國 立 交 通 大 學

應用數學系

碩 士 論 文

厄任菲斯特甕的相變現象

The total variation cutoff for Ehrenfest chains

研　究　生：范揚仁

指導教授：陳冠宇　教授

中 華 民 國 一 〇 〇 年 六 月

厄任菲斯特鏈的相變現象
The total variation cutoff for Ehrenfest chains

研 究 生：范揚仁　　Student：Yang-Jen Fan

指導教授：陳冠宇　　Advisor：Guan-Yu Chen

國 立 交 通 大 學
應 用 數 學 系

碩 士 論 文

A Thesis
Submitted to Department of Applied Mathematics
National Chiao Tung University
In partial Fulfillment of the Requirements
for the Degree of
Master
in

Applied Mathematics

June 2011

Hsinchu, Taiwan, Republic of China

中華民國一〇〇年六月

厄任菲斯特甕的相變現象

學生：范揚仁　　　　　　　　　　　　　　　　　指導教授：陳冠宇

國立交通大學應用數學系（研究所）碩士班

摘　　　要

本論文主要為研究厄任菲斯特甕的相變現象，並證明了此現象存在性的等價條件。資料來源起於指導教授陳冠宇於 2010 年發表在 Journal of Functional Analysis 的文章 The L^2-cutoff for reversible Markov processes，而此文章已證明厄任菲斯特甕的相變存在性之充分條件。利用相變現象的存在性定義，我們先是證明了其不存在的充分條件。借此我們去證明原命題的逆否命題為真。

The total variation cutoff for Ehrenfest chains

Student：Yang-Jen Fan Advisor：Dr. Guan-Yu Chen

Department（Institute）of Applied Mathematics

National Chiao Tung University

ABSTRACT

The main work of this thesis was giving a equivalent condition of the total variation cutoff for Ehrenfest chains. The source of the question came from my advisor Dr. Chen who had punished the paper, “The L^2-cutoff for reversible Markov processes”, in Journal of Functional Analysis on 2010. And the paper had proven the sufficient condition of the total variation cutoff. We derived the sufficient condition for no cutoffs. Therefore, we proved the contrapositive true.

誌　　　　謝

這篇碩士論文的誕生，最感謝的是我的指導教授陳冠宇老師。從第一次與老師討論研究內容以來，老師一步步教導我裡頭的觀念，讓我有時間釐清並且消化，僅作為一位研究的初學者，著實安心不少，並在我處理問題遇到困難時，適時地給予協助，鼓勵我持續前進，在老師身上，我看到了一位學者對於學問的專業與熱誠，而那是否即為做研究的態度與精神呢？在人生的規劃上，老師也提供我看法，於面對抉擇時，他告訴我：「抱持最大的期望、最壞的打算，之後全力以赴！」。兩年間從老師身上學到的難以三言兩語道盡，真心感謝老師在這段時間的教導。

感謝口試委員黃啟瑞教授、許順吉教授以及許元春教授在論文口試上的寶貴意見。黃啟瑞老師於口試時，所提出了許多問題，能夠把論文的結果再做推廣，非常感謝老師。許順吉老師所分享的讀書經驗，讓我緊張的心情舒緩不少，非常謝謝老師。許元春老師是我本來跟的指導教授，感謝老師在我初升碩一的階段勉勵我挑戰難度高的科目，其對爾後的研究相當有幫助，非常感謝老師的用心與包容。

感謝一起選修中研院實分析班的吳國禎、施建宏、譚識浩，大家建立了革命情感，讓我在中研院感到相當溫暖，也讓我覺得其實我並不孤單。還要感謝在交大同一個研究室裡的同學，一起寫論文寫到三更半夜的段俊旭學長、笑點超級低的蔡詩妤同學、時常自製點心來請大家的周彥伶同學、對數學教育有所抱負的楊嘉勝同學、認為碩是兩年不夠要念三年的洪湧昇同學，不論是學業上或生活上，都給我相當大的幫助，與你們一起相處非常愉快，讓我兩年的碩士生涯能夠多采多姿。

感謝父親范國泰先生，母親張碧珍女士的養育之恩，從小到大，都給予我支持，不論學業或生活上有甚麼困難，家永遠是我的避風港，讓我從新整裝出發面對挑戰。還有感謝弟弟揚文，雖然我們彼此少有時間在一起聊天，但時常提醒自己不能在你面前漏氣。

最後，有大家一路上的陪伴與幫助，我才能順利地碩士畢業，今後我會繼續勇敢向前邁進。由衷地感謝大家，謝謝你們。

范揚仁 2011/6/22　於交大

Contents

The total variation cutoff for Ehrenfest chains

Yang-Jen Fan

Abstract

We consider the family of Ehrenfest chains and provide the equivalent condition for a total variation cutoff with specified initial states. If there is a cutoff, we also give a cutoff time.

1 Introduction

A Markov process $(X_t)_{t\in T}$ is a stochastic process with the Markov property that, given the value of X_t, X_s for $s > t$ are conditionally independent of X_s for $s < t$. T denotes the time and mostly equals to $\{0, 1, \ldots\}$ or $[0, \infty)$. The Markov property says

$$P(X_{t_{n+1}} = j|X_{t_0} = i_0, \ldots, X_{t_{n-1}} = i_{n-1}, X_{t_n} = i) = P(X_{t_{n+1}} = j|X_{t_n} = i),\ \forall\ n \geq 0,$$

where $i_0, i_1, \ldots,\ i,\ j$ belong to a state space $\mathcal{S}$ and $t_0 < t_1 < t_2 < \cdots < t_{n+1}$. When $T = \{0, 1, 2, \ldots\}$, the process $(X_t)_{t\in T}$ is said to be a discrete time Markov chain with state space $\mathcal{S}$. The chain is called time-homogenous if $P(X_{n+1} = j|X_n = i) = P(X_1 = j|X_0 = i)$, for all $n \geq 0$. The matrix $K = (K(i, j))_{i,j\in\mathcal{S}}$ with $K(i, j) = P(X_1 = j|X_0 = i)$ is called the one-step transition matrix or Markov kernel.

If $T = [0, \infty)$, the transition probability $H_t(i, j) = P(X_t = j|X_0 = i)$ forms a semigroup i.e. $H_{t+s} = H_t H_s$ and $H_0 = I$. Suppose Q is an infinitesimal generator of H_t and $q = \sup_{i\in\mathcal{S}} \{-Q(i, i)\} < \infty$. Then,

$$H_t = e^{tQ} := \sum_{n=0}^{\infty} \frac{(tQ)^n}{n!}.$$

A realization of the semigroup $(H_t)_{t\geq 0}$ is to consider a discrete time Markov chain $(X_n)_{n=0}^{\infty}$ with transition matrix $K = \frac{Q+qI}{q}$ and a Poisson process $(N(t))_{t\geq 0}$ which has intensity q and is independent of $(X_n)_{n=0}^{\infty}$. Then the process $Y_t = X_{N(t)}$ is a continuous time Markov chain satisfying $P(Y_t = j|Y_0 = i) = H_t(i, j)$. A simple application of Bayes' formula yields

$$H_t(i, j) = \sum_{n=0}^{\infty} P_i(X_{N(t)} = j|N(t) = n)P(N(t) = n) = \sum_{n=0}^{\infty} e^{-qt}\frac{(qt)^n}{n!}K^n(i, j),\ \forall\ t \geq 0, i, j \in \mathcal{S}.$$

For simplicity, when we state (X_n) as a continuous time Markov chain associated with a transition matrix K, we mean $P(X_t = j|X_0 = i) = e^{-t(I-K)}(i, j)$. That is, the semigroup of the transition probability has $K - I$ as the infinitesimal generator.

A probability π is called a stationary distribution for a Markov chain with transition matrix K if $\pi K = \pi$. The following two theorems display the convergence of ergodic Markov chains to their stationary distribution. This should be related to the Perron-Frobenius theorem.

Theorem 1.1. *Let K be an irreducible and aperiodic transition matrix on $\mathcal{S}$. Suppose that K has a stationary distribution π. Then, for all $x, y \in \mathcal{S}$,*

$$\lim_{n\to\infty} K^n(x, y) = \pi(y).$$

Theorem 1.2. *Let K be an irreducible transition matrix on $\mathcal{S}$ with stationary distribution π. Then, for all $x, y \in \mathcal{S}$,*

$$\lim_{t\to\infty} H_t(x, y) = \pi(y),$$

In this paper, we consider Ehrenfest chain on $\mathcal{S}_n = \{0, 1, \dots n\}$ with transition matrix K_n given by

$$K_n(k, k+1) = 1 - \frac{k}{n},\ K_n(k+1, k) = \frac{k+1}{n},\ \forall\ 0 \le k < n. \tag{1.1}$$

It is clear that K_n is irreducible and has the unbiased binomial distribution π_n as the stationary distribution. That is $\pi_n(k) = \begin{pmatrix} n \\ k \end{pmatrix} 2^{-n}$. In discrete time case, since K_n is of period 2, we consider K_n' instead, which is defined by

$$K_n' = \frac{1}{n+1} I_{n+1} + \frac{n}{n+1} K_n, \tag{1.2}$$

where I_n is the $(n+1) \times (n+1)$ identity matrix. In the above setting K_n' is irreducible and aperiodic with stationary distribution π_n.

