find monotonicity of $\gamma(x)$, or use some gradient-free numerical optimization algorithms.

Numerical examples: Heat conduction in microscopic

- We apply our method to two Markov processes in statistical mechanics. Both of them model the 1-D microscopic heat conduction in nonequilibrium setting.
- Optimal convergence rates of these models are difficult to obtain analytically.
- Both Markov processes are reduced from deterministic dynamical systems models, in which particles have kinetic motion.
- These models aim to study the microscopic derivation of macroscopic thermodynamic laws. This is a century-old challenge to mathematicians and physicists.

Model I: Stochastic Energy Exchange Model

Model Description



- A chain of N lattice sites $\{1, \cdots, N\}$ is connected to two heat baths with temperatures T_L and T_R, named as site 0 and site $N + 1$, respectively.
- Site i carries a finite amount of energy, denoted by E_i.
- An exponential clock is associated with each pair of adjacent sites, with rate $R = \sqrt{\min\{E_i, E_{i+1}\}}$.
- When the i-th clock rings, sites i and $i + 1$ exchange energy
$$(E_i', E_{i+1}') = (p(E_i + E_{i+1}), (1 - p)(E_i + E_{i+1})) \quad p \sim U(0, 1)\,. \quad E_0 \sim Exp(T_L), \quad E_{N+1} \sim Exp(T_R)$$

Verification of conditions

- The time-1 sample chain Φ_n generated by this model is irreducible with respect to the Lebesgue measure on $\mathbb{R}_+^N$.
- Set $C = \{(E_1, \cdots, E_N) \,|\, 0.1 \leq E_i \leq 100, i = 1 \sim N\}$ is a uniform reference set of Φ_n.
- The tail of the first passage time to C, or $\mathbb{P}[\tau_C > t]$, is $\sim t^{-2}$ when starting from the numerically generated invariant measure. See Figure 1.
- For $1 < \beta < 2$, $\gamma(E)$ increases with the decrease of each E_i. See Figure 2.
- At $E^* = (0.1, \cdots, 0.1)$, which is the numerical maximal of $\gamma(E)$ for $E \in C$ and $1 \leq \beta \leq 2$, $\mathbb{P}[\tau_C > t] \sim t^{-2}$. See Figure 3.

Conclusion

- The stochastic energy exchange model admits a unique invariant measure π, which is the nonequilibrium steady-state (NESS).
- π is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}_+^N$.
- For (Lebesgue measure) almost every $x \in \mathbb{R}_+^N$, $\|P^t(x, \cdot) - \pi(\cdot)\|_{TV} \sim t^{-2}$.

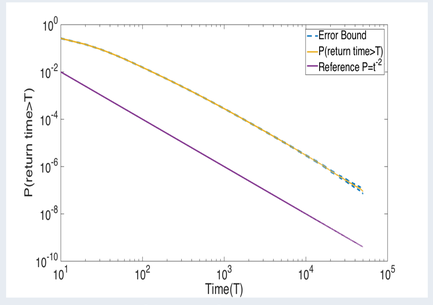


Figure: 1, $\mathbb{P}_\pi[\tau_C > t]$, π is the numerical steady-state obtained from long time averaging.

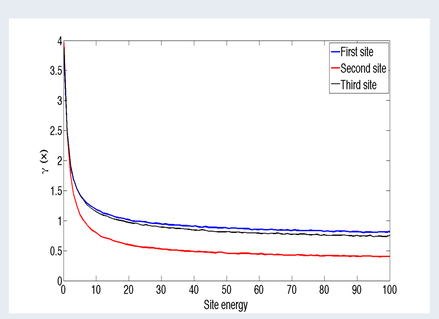


Figure: 2, Change of $\gamma(E)$ for varying E when only one site energy changes.

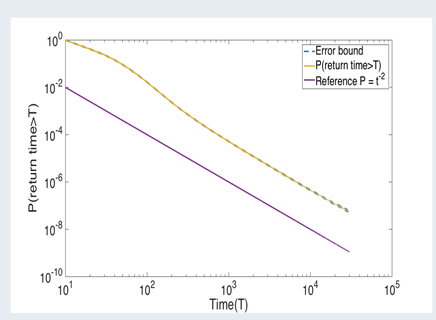
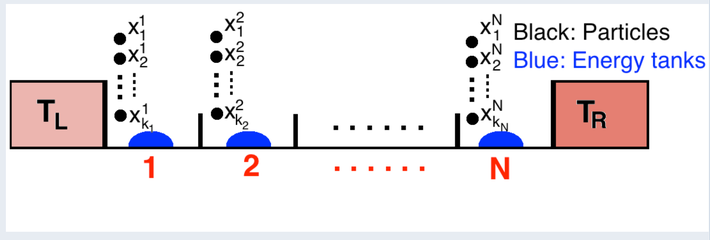


Figure: 3, $\mathbb{P}_{E^*}[\tau_C > t]$ for $E* = (0.1, 0.1, 0.1)$.

