行政院國家科學委員會專題研究計畫 成果報告

有限馬可夫鏈的 log-Sobolev 常數
研究成果報告(精簡版)

計 畫 類 別 ：個別型
計 畫 編 號 ：NSC 99-2115-M-009-008-
執 行 期 間 ：99 年 08 月 01 日至 100 年 07 月 31 日
執 行 單 位 ：國立交通大學應用數學系（所）

計 畫 主 持 人 ：陳冠宇

處 理 方 式 ：本計畫可公開查詢

中 華 民 國 100 年 10 月 07 日

成果報告中英文摘要：

（一） 中文摘要

在馬可夫鏈蒙地卡羅的理論中，一個特別選定的馬可夫鏈被用來代替其穩定分佈進行取樣。而取樣的方法就是不斷的模擬該馬可夫鏈直到其機率分佈已經足夠接近穩定分佈。至今，蒙地卡羅方法已經被廣泛地應用在其他科學領域，其中包含了統計物理學、資訊科學以及生物學。運用蒙地卡羅方法時，最直接的問題就是何時該停止馬可夫鏈的模擬。數學上，這個問題就是要找尋一個正確的停止時間使的停止時的機率分佈和穩定分佈的差異很小。

在 2010-2011 期間，我們考慮統計物理學門裡的的一個古典模型—厄任菲司特甕。Paul Ehrenfest 當年提出該模型是為了解釋熱力學第二定律—任何封閉系統的熵在時間趨近無限大時會趨近零。意即該模型會趨於穩定。而我們研究的對象是：當粒子數趨近無限大時，該模型的穩定時間為何。同時，我們也藉由這個例子來探討馬可夫鏈收斂時的相變現象。

關鍵詞：馬可夫鏈，厄任菲司特甕。

（二）英文摘要

In Markov chain Monte Carlo (MCMC, for short) theory, a particular Markov chain is run for a very long time, say T, until its distribution is close enough to the stationarity. In practice, one is interested in the time T to stop the simulation and choose a random sample, and T is closely related to the mixing time. The most widely used in measuring the convergence rate of the MCMC method includes the total variation, separation, entropy, and the L^p-norm. In recent years, for models of statistical mechanics and of theoretical computer science and many others, there has been a flourishing of new mathematical ideas to rigorously control the time T.

We consider families of Ehrenfest chains and provide a simple criterion on the L^p-cutoff and the L^p -precutoff with specified initial states for $1\leq p<\infty$. For the family with an L^p -cutoff, a cutoff time is described and a possible window is given. For the family without an L^p -precutoff, the exact order of the L^p -mixing time is determined. The result is consistent with the well-known conjecture on cutoffs of Markov chains proposed by Peres in 2004, which says that a cutoff exists if and only if the multiplication of the spectral gap and the mixing time tends to infinity.

Keywords: Markov chains, Ehrenfest chains.

THE REPORT OF GRANT NSC99-2115-M-009-008

The research on the mixing of Markov chains during 2010-2011 has fruitful results and they are collected in [4]. In this article, we consider families of Ehrenfest chains and provide a simple criterion on the L^p-cutoff and the L^p-precutoff with specified initial states for $1 \le p < \infty$. For the family with an L^p-cutoff, a cutoff time is described and a possible window is given. For the family without an L^p-precutoff, the exact order of the L^p-mixing time is determined. The result is consistent with the well-known conjecture on cutoffs of Markov chains proposed by Peres in 2004, which says that a cutoff exists if and only if the multiplication of the spectral gap and the mixing time tends to infinity.

Consider a time-homogeneous Markov chain on a finite set Ω with one-step transition matrix K. Let $K^t(x,\cdot)$ denote the probability distribution of the chain at time t started at x. It is well-known that if K is ergodic (irreducible and aperiodic), then

$$\lim_{t\to\infty} K^t(x,y) = \pi(y) \quad \forall x, y \in \Omega,$$

where π is the stationary distribution of K on Ω. Denote by k_x^t the relative density of $K^t(x,\cdot)$ with respect to π, that is, $k_x^t(y) = K^t(x,y)/\pi(y)$. For $1 \le p < \infty$, define the L^p-distance by

$$D_p(x,t) = \|k_x^t - 1\|_{L^p(\pi)} = \left(\sum_{y\in\Omega} |k_x^t(y) - 1|^p \pi(y) \right)^{1/p}.$$

For $p = \infty$, the L^∞-distance is set to be $D_\infty(x,t) = \max_y |k_x^t(y) - 1|$. In the case $p = 1$, this is exactly twice of the total variation distance between $K^t(x,\cdot)$ and π, which is defined by

$$D_{\mathrm{TV}}(x,t) = \|K^t(x,\cdot) - \pi\|_{\mathrm{TV}} = \max_{A\subset\Omega} \{K^t(x,A) - \pi(A)\}.$$