As a consequence of Theorem 1.1-1.2, $K_n'^t(x, y)$ and $e^{-t(I-K_n)}(x, y)$ converge to $\pi_n(y)$ as $t \to \infty$. A natural question arises: How fast the convergence? The first thing is to set up a measurement on distributions. For example, the total variation distance given by

$$\|K_n^t(k, \cdot) - \pi_n(\cdot)\|_{TV} := \max_{A \subseteq \mathcal{S}_n} \{K_n^t(k, A) - \pi_n(A)\}.$$

The $L^2(\pi)$-distance is defined by

$$\|\frac{K_n^t(k, \cdot)}{\pi_n(\cdot)} - 1\|_2 := \left(\sum_{i \in \mathcal{S}_n} |\frac{K_n^t(k, i)}{\pi_n(i)} - 1|^2 \pi_n(i) \right)^{\frac{1}{2}}.$$

The convergence to stationarity has a cutoff in total variation if there is some sequence (t_n) such that

$$\lim_{n\to\infty} \|K_n^{ct_n}(k,\cdot) - \pi_n(\cdot)\|_{TV} = \begin{cases} 1 & \textit{if} \quad c \in (0,1) \\ 0 & \textit{if} \quad c > 1 \end{cases}.$$

Similarly, in the $L^2(\pi)$-distance, the cutoff means the existence of some sequence (t_n) such that

$$\lim_{n\to\infty} \|\frac{K_n^{ct_n}(k,\cdot)}{\pi_n(\cdot)} - 1\|_2 = \begin{cases} \infty & \textit{if} \quad c \in (0,1) \\ 0 & \textit{if} \quad c > 1 \end{cases}.$$

(t_n) is closely related to the mixing time which is defined by

$$T_n(x_n, \varepsilon) := \inf\{t \geq 0 | D_n(x_n, t) \leq \varepsilon\},$$

where D_n is any distance we defined above. We refer the reader to Section 2.1 or the Chen[1] for more details.

We quote Chen and Saloff-Coste[2] for illustration. In the discrete time case, let $t_n = (n/2)\log(|2x_n - n|/\sqrt{n})$, if

$$\frac{|2x_n - n|}{\sqrt{n}} \longrightarrow \infty, \tag{1.3}$$

then there are constant $\beta > 0$ and N such that for all $n \geq N$,

$$e^{-c} \leq \|\frac{K_n'^{\,t_n + cn}(x_n,\cdot)}{\pi_n(\cdot)} - 1\|_2 \leq \beta e^{-2c}.$$

The first inequality holds for $c < 0$ and the second inequality is true for $c > 0$. If (1.3) holds, then there is a $L^2(\pi)$-cutoff. In fact, (1.3) is also necessary for the $L^2(\pi)$-cutoff. The main goal of this thesis is to show the following theorem.

Theorem 1.3. *Fix $(x_n)_{n\geq 1}$ for all $x_n \in \mathcal{S}_n$. Let $\mathcal{F} = \{(\mathcal{S}_n, K_n', \pi_n) | n = 1, 2, \ldots\}$ be the family of discrete time Ehrenfest chains and $\mathcal{F}_c = \{(\mathcal{S}_n, H_{n,t}, \pi_n) | n = 1, 2, \ldots\}$ be the continuous time case. The following three things are equivalent:*

(i) $\mathcal{F}$(resp, $\mathcal{F}_c$) has a total variation precutoff.

(ii) $\mathcal{F}$(resp, $\mathcal{F}_c$) has a total variation cutoff.

(iii) $|n - 2x_n|/\sqrt{n} \to \infty$.

1.1 Description of the paper

In section 2, we have two parts. At the first part, we introduce the distance between two probabilities on finite set and the mixing time. At the second part, we introduce the notion of the total variation cutoff and precutoff. Theorem 2.1 provides an equivalent condition on precutoff. In section 3, we introduce the technique of spectral analysis and the coupling method to bound the total variation. The main results including theorem 1.3 are discussed in section 4.

2 Terminology

This section will give some definitions and propositions which concern the cutoff.

2.1 Distances and mixing times

In the introduction, we mention the total variation distance and the $L^2(\pi)$-distance which is related to the chi-square distance. Actually, those come from more general definitions.

Definition 2.1. *Let μ,ν be probabilities on a finite set Ω.*

1. *The total variation distance between μ and ν is defined by*

$$\|\mu - \nu\|_{TV} := \sup_{A\subseteq\Omega} |\mu(A) - \nu(A)| = \sup_{A\subseteq\Omega} \{\mu(A) - \nu(A)\}.$$

   In the following, assume that $\nu > 0$ and set $h = \frac{\mu}{\nu}$.

2. *For $p \in [1, \infty)$, the $L^p(\nu)$-distance between μ and ν is defined by*

$$\|h - 1\|_p := \left(\sum_{x\in\Omega} |h(x) - 1|^p \nu(x)\right)^{\frac{1}{p}}.$$

3. *The $L^\infty(\nu)$-distance is given by*

$$\|h(x) - 1\|_\infty := \sup_{x\in\Omega} |h(x) - 1|.$$

Proposition 2.1. *[3, Lemma 2.4.1] Let ν and μ be probabilities on a finite set Ω. Assume that $\nu > 0$ and set $h = \frac{\mu}{\nu}$.*

1. *Set $\nu_* = \inf_{x\in\Omega} \nu(x)$. For $1 \le r \le s \le \infty$,*

$$\|h - 1\|_r \le \|h - 1\|_s \le \nu_*^{\frac{1}{s} - \frac{1}{r}} \|h - 1\|_r.$$

2. $\|\mu - \nu\|_{TV} = \frac{1}{2}\|h-1\|_1$.

Let (X_n) be an irreducible Markov chain on a finite state space $\mathcal{S}$ with initial distribution μ and stationary distribution π. Let K be the transition matrix and $\|\cdot\|_p$ be the measurements introduced in Definition 2.1. The $L^p(\pi)$-distance of the chain (X_n) is defined to be a real valued function $D_p(\mu, n) := \|\mu K^n - \pi\|_p$. If μ is a dirac mass function δ_x, i.e. $\delta_x(x) = 1$ and $\delta_x(y) = 0$ if $y \neq x$, we write briefly $D_p(x, n)$ for $D_p(\delta_x, n)$. The max-$L^p(\pi)$-distance is defined by $D_p(n) := \sup_{\mu} D_p(\mu, n)$. Clearly, $D_p(n) := \sup_{x} D_p(x, n)$.

Definition 2.2. *Let $D_p(\mu, n)$ and $D_p(n)$ be the L^p-distance.*

1. *The $L^p(\pi)$-mixing time is defined to be*

$$T_p(\mu, \varepsilon) := \inf\{n \geq 0 | D_p(\mu, n) \leq \varepsilon\}, \ \forall \varepsilon > 0,$$

   and $\inf \emptyset := \infty$.

2. *The max-$L^p(\pi)$-mixing time is defined by*

$$T_p(\varepsilon) := \inf\{n \geq 0 | D_p(n) \leq \varepsilon\} = \sup_{\mu} T_p(\mu, \varepsilon) = \sup_{x} T_p(x, \varepsilon), \ \forall \varepsilon > 0.$$

The case for continuous-time chains goes in the same way except the domain is replaced by $[0, \infty)$ and we use D_p^c and T_p^c to represent the $L^p(\pi)$-distance and $L^p(\pi)$-mixing time respectively.

Proposition 2.2. *Suppose $D_p(\mu, \cdot)$, $D_p(\cdot)$, $T_p(\mu, \cdot)$, $T_p(\cdot)$ are those in Definition 2.2.*

1. *$D_p(\mu, \cdot)$, $D_p(\cdot)$ are non-increasing.*

2. *$T_p(\mu, \cdot)$, $T_p(\cdot)$ are non-increasing.*

3. *If $T_p(\mu, \varepsilon) < \infty$, then $D_p(\mu, T_p(\mu, \varepsilon)) \leq \varepsilon$. If $T_p(\varepsilon) < \infty$, then $D_p(T_p(\varepsilon)) \leq \varepsilon$.*

Remark 2.1. *The above proposition also applies for the continuous time case. In Proposition 2.2, if $0 < T_p^c(\mu, \varepsilon) < \infty$, then $D_p^c(\mu, T_p^c(\mu, \varepsilon)) = \varepsilon$. Similarly, if $0 < T_p^c(\varepsilon) < \infty$, $D_p^c(T_p^c(\varepsilon)) = \varepsilon$.*

Remark 2.2. *The definition of the distance and mixing time in total variation is the same as in Definition 2.2. Proposition 2.2 and Remark 2.1 also apply for the total variation.*

2.2 Cutoffs

Let $\mathcal{F} = \{(\mathcal{S}_m, K_m, \pi_m); m = 1, 2, \ldots\}$ be a family of ergodic Markov chains.