Model II: Random Halves Model

Model Description



- A chain of N lattice sites $\{1, \cdots, N\}$ is connected to two heat baths with temperatures T_L and T_R. One "energy tank" is placed at each site.
- Particles are injected from heat baths with rates ρ_L and ρ_R. The initial distribution of each particle's energy depends on bath temperatures.
- An exponential clock with rate $\sqrt{\text{particle energy}}$ is associated with each particle.
- When one clock rings, $\mathbb{P}[$ jump to neighbors $] = p_m$, $\mathbb{P}[$ exchange energy with tank $] = 1 - p_m$. A particle exits the system if it jumps to the bath.
- Rule of energy exchange: $x \sim$ particle energy, $s \sim$ tank energy.
$$(x', s') = ((1 - u^2)x + s, u^2x) \quad u \sim U(0, 1)\,.$$
- State space: $\mathbf{\Omega}$. A state of the model: $\omega = \{(k_1, \cdots, k_N), (\{x_1^1, \cdots, x_{k_1}^1\}, s_1), \cdots, (\{x_1^N, \cdots, x_{k_N}^N\}, s_N)\}$. $k_i \sim$ number of particles at site i.

Verification of conditions

- The time-1 sample chain Φ_n generated by this model is irreducible with respect to a reference measure λ on $\mathbf{\Omega}$. λ resembles the Lebesgue measure.
- Set $C = \left\{\omega \in \mathbf{\Omega} \,|\, k_i \leq 40, 0.1 \leq x_{k_i}^j \leq 100, s_j \leq 100, i = 0 \sim 40, j = 1 \sim N\right\}$ is a uniform reference set. (Particle number ≤ 40, tank energy ≤ 100, $0.1 \leq$ particle energy ≤ 100).
- The tail of the first passage time to C, or $\mathbb{P}[\tau_C > t]$, is $\sim t^{-2}$ when starting from the numerically generated invariant measure. See Figure 4.
- For $1 < \beta < 2$, $\gamma(\omega), \omega \in C$ increases with (i) decreasing site energy at each site, (ii) decreasing particle energy at each site, and (iii) increasing number of particles in each site. See Figures 5, 6, and 7.
- At $\omega^* = \{(40, \cdots, 40), (\{0.1, \cdots, 0.1\}, 0), \cdots, (\{0.1, \cdots, 0.1\}, 0)\}$ (40 particles with energy 0.1 at each site), the numerical maximal of $\gamma(\omega), \omega \in C$, $\mathbb{P}[\tau_C > t] \sim t^{-2}$. See Figure 8.

Conclusion

- The stochastic energy exchange model admits a unique invariant measure π, which is the nonequilibrium steady-state (NESS).
- π is absolutely continuous with respect to the reference measure λ.
- For λ-almost every $\omega \in \mathbf{\Omega}$, $\|P^t(\omega, \cdot) - \pi(\cdot)\|_{TV} \sim t^{-2}$.

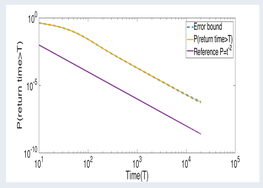


Figure: 4, $\mathbb{P}_\pi[\tau_C > t]$, π is the numerical steady-state obtained from long time averaging.

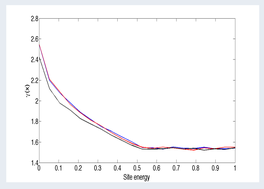


Figure: 5, Change of $\gamma(\omega)$ Change of $\gamma(\omega)$ for varying ω when site energy at only one site changes.

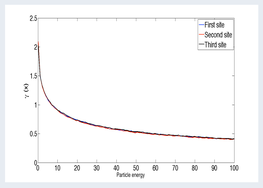


Figure: 6, Change of $\gamma(\omega)$ Change of $\gamma(\omega)$ for varying ω when particle energy at only one site changes.

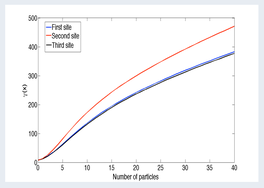


Figure: 7, Change of $\gamma(\omega)$ Change of $\gamma(\omega)$ for varying ω when number of particles at only one site changes.

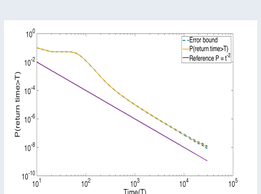


Figure: 8, $\mathbb{P}_{\omega^*}[\tau_C > t]$, ω^* is the numerically obtained maximal of $\gamma(\omega)$ when $\omega \in C$.

Reference

[1] Nummelin, E. and Tuominen, P., *The rate of convergence in Orey's theorem for Harris recurrent Markov chains with applications to renewal theory*. Stochastic Processes and Their Applications, 15(3), 1983.

[2] S Sean P Meyn and Richard L Tweedie, *Markov chains and stochastic stability*, Cambridge University Press, 2009.

[3] Li, Y. and Xu, H., *Numerical simulation of polynomial-speed convergence phenomenon* , Under Review (arXiv:1703.04008).