In the case $p = 2$, it is the so-called chi-square distance. For any $\epsilon > 0$ and $1 \le p \le \infty$, define the L^p-mixing time by

$$T_p(x,\epsilon) = \min\{t \ge 0 : D_p(x,t) \le \epsilon\}.$$

Consider a family of finite ergodic Markov chains (Ω_n, K_n, π_n) with specified initial states x_n. For $1 < p \le \infty$, the family is said to present an L^p-cutoff with cutoff time t_n if

$$\lim_{n\to\infty} D_{n,p}(x_n, (1+a)t_n) = \begin{cases} 0 & \text{if } a > 0 \\ \infty & \text{if } -1 < a < 0 \end{cases},$$

where $D_{n,p}$ denotes the L^p-distance for the nth Markov chain. In total variation, separation and L^1-distance, the cutoff is the same expect the replacement of the limit ∞ with 1 in the first two and 2 in the last. The concept of cutoffs was introduced by Aldous and Diaconis in [1, 2, 3] to capture the fact that many ergodic Markov chains converge abruptly to their stationary distributions. We refer the reader to [5, 6, 7, 8, 9] for details and further discussions on variant examples.

In [4], we treat the Ehrenfest chains, a classical example introduced by Paul Ehrenfest to remark the second law of thermodynamics. In detail, let $\Omega_n = \{0, 1, ..., n\}$ and K_n be the Markov kernel of the Ehrenfest chain on Ω_n given by

$$K_n(i, i+1) = 1 - \frac{i}{n}, \quad K_n(i+1, i) = \frac{i+1}{n}, \quad \forall 0 \le i \le n-1. \tag{0.1}$$

Key words and phrases. Cutoff phenomenon, Ehrenfest chains.

Clearly, the unbiased binomial distribution, $\pi_n(i) = \binom{n}{i}2^{-n}$, is the stationary distribution of K_n and the pair (K_n, π_n) is reversible, i.e. $\pi_n(i)K_n(i,j) = \pi_n(j)K_n(j,i)$ for all $i, j \in \Omega_n$. By lifting the chain to a random walk on the hypercube, one may use the group representation of $(\mathbb{Z}_2)^n$ to identify the eigenvalues and eigenvectors of K_n. The following is our main result achieved in [4, Theorems 3.1-4.1].

Theorem 0.1. *Let K_n be defined in* (0.1) *and set $K'_n = (I + nK_n)/(n+1)$, $\pi_n(i) = \binom{n}{i}2^{-n}$. For $p \in [1, \infty)$, the following are equivalent.*

1. *The family $\{(\Omega_n, K'_n, \pi_n) : n = 1, 2, ...\}$ with starting states $(x_n)_{n=1}^{\infty}$ has an L^p-cutoff.*
2. *The family $\{(\Omega_n, K'_n, \pi_n) : n = 1, 2, ...\}$ with starting states $(x_n)_{n=1}^{\infty}$ has an L^p-precutoff.*
3. *$|n - 2x_n|/\sqrt{n} \to \infty$ as $n \to \infty$.*

Moreover, if (3) *holds, then, as $n \to \infty$,*

$$T_{n,p}(x_n, \epsilon) = \frac{n}{2} \log \frac{|n - 2x_n|}{\sqrt{n}} + O(n), \quad \forall \epsilon > 0,\ p \in (1, \infty),$$

where $O(n)$ denotes a function of order less than or equal to n. For $p = 1$, the above identity remains true with $\epsilon \in (0, 2)$.

For the total variation, [6, Theorem 6.5] provides a sufficient condition on cutoffs, while [4, Theorem 3.1] proves that such a condition is necessary. For the L^p-cutoff with $1 < p < \infty$, Theorem [6, Theorem 6.5] gives the L^2 case, while [4, Theorem 4.1] gives the L^p case. It is worthwhile to remark that if there is a cutoff, then the main term of the mixing time is the same for all $1 \le p < \infty$. If the sequence in Theorem 0.1(3) is bounded, then there is no precutoff and the order of mixing time is of order n, the size of Ω_n. We would like to point out that in Appendix A of [4], we derive the limiting distribution of the hitting probability for the simple random walk on $\mathbb{Z}$. This is not only of interests by itself but also plays an important role in proving the total variation cutoff.

References

[1] David Aldous. Random walks on finite groups and rapidly mixing Markov chains. In *Seminar on probability, XVII*, volume 986 of *Lecture Notes in Math.*, pages 243–297. Springer, Berlin, 1983.

[2] David Aldous and Persi Diaconis. Shuffling cards and stopping times. *Amer. Math. Monthly*, 93(5):333–348, 1986.

[3] David Aldous and Persi Diaconis. Strong uniform times and finite random walks. *Adv. in Appl. Math.*, 8(1):69–97, 1987.

[4] Guan-Yu Chen, Yang-Jen Fang, and Yuan-Chung Sheu. The cutoff phenomenon for ehrenfest chains. Submitted, 2011.