Definition 2.3. *We called $\mathcal{F}$ has a total variation cutoff if for all $\varepsilon, \delta \in (0,1)$,*

$$\lim_{m\to\infty} \frac{T_{TV}(x_m,\varepsilon)}{T_{TV}(x_m,\delta)} = 1.$$

Remark 2.3. *The continuous time case is defined by the similar way.*

The following proposition provides an equivalent description on the total variation cutoff. When discussing the discrete time family $\mathcal{F}$, we assume

$$\lim_{m\to\infty} T_{TV}(x_m,\delta_0) = \infty \;\; \text{for some} \;\; 0<\delta_0<1. \tag{2.1}$$

Proposition 2.3. *Assume (2.1) holds. $\mathcal{F}$ has a total variation cutoff if and only if there is a sequence of positive integers (a_m) s.t.*

$$\lim_{m\to\infty} D_{TV}(x_m, \lfloor ca_m \rfloor) = 1 \;\; \text{if} \;\; c\in(0,1),$$

and

$$\lim_{m\to\infty} D_{TV}(x_m, \lceil ca_m \rceil) = 0 \;\; \text{if} \;\; c>1.$$

In particular, $(a_m)_{m=1}^{\infty}$ can be chosen to $(T_{TV}(x_m,\delta))_{m=1}^{\infty}$ for any $0<\delta<1$.

Proof. Suppose there is a total variation cutoff. Fix $\varepsilon \in (0,1)$ and $\delta \le \delta_0$. Given $\eta \in (0,\frac{1}{2})$, there is a $M>0$ such that,

$$\left|\frac{T_{TV}(x_m,\varepsilon)}{T_{TV}(x_m,\delta)} - 1\right| < \eta \; \forall \; m \ge M.$$

Expand the above inequality, we obtain

$$T_{TV}(x_m,\delta)(1-\eta) < T_{TV}(x_m,\varepsilon) < T_{TV}(x_m,\delta)(1+\eta). \tag{2.2}$$

By the first inequality of (2.2), we have

$$D_{TV}(x_m, \lfloor T_{TV}(x_m,\delta)(1-\eta)\rfloor) \ge D_{TV}(x_m, T_{TV}(x_m,\varepsilon)-1) \ge \varepsilon.$$

Take the limit inferior, we have

$$\varliminf_{m\to\infty} D_{TV}(x_m, \lfloor T_{TV}(x_m,\delta)(1-2\eta)\rfloor) \ge 1.$$

By the second inequality of (2.2), we have

$$\varepsilon \ge D_{TV}(x_m, T_{TV}(x_m,\varepsilon)) \ge D_{TV}(x_m, \lceil T_{TV}(x_m,\delta)(1+\eta)\rceil).$$

This implies

$$\varlimsup_{m\to\infty} D_{TV}(x_m, \lceil T_{TV}(x_m,\delta)(1+\eta)\rceil) \le 0.$$

For the converse, at first, we have observed if $\varepsilon \geq \delta_0$, then

$$\overline{\lim_{m\to\infty}} \frac{T_{TV}(x_m,\varepsilon)}{T_{TV}(x_m,\delta_0)} \leq 1. \tag{2.3}$$

And if $\varepsilon \leq \delta_0$, then

$$\underline{\lim_{m\to\infty}} \frac{T_{TV}(x_m,\varepsilon)}{T_{TV}(x_m,\delta_0)} \geq 1. \tag{2.4}$$

Fix $c > 1$. Given $\varepsilon \in [\delta_0, 1)$, there exists a $M > 0$ such that for all $m \geq M$

$$D_{TV}(x_m, \lceil ca_m \rceil) < \delta_0 < D_{TV}(x_m, T_{TV}(x_m, \delta_0) - 1),$$

and

$$1 - D_{TV}(x_m, \lfloor \frac{1}{c} a_m \rfloor) < 1 - \varepsilon < 1 - D_{TV}(x_m, T_{TV}(x_m, \varepsilon)).$$

Those imply

$$\lfloor \frac{1}{c} a_m \rfloor \leq T_{TV}(x_m, \varepsilon) \leq T_{TV}(x_m, \delta_0) \leq \lceil ca_m \rceil + 1, \ \forall\, m \geq M.$$

According the last inequality, we know $a_m \longrightarrow \infty$. Besides, we get

$$\frac{T_{TV}(x_m,\varepsilon)}{T_{TV}(x_m,\delta_0)} \geq \frac{\lfloor \frac{1}{c} a_m \rfloor}{\lceil ca_m \rceil + 1}, \ \forall\, m \geq M.$$

Take limit inferior

$$\underline{\lim_{m\to\infty}} \frac{T_{TV}(x_m,\varepsilon)}{T_{TV}(x_m,\delta_0)} \geq \frac{1}{c^2}, \ \forall\, c > 1.$$

By (2.3), we have

$$\lim_{m\to\infty} \frac{T_{TV}(x_m,\varepsilon)}{T_{TV}(x_m,\delta_0)} = 1, \ \forall\, \varepsilon \in [\delta_0, 1).$$

Given $\varepsilon \in (0, \delta_0)$, there is a $M' > 0$ such that for all $m \geq M'$

$$D_{TV}(x_m, \lceil ca_m \rceil) < \varepsilon \leq D_{TV}(x_m, T_{TV}(x_m, \varepsilon) - 1),$$

and

$$1 - D_{TV}(x_m, \lfloor \frac{1}{c} a_m \rfloor) < 1 - \delta_0 \leq 1 - D_{TV}(x_m, T_{TV}(x_m, \delta_0)).$$

We have inequalities

$$\lfloor \frac{1}{c} a_m \rfloor \leq T_{TV}(x_m, \delta_0) \leq T_{TV}(x_m, \varepsilon) \leq \lceil ca_m \rceil + 1, \ \forall\, m \geq M'.$$

Hence, we will similarly have

$$\overline{\lim_{m\to\infty}} \frac{T_{TV}(x_m,\varepsilon)}{T_{TV}(x_m,\delta_0)} \leq c^2, \ \forall\, c > 1.$$

By (2.4), we have

$$\lim_{m\to\infty} \frac{T_{TV}(x_m,\varepsilon)}{T_{TV}(x_m,\delta_0)} = 1, \ \forall\, \varepsilon \in (0, \delta_0).$$

$\square$

Remark 2.4. *The Proposition 2.3 also holds in continuous time cases without the assumption* (2.1).

Definition 2.4. *We called $\mathcal{F}$ has a total variation precutoff if there exists constants $c \geq 1$ and $\varepsilon > 0$ s.t. for all $0 < \delta < \varepsilon$,*

$$\overline{\lim_{m\to\infty}} \frac{T_{TV}(x_m,\delta)}{T_{TV}(x_m,\varepsilon)} \leq c. \tag{2.5}$$

Proposition 2.4. *Assume* (2.1) *holds. $\mathcal{F}$ has a total variation precutoff if and only if there is a sequence of positive integers (a_m) and $c > 1$ s.t.*

$$\overline{\lim_{m\to\infty}} D_{TV}(x_m, \lceil ca_m \rceil) = 0 \quad \textit{and} \quad \underline{\lim_{m\to\infty}} D_{TV}(x_m, \lfloor a_m \rfloor) > 0. \tag{2.6}$$

Proof. Suppose $\mathcal{F}$ has a precutoff. Let $\varepsilon_1 = \min\{\varepsilon, \delta_0\}$. By (2.5), given $\delta \in (0, \varepsilon_1)$, we may find a $M > 0$ s.t.

$$\sup_{m\geq M} \frac{T_{TV}(x_m,\delta)}{T_{TV}(x_m,\varepsilon_1)} \leq 2c.$$

This implies

$$T_{TV}(x_m,\varepsilon_1) \leq T_{TV}(x_m,\delta) \leq 2cT_{TV}(x_m,\varepsilon_1),\ \forall\ m \geq M. \tag{2.7}$$

By the second inequality of (2.7),

$$D_{TV}(x_m, \lceil 2cT_{TV}(x_m,\varepsilon_1)\rceil) \leq D_{TV}(x_m, T_{TV}(x_m,\delta)) \leq \delta,\ \forall\ m \geq M.$$

Then

$$\overline{\lim_{m\to\infty}} D_{TV}(x_m, \lceil 2cT_{TV}(x_m,\varepsilon_1)\rceil) \leq \delta,\ \forall\ \delta \in (0,\varepsilon_1).$$

Let δ tend to 0, we get

$$\overline{\lim_{m\to\infty}} D_{TV}(x_m, \lceil 2cT_{TV}(x_m,\varepsilon_1)\rceil) = 0.$$

Since $T_{TV}(x_m,\delta_0) \leq T_{TV}(x_m,\varepsilon_1)$, we know $T_{TV}(x_m,\varepsilon_1) \longrightarrow \infty$. Therefore,

$$\frac{T_{TV}(x_m,\varepsilon_1) - 1}{T_{TV}(x_m,\varepsilon_1)} \longrightarrow 1.$$

Choose $M' > 0$ s.t., for all $m \geq M'$,

$$\frac{1}{2c} T_{TV}(x_m,\varepsilon_1) \leq T_{TV}(x_m,\varepsilon_1) - 1.$$

Then

$$D_{TV}(x_m, \lfloor \frac{1}{2c} T_{TV}(x_m,\varepsilon_1)\rfloor) \geq D_{TV}(x_m, T_{TV}(x_m, \lfloor \varepsilon_1) - 1\rfloor) \geq \varepsilon_1 > 0,\ \forall\ m \geq M'.$$

It implies

$$\underline{\lim_{m\to\infty}} D_{TV}(x_m, \lfloor \frac{1}{2c} T_{TV}(x_m,\varepsilon_1)\rfloor) > 0$$

Choosing $a_m = \lfloor \frac{1}{2c} T_{TV}(x_m, \varepsilon_1) \rfloor$ and the new c is $2c^2 > 1$ gives the desired (2.6).

To show the converse, suppose there are a sequence (a_m) and $c > 1$ s.t. (2.6) holds. Given $\varepsilon \in (0, \delta_0)$, we can find $M > 0$ s.t.

$$\sup_m D_{TV}(x_m, \lceil ca_m \rceil) < \varepsilon < \delta_0 \leq D_{TV}(x_m, T_{TV}(x_m, \delta_0) - 1),\ \forall\, m \geq M.$$

Hence, we know

$$\lceil ca_m \rceil \geq T_{TV}(x_m, \delta_0) - 1,\ \forall\, m \geq M.$$

Then $a_m \longrightarrow \infty$.

By the inequality of (2.6), there exists a $0 < \varepsilon \leq \varliminf_{m\to\infty} D_{TV}(x_m, \lfloor a_m \rfloor)$. Let $\delta \in (0, \varepsilon)$, we find a $M^{'} > 0$ s.t.

$$\sup_m D_{TV}(x_m, \lceil ca_m \rceil) \leq \delta < \varepsilon \leq \inf_m D_{TV}(x_m, \lfloor a_m \rfloor),\ \forall\, m \geq M^{'}.$$

Observe that for all $m \geq M^{'}$,

$$\sup_m D_{TV}(x_m, \lceil ca_m \rceil) \geq D_{TV}(x_m, \lceil ca_m \rceil + 1),$$

and

$$\inf_m D_{TV}(x_m, \lfloor a_m \rfloor) \leq D_{TV}(x_m, \lfloor a_m \rfloor - 1).$$

Therefore, we have following inequalities

$$\lceil ca_m \rceil + 1 \geq T_{TV}(x_m, \delta) \geq T_{TV}(x_m, \varepsilon) \geq \lfloor a_m \rfloor - 1,\ \forall\, m \geq M^{'}.$$

And then

$$\frac{T_{TV}(x_m, \delta)}{T_{TV}(x_m, \varepsilon)} \leq \frac{\lceil ca_m \rceil + 1}{\lfloor a_m \rfloor - 1},\ \forall\, m \geq M^{'}.$$

Hence

$$\varlimsup_{m\to\infty} \frac{T_{TV}(x_m, \delta)}{T_{TV}(x_m, \varepsilon)} \leq c.$$

□

Remark 2.5. *If $a_m \longrightarrow \infty$, then it makes no difference to replace $\lfloor\ \rfloor$ to $\lceil\ \rceil$ or $\lceil\ \rceil$ to $\lfloor\ \rfloor$.*

Remark 2.6. *The definition of the total variation precutoff also applies for continuous times and Proposition 2.4 holds without the assumption* (2.1).

For two sequences of positive integers, s_m and t_m, $s_m = O(t_m)$ means that there are $C > 0$ and $M > 0$ such that $s_m \leq Ct_m$ for all $m \geq M$. $s_m \asymp t_m$ means that $s_m = O(t_m)$ and $t_m = O(s_m)$.

Theorem 2.1. *Fix a sequence (x_m) for all $x_m \in \mathcal{S}_m$. Assume (2.1) holds, and the sequence (a_m) satisfying $a_m \asymp T_{TV}(x_m, \varepsilon)$ for all $\varepsilon > 0$. We say $\mathcal{F}$ has a total variation precutoff if and only if there is a constant $c > 1$ s.t.*

$$\overline{\lim_{m\to\infty}} D_{TV}(x_m, \lceil ca_m \rceil) = 0. \tag{2.8}$$

Proof. Assume $\mathcal{F}$ has a precutoff. By definition 2.4, we have constants $c > 1$, $\varepsilon > 0$ and let $\varepsilon_1 = \min\{\varepsilon, \delta_0\}$; therefore, when we give a $\delta \in (0, \varepsilon_1)$, there is a $M_1 > 0$ s.t.

$$T_{TV}(x_m, \varepsilon_1) \leq T_{TV}(x_m, \delta) \leq cT_{TV}(x_m, \varepsilon), \ \forall\ m \geq M_1. \tag{2.9}$$

And there are two constants $0 < A,\ M_2 < \infty$ s.t. for all $m \geq M_2$,

$$T_{TV}(x_m, \varepsilon_1) \leq \lceil Aa_m \rceil.$$

Let $M = \max\{M_1, M_2\}$. By (2.9), for all $m \geq M$,

$$T_{TV}(x_m, \delta) \leq cT_{TV}(x_m, \varepsilon_1) \leq c\lceil Aa_m \rceil.$$

By above two inequalities, we have

$$D_{TV}(x_m, c\lceil Aa_m \rceil) \leq D_{TV}(x_m, T_{TV}(x_m, \delta)) \leq \delta, \ \forall\ m \geq M.$$

Then

$$\overline{\lim_{m\to\infty}} D_{TV}(x_m, c\lceil Aa_m \rceil) = 0.$$

To show the converse, assume there is $c > 1$ s.t. (2.8) holds. We have given $\delta_0 > 0$, then there is a $M_0 > 0$ s.t. for all $m \geq M_0$, $D_{TV}(x_m, \lceil ca_m \rceil) < \delta_0$. And then $\lceil ca_m \rceil \geq T_{TV}(x_m, \delta_0)$ for all $m \geq M_0$. We have $a_m \longrightarrow \infty$.

Since $a_m \asymp T_{TV}(x_m, \varepsilon)$ for all $\varepsilon > 0$, we fix $\varepsilon > 0$, there are $0 < c' < 1$ and $M_1 > 0$ s.t.

$$\inf_{m \geq M_1} \frac{T_{TV}(x_m, \varepsilon)}{a_m} \geq c' > 0.$$

That is

$$T_{TV}(x_m, \varepsilon) \geq c' a_m, \ \forall\ m \geq M_1.$$

Let $\delta \in (0, \varepsilon)$, and we can find a $M_2 > 0$ s.t.

$$\sup_{m \geq M_2} D_{TV}(x_m, \lceil ca_m \rceil) < \delta < \varepsilon.$$

Let $M = \max\{M_1, M_2\}$. We have

$$\lceil ca_m \rceil \geq T_{TV}(x_m, \delta) \geq T_{TV}(x_m, \varepsilon) \geq c' a_m, \ \forall\ m \geq M.$$

Then

$$\frac{T_{TV}(x_m,\delta)}{T_{TV}(x_m,\varepsilon)} \le \frac{\lceil ca_m \rceil + 1}{c' a_m}, \ \forall\, m \ge M.$$

And we have

$$\varlimsup_{m\to\infty} \frac{T_{TV}(x_m,\delta)}{T_{TV}(x_m,\varepsilon)} \le \frac{c}{c'}.$$

□

Corollary 2.1. *[4, Corrollary 2.3] Suppose that there is $\varepsilon > 0$ and $a_m \to \infty$ such that $T_{TV}(x_m,\varepsilon) \asymp a_m$ and $T_{TV}(x_m,\delta) = O(a_m)$ for all $0 < \delta < \varepsilon$. Then, the following are equivalent:*

1. *No subfamily of $\mathcal{F}$ has a total variation precutoff.*
2. *For all $c > 0$,*

$$\varliminf_{m\to\infty} D_{TV}(x_m, \lfloor ca_m \rfloor) > 0.$$

3. *As $\delta \to 0$,*

$$\varliminf_{m\to\infty} \frac{T_{TV}(x_m,\delta)}{a_m} = \infty.$$

Remark 2.7. *The Theorem 2.1 still holds in the continuous time case without the assumption* (2.1).