[5] Guan-Yu Chen and Laurent Saloff-Coste. The cutoff phenomenon for ergodic markov processes. *Electron. J. Probab.*, 13:26–78, 2008.

[6] Guan-Yu Chen and Laurent Saloff-Coste. The L^2-cutoff for reversible Markov processes. *J. Funct. Anal.*, 258(7):2246–2315, 2010.

[7] David A. Levin, Yuval Peres, and Elizabeth L. Wilmer. *Markov chains and mixing times*. American Mathematical Society, Providence, RI, 2009. With a chapter by James G. Propp and David B. Wilson.

[8] Laurent Saloff-Coste. Lectures on finite Markov chains. In *Lectures on probability theory and statistics (Saint-Flour, 1996)*, volume 1665 of *Lecture Notes in Math.*, pages 301–413. Springer, Berlin, 1997.

[9] Laurent Saloff-Coste. Random walks on finite groups. In *Probability on discrete structures*, volume 110 of *Encyclopaedia Math. Sci.*, pages 263–346. Springer, Berlin, 2004.

國科會補助計畫衍生研發成果推廣資料表

日期:2011/10/07

國科會補助計畫	計畫名稱: 有限馬可夫鏈的log-Sobolev常數
	計畫主持人: 陳冠宇
	計畫編號: 99-2115-M-009-008- 學門領域: 機率

無研發成果推廣資料

99 年度專題研究計畫研究成果彙整表

計畫主持人：陳冠宇	計畫編號：99-2115-M-009-008-
計畫名稱：有限馬可夫鏈的 log-Sobolev 常數	

成果項目			量化			單位	備註（質化說明：如數個計畫共同成果、成果列為該期刊之封面故事...等）
			實際已達成數（被接受或已發表）	預期總達成數(含實際已達成數)	本計畫實際貢獻百分比		
國內	論文著作	期刊論文	0	0	100%	篇	
		研究報告/技術報告	0	0	100%		
		研討會論文	0	0	100%		
		專書	0	0	100%		
	專利	申請中件數	0	0	100%	件	
		已獲得件數	0	0	100%		
	技術移轉	件數	0	0	100%	件	
		權利金	0	0	100%	千元	
	參與計畫人力（本國籍）	碩士生	1	1	100%	人次	
		博士生	0	0	100%		
		博士後研究員	0	0	100%		
		專任助理	0	0	100%		
國外	論文著作	期刊論文	0	1	100%	篇	The draft has been submitted to journal for publication.
		研究報告/技術報告	0	0	100%		
		研討會論文	0	0	100%		
		專書	0	0	100%	章/本	
	專利	申請中件數	0	0	100%	件	
		已獲得件數	0	0	100%		
	技術移轉	件數	0	0	100%	件	
		權利金	0	0	100%	千元	
	參與計畫人力（外國籍）	碩士生	0	0	100%	人次	
		博士生	0	0	100%		
		博士後研究員	0	0	100%		
		專任助理	0	0	100%		

其他成果 (無法以量化表達之成果如辦理學術活動、獲得獎項、重要國際合作、研究成果國際影響力及其他協助產業技術發展之具體效益事項等，請以文字敘述填列。)	無

	成果項目	量化	名稱或內容性質簡述
科教處計畫加填項目	測驗工具(含質性與量性)	0	
	課程/模組	0	
	電腦及網路系統或工具	0	
	教材	0	
	舉辦之活動/競賽	0	
	研討會/工作坊	0	
	電子報、網站	0	
	計畫成果推廣之參與（閱聽）人數	0	

國科會補助專題研究計畫成果報告自評表

請就研究內容與原計畫相符程度、達成預期目標情況、研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）、是否適合在學術期刊發表或申請專利、主要發現或其他有關價值等，作一綜合評估。

1. 請就研究內容與原計畫相符程度、達成預期目標情況作一綜合評估

■達成目標

□未達成目標（請說明，以 100 字為限）

□實驗失敗

□因故實驗中斷

□其他原因

說明：

2. 研究成果在學術期刊發表或申請專利等情形：

論文：□已發表 ■未發表之文稿 □撰寫中 □無

專利：□已獲得 □申請中 ■無

技轉：□已技轉 □洽談中 ■無

其他：(以 100 字為限)

3. 請依學術成就、技術創新、社會影響等方面，評估研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）(以 500 字為限)

我們考慮統計物理學門裡的的一個古典模型—厄任菲司特甕。Paul Ehrenfest 當年提出該模型是為了解釋熱力學第二定律—任何封閉系統的熵在時間趨近無限大時會趨近零。意即該模型會趨於穩定。而我們研究的對象是：當粒子數趨近無限大時，該模型的穩定時間為何。在統計物裡上，我們嚴格地證明了穩定時間。在數學上，我們藉由這個例子來探討馬可夫鏈收斂時的相變現象。我們所運用的技巧,在理論的建立與推廣上具有啟發性的作用。