Remark 2.8. *Definitions of L^p-cutoff and L^p-precutoff, $1 < p \le \infty$, are similar to the cutoff in total variation. All propositions and theorems in this subsection hold for the L^p-cutoff and L^p-precutoff. See [1] for details.*

3 Bounding the mixing time in total variation

3.1 Spectral analysis of reversible Markov chains

We perform the spectral information to describe the transition matrix.

Lemma 3.1. *[3, Lemma 1.2.9] Let K be an irreducible Markov kernel on Ω with stationary distribution π. Suppose that K is reversible and $\beta_0 = 1, \beta_1, \ldots, \beta_{|\Omega|-1}$ are eigenvalues of K with associated $L^2(\pi)$-orthonormal eigenvectors $\psi_0 = \mathbf{1}, \psi_1, \ldots, \psi_{|\Omega|-1}$. Then, for $x, y \in \Omega$,*

$$\frac{K^n(x,y)}{\pi(y)} = \sum_{i=0}^{|\Omega|-1} \beta_i^n \psi_i(x)\psi_i(y)$$

and

$$\frac{e^{-t(I-K)}(x,y)}{\pi(y)} = \sum_{i=0}^{|\Omega|-1} e^{-t(1-\beta_i)} \psi_i(x)\psi_i(y).$$

Proposition 3.1. *[3, Lemma 1.3.3] Let K be an irreducible and reversible Markov kernel on a finite set Ω with a stationary distribution π. If $\beta_0 = 1, \beta_1, \ldots, \beta_{|\Omega|-1}$ are eigenvalues of K and $\psi_0 = \mathbf{1}, \psi_1, \ldots, \psi_{|\Omega|-1}$ are corresponding $L^2(\pi)$-orthonormal eigenvectors, then*

$$D_2(\mu, n) = \left(\sum_{i \geq 1} |\mu(\psi_i)|^2 \beta_i^{2n} \right)^{\frac{1}{2}}$$

and

$$D_2^c(\mu, t) = \left(\sum_{i \geq 1} |\mu(\psi_i)|^2 e^{-2t(1-\beta_i)} \right)^{\frac{1}{2}}.$$

Proposition 3.1 gives us the formula of L^2-distance, then, by Proposition 2.1, we make L^2-distance be the upper bound of total variation distance. And we use the following proposition to get a lower bound.

Proposition 3.2. *Let Ω be a finite set and μ, ν be probabilities on Ω. Assume that f is a function on Ω (complex values are allowed) satisfying $\mu(f) \neq 0$ and $\nu(f) = 0$. Then*

$$\|\mu - \nu\|_{TV} \geq 1 - \frac{4(Var_\mu(f) - Var_\nu(f))}{|\mu(f)|^{\frac{1}{2}}}.$$

Proof. Set $s = \frac{|\mu(f)|}{2}$ and $A = \{x \in \Omega \,|\, |f(x)| \geq s\}$. Then for $x \in A^c$, $|f(x) - \mu(f)| \geq s$. This implies

$$\mu(A^c) = \mu(I_{\{A^c\}}) \leq \mu(I_{\{A^c\}} \frac{|f - \mu(f)|^2}{s^2}) \leq \frac{Var_\mu(f)}{s^2}.$$

Using the Chebyshev inequality, it is obvious that

$$\nu(A) = \mathbb{P}_\nu(|f| \geq s) = \mathbb{P}_\nu(|f - \nu(f)| \geq s) \leq \frac{Var_\nu(f)}{s^2}.$$

Putting all above together gives the desired inequality. □

We give a examples to illustrate how to use them.

Example 3.1 (The continuous-time case)**.** *Let $G = (\mathbb{Z}_2)^n$ and $\mathbb{P} : G \longrightarrow [0, 1]$ be defined by $\mathbb{P}(e_i) = \frac{1}{n}$, for all $1 \leq i \leq n$, where e_i is the vector which only the ith entry is 1 and others are 0. Let $\beta_{n,x} = 1 - \frac{2|x|}{n}$ and $\phi_{n,x}(y) = (-1)^{x \cdot y}$.*

$$D_{TV}^c(\delta_x, t) \leq D_2^c(\delta_x, t) = \sqrt{\sum_{i=1}^{n} \binom{n}{i} e^{\frac{-4ti}{n}}} = \sqrt{(1 + e^{\frac{-4t}{n}})^n - 1}.$$

Clearly,

$\lim_{n\to\infty} D_2^c(\delta_x, \frac{1}{4}n(\log n + c)) = \sqrt{e^{e^{-c}} - 1}$
$\Rightarrow T_{TV}^c(\varepsilon) \le T_2^c(\varepsilon) = \frac{1}{4}n\log n + O_\varepsilon(n)$ *as* $n \longrightarrow \infty$ *for all* $\varepsilon > 0$.

Set $f = \sum_{i=1}^{n} \phi_{n,e_i}$. *That is,* $f(y) = \sum_{i=1}^{n} (-1)^{y_i}$ *for* $y = (y_1, y_2, \ldots, y_n)$. *By letting* $\mu = H_t(0,\cdot)$ *and* $\nu = \pi$, *we have*

$$\mu(f) = ne^{-2t/n}, \quad \nu(f) = 0, \quad \mu(f^2) = n(n-1)e^{-4t/n} + n, \quad \nu(f^2) = n.$$

By proposition 3.2,

$$D_{TV}^c(0,t) \ge 1 - \frac{4n(2 - e^{-4t/n})}{n^2e^{-4t/n}} \ge 1 - \frac{8}{ne^{-4t/n}}$$

and

$$\lim_{n\to\infty} D_{TV}^c(0, \frac{1}{4}n(\log n + c)) \ge 1 - 8e^c.$$

$\Rightarrow T_{TV}^c(\varepsilon) \ge \frac{1}{4}n\log n + O_\varepsilon(n)$ *as* $n \longrightarrow \infty$.
Consequently, we obtained $T_{TV}^c(\varepsilon) = \frac{1}{4}n\log n + O_\varepsilon(n)$. *Noted that* $T_2^c(\varepsilon) \ge T_{TV}^c(\varepsilon)$, *the* $T_2^c(\varepsilon) = \frac{1}{4}n\log n + O_\varepsilon(n)$.

3.2 Coupling of Markov chains

From the viewpoint of probability, the coupling is useful in bounding the total variation distance. The coupling provides a probabilistic character on the total variation distance.

Definition 3.1. *Let K be a transition matrix on a finite state space S. A coupling of Markov chains with K is a process $(X_t, Y_t)_{t\ge0}$ of which marginals X_t, Y_t are Markov chains with K and initial distributions μ, ν, respectively; besides, the coupling satisfies following.*

$$\text{If } X_t = Y_t, \text{ then } X_{t+k} = Y_{t+k} \text{ for all } k \ge 0.$$

And we define the coupling time $T := \inf\{t \ge 0 | X_t = Y_t\}$.

Remark 3.1. *We write $P_{\mu,\nu}$ if X_0, Y_0 have distributions μ, ν and simplify it as $P_{x,y}$ if $X_0 = x$ and $Y_0 = y$.*

Theorem 3.1. *Let (X_t, Y_t) be a coupling with $X_0 = x$ and $Y_0 \cong \pi$, where π is a stationary distribution. Then,*

$$D_{TV}(x,t) \le P_{x,\pi}(T > t).$$

Example 3.2. *Let $\Omega = \mathbb{Z}_n$ and K be the kernel of the lazy simple random walk on the n-cycle. That is, $K(x,x) = \frac{1}{2}$ and $K(x, x \pm 1 \pmod n) = \frac{1}{4}$ for all $x \in \Omega$. Consider the following* *coupling. Let* $\{W_k, Z_k | k = 0, 1, \dots\}$ *be independent fair coins, where 1 and 0 denote heads and tails. Set* $X_0 = x$, $Y_0 \cong \pi \equiv 1/n$ *and*

$$X_{k+1} = X_k + W_k \cdot (-1)^{Z_k} \pmod{n}, \forall k = 0, 1, \dots.$$

If $X_k \neq Y_k$*, define*

$$Y_{k+1} = Y_k + (1 - W_k) \cdot (-1)^{Z_k} \pmod{n},$$

while $Y_{k+1} = X_{k+1}$ *if* $X_k = Y_k$*. Noted that* X_k *and* Y_k *forms a Markovian coupling of lazy random walks on n-cycles.*

Consider the clockwise distance C_k *from* x *to* y*, that is,* $C_k = Y_k - X_k \pmod{n}$*. It is easy to check that* C_k *is a simple random walk on* $\{0,1,\dots,n\}$ *with absorbing states 0 and n and initial state* $y - x \pmod{n}$*. As a conclusion of Gambler's ruin problem, if* T *is the coupling time, then*

$$D_{TV}(k) \leq \max_x P_{x,\pi}(T > k) \leq \frac{\sup_x E_{x,\pi}[T]}{k} \leq \sup_{0 \leq i \leq n} \frac{i(n-i)}{k} \leq \frac{n^2}{4k},$$

where the last second inequality is derived from

$$E_{x,\pi}[T] = \frac{1}{n} \sum_y E_{x,\pi}[T | Y_0 = y] \leq \sup_{x,y} E_{x,y}[T].$$

This implies $T_{TV}(\varepsilon) \leq \lceil n^2/4\varepsilon \rceil$*. In fact, the exact order of the total variation mixing time is* n^2*. To see this, note that* $f(x) = \cos(2\pi x/n)$ *is an eigenvector of the transition matrix with eigenvalue* $[1 + \cos(2\pi/n)]/2$*. This implies*

$$\begin{aligned} 2D_{TV}(k) &= \sup_x D_1(x,k) = \sup_x \|K^k(x,\cdot)/\pi - 1\|_1 \\ &\geq \frac{\langle \frac{K^k(0,\cdot)}{\pi} - 1, f \rangle_\pi}{\|f\|_\infty} = K^k(0,f) = \left(\frac{1 + \cos(2\pi/n)}{2} \right)^k. \end{aligned}$$

Using the inequality $\cos\theta \geq 1 - \frac{\theta^2}{2}$ *for all* $\theta \in \mathbb{R}$ *implies that*

$$D_{TV}(k) \geq \tfrac{1}{2}(1 - \tfrac{\pi^2}{n^2})^k \geq \tfrac{1}{2} e^{-2\pi^2 k/n^2}, \quad \forall\ n \geq 7,$$

where the last inequality applies the fact $\ln(1 - \theta) \geq -2\theta$ *for* $\theta < 1/2$*. Thus,* $T_{TV}(\varepsilon) \geq [(\ln(1/2\varepsilon))/2\pi^2] n^2$*. Putting the upper and lower bounds together, one may conclude that the lazy simple random walk on n-cycles has no total variation cutoff.*

4 The total variation cutoff of Ehrenfest chains

At the beginning, we recall (1.1). The Ehrenfest chain on $\mathcal{S}_n = \{0, 1, \ldots, n\}$ with the transition matrix K_n satisfying

$$K_n(k, k+1) = 1 - \tfrac{k}{n}, \quad K_n(k+1, k) = \tfrac{k+1}{n}, \quad \forall\ 0 \le k < n,$$

and its stationary distribution is $\pi_n(k) = \begin{pmatrix} n \\ k \end{pmatrix} 2^{-n}$ for all $0 \le k \le n$. Concerning the periodicity, we consider $K_n^{'}$ in the discrete time case, given by

$$K_n^{'} = \frac{1}{n+1} I_n + \frac{n}{n+1} K_n,$$

where I_n is an $(n+1) \times (n+1)$ identity matrix. In the continuous time case, we consider the semigroup accosiated witn K_n i.e.

$$H_{n,t} = e^{-t(I-K_n)} = \sum_{j=0}^{\infty} \left(e^{-t} \frac{t^j}{j!} \right) K_n^j.$$

There is a result giving a description on the eigenvalues and eigenvectors of K_n in Chen and Saloff-Coaste[2].

Lemma 4.1. *[2, Theorem 6.1] The matrix K_n has eigenvalues*

$$\beta_{n,i} = 1 - \tfrac{2i}{n}, \quad \forall\ 0 \le i \le n,$$

with $L^2(\pi_n)$-normalized right eigenvectors

$$\psi_{n,i}(x) = \begin{pmatrix} n \\ i \end{pmatrix}^{-1/2} \sum_{k=0}^{i} (-1)^k \begin{pmatrix} x \\ k \end{pmatrix} \begin{pmatrix} n-x \\ i-k \end{pmatrix}, \quad \forall\ 0 \le i, x \le n.$$

Remark 4.1. *Krawtchouk polynomials[5] defined by*

$$P_i(x, p, n) = {}_2F_1 \left(\begin{array}{c|c} -i, -x & \\ & \frac{1}{p} \\ -n & \end{array} \right), \quad \forall\ i \in \{0, 1, \ldots, n\},$$

then the eigenvector $\psi_{n,i}$ of K_n can be rewritten as

$$\psi_{n,i}(x) = \begin{pmatrix} n \\ i \end{pmatrix}^{1/2} P_i(x, 1/2, n).$$

Remark 4.2. *By lemma 4.1, $K_n^{'}$ has eigenvalues $\beta_{n,i}^{'} = 1 - \frac{2i}{n+1}$ with corresponding eigenvectors $\psi_{n,i}$ given by K_n.*

4.1 The order of total variation mixing time

In this section, we treat the case $\frac{|n-2x_n|}{\sqrt{n}} \to B$ and derive the order of total variation mixing time for $B > 0$.

By proposition 2.1, $2D_{TV}(x_n, t) \le D_2(x_n, t)$ and $2D^c_{TV}(x_n, t) \le D^c_2(x_n, t)$, where

$$\begin{aligned}
[D_2(x_n,t)]^2 &= \sum_{i=1}^{n} |\psi_{n,i}(x_n)|^2 |1 - \frac{2i}{n+1}|^{2t} \\
&\le 2\sum_{i=1}^{\lfloor \frac{n}{2} \rfloor} |\psi_{n,i}(x_n)|^2 |1 - \frac{2i}{n+1}|^{2t} + |1 - \frac{2}{n+1}|^{2t} \\
&\le 2\sum_{i=1}^{\lfloor \frac{n}{2} \rfloor} |\psi_{n,i}(x_n)|^2 e^{-4ti/(n+1)} + e^{-4nt/(n+1)}, \qquad (4.1)
\end{aligned}$$

where the first inequality is derived from $\psi_{n,n-i}(x) = (-1)^x \psi_{n,i}(x)$ for all $x, i \in \{0, \dots, n\}$[2].

By recurrence relation of $P_i(x, 1/2, n)$, we have, for all $i \in \{1, \dots, n-1\}$,

$$\psi_{n,i+1}(x_n) = \frac{n-2x_n}{\sqrt{n}} A_{n,i}\psi_{n,i}(x_n) - B_{n,i}\psi_{n,i-1}(x_n),$$

where $A_{n,i} = \sqrt{\frac{n}{(i+1)(n-i)}}$, $B_{n,i} = \sqrt{\frac{i(n-i+1)}{(i+1)(n-i)}}$.

Observe that $A_{n,i} \le 1$, $B_{n,i} \le 1$ for all $1 \le i < n$. Let $\gamma = \sup_n \{|n-2x_n|/\sqrt{n}\} \vee 1 < \infty$. Then

$$|\psi_{n,i+1}(x_n)| \le \gamma |\psi_{n,i}(x_n)| + |\psi_{n,i-1}|, \quad \forall\ 1 \le i < n,\ n \ge 1.$$

Since we know the initial state $|\psi_{n,0}(x_n)| = 1$, we use induction to obtain, for all $i \in \{0, 1, \dots, n\}$,

$$|\psi_{n,i}(x_n)| \le (\gamma+1)^i.$$

Let $t = c(n+1)$, where $c \ge \frac{1}{4}\ln 2(\gamma+1)^2$, we have

$$\begin{aligned}
(4.1) &\le 2\sum_{i=1}^{\lfloor \frac{n}{2} \rfloor} \left[|(\gamma+1)|^2 e^{-4c}\right]^i + e^{-4nc} \le 2\sum_{i=1}^{\infty} \left[|(\gamma+1)|^2 e^{-4c}\right]^i \\
&= 2\left[\frac{(\gamma+1)^2 e^{-4c}}{1-(\gamma+1)^2 e^{-4c}}\right] \le 2\left[\frac{(\gamma+1)^2 e^{-4c}}{1-\frac{1}{2}}\right] \\
&\le 4(\gamma+1)^2 e^{-4c}.
\end{aligned}$$

Since $2D_{TV}(x_n, t) \le D_2(x_n, t)$, $D_{TV}(x_n, c(n+1)) \le (\gamma+1)e^{-2c}$ for all $c \ge \frac{1}{4}\ln 2(\gamma+1)^2$. Then, for all $\varepsilon \in (0, 1)$,

$$T_{TV}(x_n, \varepsilon) \le \frac{1}{2}(\ln\frac{\gamma+1}{\varepsilon})(n+1).$$

Recall the proposition 3.2, let $f(x) = \psi_{n,1}(x)$, $\mu = {K'_n}^m(x_n, \cdot)$ and $\nu = \pi_n$.

$\mu(f) = (\beta'_{n,1})^m \psi_{n,1}(x_n) = (1 - \frac{2}{n+1})^m (\frac{n-2x_n}{\sqrt{n}})$.

$\nu(f) = \sum_{k=0}^{n} \psi_{n,1}(k)\pi(k) = \sum_{k=0}^{n} (\frac{n-2k}{\sqrt{n}}) \begin{pmatrix} n \\ k \end{pmatrix} 2^{-n} = 0$

$\mu(f^2) = \sum_{k=0}^{n} \psi_{n,1}^2(k) K_n^m(x_n, k) = \sum_{k=0}^{n} (c_1 \psi_{n,2}(k) + c_2) K_n^m(x_n, k) = c_1 \beta_{n,2}^m \psi_{n,2}(x_n) + c_2,$

where $c_1 = \sqrt{2(1 - 1/n)}$ and $c_2 = 1$.

$$\begin{aligned} \nu(f^2) &= \sum_{k=0}^{n} \psi_{n,1}^2(k)\nu_n(k) = \sum_{k=0}^{n} \frac{(n-2k)^2}{n} \begin{pmatrix} n \\ k \end{pmatrix} (\frac{1}{2})^n \\ &= \frac{1}{n} \sum_{k=0}^{n} (n^2 - 4nk + 4k^2) \begin{pmatrix} n \\ k \end{pmatrix} (\frac{1}{2})^n \\ &= \frac{1}{n} \left[n^2 - 4n \sum_{k=0}^{n} k \begin{pmatrix} n \\ k \end{pmatrix} (\frac{1}{2})^n + 4 \sum_{k=0}^{n} k^2 \begin{pmatrix} n \\ k \end{pmatrix} (\frac{1}{2})^n \right] = 1 \end{aligned}$$

Hence, $Var_\mu(f) = (1 - \frac{4}{n+1})^m \left[(\frac{n-2x_n}{\sqrt{n}})^2 - 1 \right] + 1 - (\frac{n-2x_n}{\sqrt{n}})^2 (1 - \frac{2}{n+1})^{2m}$ and $Var_\nu(f) = 1$.

$$D_{TV}(x_n, m) \geq 1 - 4 \cdot \frac{2 + (1 - \frac{4}{n+1})^m [(\frac{n-2x_n}{\sqrt{n}})^2 - 1] - (\frac{n-2x_n}{\sqrt{n}})^2 (1 - \frac{2}{n+1})^{2m}}{(1 - \frac{2}{n+1})^{2m} (\frac{n-2x_n}{\sqrt{n}})^2}$$

Let $m = c(n+1)$, $c > 0$. Then

$$D_{TV}(x_n, m) \geq 1 - 4 \cdot \frac{2 + (1 - \frac{4}{n+1})^{c(n+1)} [(\frac{n-2x_n}{\sqrt{n}})^2 - 1] - (\frac{n-2x_n}{\sqrt{n}})^2 (1 - \frac{2}{n+1})^{2c(n+1)}}{(1 - \frac{2}{n+1})^{2c(n+1)} (\frac{n-2x_n}{\sqrt{n}})^2}$$

We have

$$\varliminf_{n \to \infty} D_{TV}(x_n, c(n+1)) \geq 1 - \frac{4}{B^2}(2e^{4c} - 1).$$

Therefore, $T_{TV}(x_n, \varepsilon) \asymp n$ for all ε small enough and B large enough if $|n - 2x_n|/\sqrt{n} \to B$, where $0 < B < \infty$.

By proposition 3.1 and lemma 4.1,

$$\begin{aligned} [D_2^c(x_n, t)]^2 &= \sum_{i=1}^{n} |\psi_{n,i}(x_n)|^2 e^{-4ti/n} \leq \sum_{i=1}^{n} |\psi_{n,i}(x_n)|^2 e^{-4ti/(n+1)} \\ &\leq 4(\gamma + 1)^2 e^{-4c}, \end{aligned}$$

where the last inequality is given by the discrete time case when $t = c(n+1)$, $c \geq \frac{1}{4}\ln 2(\gamma+1)^2$.

In the similar way, we use proposition 3.2, let $f(x) = \psi_{n,1}(x)$, $\mu = H_{n,t}(x_n, \cdot)$, $\nu = \pi_n$.

$\mu(f) = e^{-t(1-\beta_{n,1})}\psi_{n,1}(x_n) = e^{-2t/n}(\frac{n-2x_n}{\sqrt{n}})$.

$\nu(f) = 0$.

$\mu(f^2) = c_1 e^{-t(1-\beta_{n,2})}\psi_{n,2}(x_n) + c_2$, where $c_1 = \sqrt{2(1-1/n)}$ and $c_2 = 1$.

$\nu(f^2) = 1$.

Hence, $Var_\mu(f) = e^{-4t/n}\left[(\frac{n-2x_n}{\sqrt{n}})^2 - 1\right] + 1 - e^{-4t/n}(\frac{n-2x_n}{\sqrt{n}})^2$ and $Var_\nu(f) = 1$.
Let $t = cn$, $c > 0$, we obtain

$$\underline{\lim}_{n\to\infty} D_{TV}(x_n, c(n+1)) \geq 1 - \frac{4}{B^2}(2e^{4c} - 1).$$

Observe that the continuous time case has the same bound with the discrete time case, $T^c_{TV}(x_n, \varepsilon) \asymp n$ for all ε small enough and B large enough. If $\gamma < \frac{1}{4}\sqrt{B^2+4} - 1$, then there is no total variation cutoff. But it's not sufficient to conclude the precutoff.

For $B = 0$, we can find the upper bound of the total variation distance,

$$D_{TV}(x_n, c(n+1)) \leq (\gamma+1)e^{-2c},$$

where $c \leq \frac{1}{4}\ln 2(\gamma+1)^2$. But the lower bound is not available to use the same method.

4.2 Proof of Theorem 1.3

We have known $(iii) \Rightarrow (ii)$ by Chen[2] and $(ii) \Rightarrow (i)$ by their definitions. Therefore, it remains to show $(i) \Rightarrow (iii)$. For that purpose, our idea is to show the contrapositive true. Suppose $|n - 2x_n|/\sqrt{n}$ is bounded for all $n > 0$, we need to show $\mathcal{F}$ and $\mathcal{F}_c$ have no total variation precutoff. The proof is divided into two cases.

1. The discrete time case.

The discrete time Ehrenfest chain K_n' still has a probability $n/(n+1)$ to develop to the chain K_n, and it's more possible than staying still. The period of a reversible and irreducible chain is just 1 or 2. Since a reversibility, a chain is periodic if and only if -1 is an eigenvalue of its transition matrix. By lemma 4.1, the respective eigenvector is $\psi_{n,n}(x) = (-1)^x$. Let $A_n = \{x \in \mathcal{S}_n | x \text{ is even}\}$. Clearly, $\psi_{n,n} = 2 \cdot I_{A_n} - 1$.

$$\begin{aligned}
D_{TV}(x_n, m) &\geq |K_n'^{\,m}(x_n, A_n) - \pi_n(A_n)| \\
&= \frac{1}{2}|[K_n'^{\,m}(x_n, \cdot) - \pi_n(\cdot)](2 \cdot I_{A_n} - 1)| \\
&= \frac{1}{2}\left|\sum_{y=0}^{n}\sum_{i=0}^{n} \beta_{n,i}'^{\,m} \psi_{n,i}(x_n)\psi_{n,i}(y)\psi_{n,n}(y) - \sum_{y=0}^{n} \pi_n(y)\psi_{n,n}(y)\right| \\
&= \frac{1}{2}|(-1)^{x_n}\beta_{n,n}'^{\,m}| = \frac{1}{2}|1 - \frac{2n}{n+1}|^m \\
&\geq \frac{1}{2}e^{-2m/(n+1)},
\end{aligned}$$

for $n \geq 3$, where the last inequality is given by $\ln(1-t) \geq -2t$ for $t \in [0, 1/2]$.

Then, for $0 < \varepsilon \leq \frac{1}{2e^2}$ and $n \geq 3$,

$$T_{TV}(x_n, \varepsilon) \geq \lfloor \frac{1}{2}\ln(\frac{1}{2\varepsilon}) \rfloor (n+1).$$

It's important that we don't need to assume $|n - 2x_n|/\sqrt{n}$ converge to a nonzero real value. By subsection 4.1, we have $T_{TV}(x_n, \varepsilon) \asymp n$, $0 < \varepsilon \leq \frac{1}{2e^2}$. For $n \geq 3$ and $c \geq 1$,

$$D_{TV}(x_n, \lfloor cn \rfloor) \geq D_{TV}(x_n, \lfloor c(n+1) \rfloor) \geq \frac{1}{2}e^{-\lfloor 2c \rfloor}.$$

By Corollary 2.1, there is no total variation precutoff.

2. The continuous time case.

By Corollary 2.1, we need to show

$$\varliminf_{n\to\infty} D_{n,TV}^c(x_n, an) > 0, \quad \forall\, a > 0.$$

At the beginning, we consider the Ehrenfest chain $(X_n)_{n=0}^{\infty}$ on $\mathcal{S}_{2n} = \{0, 1, \ldots, 2n\}$ and a sequence (x_{2n}) satisfying $|2n - 2x_{2n}|/\sqrt{2n} \to c$.

Part 1. $c > 0$

By the symmetry of Ehrenfest chain, we assume the initial states $x_{2n} = n - c_n\sqrt{n}$, $c_n \to c/\sqrt{2}$ and choose the proper testing set $A_{2n} = \{0, 1, \ldots, n\}$ s.t.

$$\begin{aligned}
\|H_{2n,t}(x_{2n}, \cdot) - \pi_{2n}(\cdot)\|_{TV} &\geq H_{2n,t}(x_{2n}, A_{2n}) - \pi_{2n}(A_{2n}) \\
&= e^{-t}\sum_{j=0}^{\infty}\frac{t^j}{j!}K_{2n}^j(x_{2n}, A_{2n}) - \pi_{2n}(A_{2n}). \qquad (4.2)
\end{aligned}$$

Set $T_n = \inf\{t \geq 0 | X_t = n\}$. Using the strong Markov property, the $K_{2n}^j(x_{2n}, A_{2n})$ can be derived as follow.

$$\begin{aligned} K_{2n}^j(x_{2n}, A_{2n}) &= \sum_{i=0}^{\infty} P_{x_{2n}}(X_j \in A_{2n} | T_n = i) P_{x_{2n}}(T_n = i) \\ &= \sum_{i=0}^{j} K_{2n}^{j-i}(n, A_{2n}) P_{x_{2n}}(T_n = i) + P_{x_{2n}}(T_n > j) \end{aligned} \tag{4.3}$$

The next inequality will apply the following proposition.

Proposition 4.1. *Let K_n is the transition matrix of the Ehrenfest chain on $\mathcal{S}_n$. Suppose $A = \{0, 1, \ldots, \lceil n/2 \rceil\}$. Then $K_n^t(x, A) \geq 1/2$ for all $x \in A$, $t \geq 0$.*

See the appendix for a proof of this proposition. Using Proposition 4.1, we have

$$(4.3) \geq \frac{1}{2} P_{x_{2n}}(T_n \leq j) + P_{x_{2n}}(T_n > j) = \frac{1}{2} + \frac{1}{2} P_{x_{2n}}(T_n > j).$$

Observe that $\pi_{2n}(A_{2n}) = \frac{1}{2} + \frac{1}{2}\pi_{2n}(n) = \frac{1}{2} + O(\frac{1}{\sqrt{n}})$. And then we combine above information to (4.2),

$$\begin{aligned} D_{TV}^c(x_{2n}, t) &\geq \frac{1}{2} e^{-t} \sum_{j=0}^{\infty} \frac{t^j}{j!} P_{x_{2n}}(T_n > j) + O(\frac{1}{\sqrt{n}}) \\ &\geq \frac{1}{2} e^{-t} \sum_{j=0}^{l} \frac{t^j}{j!} P_{x_{2n}}(T_n > l) + O(\frac{1}{\sqrt{n}}), \end{aligned} \tag{4.4}$$

where for all $l \geq 0$.

Note that if we let $l = \lceil bn \rceil$, $b > a > 0$, then

$$\varliminf_{n \to \infty} D_{TV}^c(x_{2n}, an) \geq \frac{1}{2} \varliminf_{n \to \infty} P_{x_{2n}}(T_n > \lceil bn \rceil).$$

That exhibits that we suffice to show $\varliminf_{n \to \infty} P_{x_{2n}}(T_n > \lceil bn \rceil) \geq C$ for some $C > 0$. For that reason, we compare the Ehrenfest chain with the simple random walk on $\mathbb{Z}$.

Set the $T = \inf\{m \geq 0 | X_m = n \text{ or } X_m = 2x_{2n} - n\}$. Clearly, $P_{x_{2n}}(T_n > l) \geq P_{x_{2n}}(T > l)$ for all $l \geq 0$. Let $(Y_m)_{m=0}^{\infty}$ be the simple random walk on $\mathbb{Z}$ and $\widetilde{T} = \inf\{m \geq 0 | Y_m = n \text{ or } Y_m = 2x_{2n} - n\}$. Write

$$P_{x_{2n}}(T > l) \geq \alpha_n(l) \widetilde{P}_{x_{2n}}(\widetilde{T} > l),$$

where

$$\alpha_n(l) = \left(1 - \frac{2n - 2x_{2n}}{n}\right)^{n - x_{2n}} \left[1 - \left(\frac{2n - 2x_{2n}}{n}\right)^2\right]^l.$$

We call w is a path if w is a sequence $(w_i)_{i=0}^{\infty}$ for all $w_i \in \mathbb{Z}$ satisfying $|w_{i+1} - w_i| = 1$ for all $i \geq 0$ and $2x_{2n} - n < w_i < n$ for all $0 \leq i \leq l$. And we use the notion (w_i, w_{i+1}) to describe a edge in w from w_i to w_{i+1}. Clearly, the probability of (w_i, w_{i+1}) is $K_{2n}(w_i, w_{i+1})$. Define $p(x : y, l)$, $\forall x, y \in \{2x_{2n} - n + 1, \ldots, n - 1\}$, $l \in \mathbb{Z}^+$, be a collection of paths of length l which started at x and ended at y. Hence

$$P_{x_{2n}}(T > l) = \sum_{2x_{2n} - n < k < n} P_{x_{2n}}(p(x_{2n} : k, l)).$$

In each $w \in p(x_{2n} : k, l)$ for $2x_{2n} - n < k < n$, we collect all edges of w and we can partition this collection into two subcollections. One is $A(w) = \{(i, i+1) | x_{2n} \leq i < k\}$ if $k > x_{2n}$; instead, $A(w) = \{(i, i-1) | x_{2n} \geq i > k\}$. And the other one $B(w)$ is a union of $\{(i, i+1), (i+1, i)\}$ for $2x_{2n} - n + 1 < i < n - 1$. For all $2x_{2n} - n < i < n$,

$$1 - \frac{i}{2n} \geq \frac{i}{2n} \geq \frac{1}{2}\left(\frac{2x_{2n} - n}{n}\right) = \frac{1}{2}\left(1 - \frac{2n - 2x_{2n}}{n}\right),$$

and

$$\begin{aligned} \frac{i}{2n}\left(1 - \frac{i-1}{2n}\right) = \frac{2ni - i^2 + i}{4n^2} &\geq \frac{1}{4}\left(\frac{n^2 - n^2 + 2ni - i^2}{n^2}\right) \\ &\geq \frac{1}{4}\left[1 - \left(\frac{2n - 2x_{2n}}{n}\right)^2\right]. \end{aligned}$$

Now we can make a conclusion that $P_{x_{2n}}(w) \geq \alpha_n(l)\widetilde{P}_{x_{2n}}(w)$ for all $w \in p(x_{2n} : k, l)$ and $2x_{2n} - n < k < n$, where

$$\alpha_n(l) = \left(1 - \frac{2n - 2x_{2n}}{n}\right)^{n - x_{2n}} \left[1 - \left(\frac{2n - 2x_{2n}}{n}\right)^2\right]^l.$$

Let $b = \lceil a + \frac{c^2}{2} \rceil$,

$$\begin{aligned} P_{x_{2n}}(T > bn) &\geq \alpha_n(bn)\widetilde{P}_{x_{2n}}(\widetilde{T} > bn) \\ &= \alpha_n(bn)\widetilde{P}_0(\widetilde{T}_{\lfloor c_n\sqrt{n}\rfloor} > bn) \\ &\geq \alpha_n(bn)\exp\left\{\frac{-2b}{\lfloor c_n \rfloor^2}\right\}. \end{aligned}$$

where $\widetilde{T}_{\lfloor c_n\sqrt{n}\rfloor} = \inf\{m \geq 0 | Y_m = \lfloor c_n\sqrt{n}\rfloor \text{ or } Y_m = -\lfloor c_n\sqrt{n}\rfloor\}$ and the last inequality is given by Lemma A.1. Applying this to (4.4), we have

$$D_{TV}^c(x_{2n}, an) \geq \frac{1}{2}e^{-an}\left(\sum_{j=0}^{bn} \frac{(an)^j}{j!}\right)\alpha_n(bn)\exp\left\{\frac{-2b}{\lfloor c_n \rfloor^2}\right\} + O(\frac{1}{\sqrt{n}}).$$

By the Lemma A.2, we obtain

$$\varliminf_{n\to\infty} D_{TV}^c(x_{2n}, an) \geq \tfrac{1}{2}e^{-(3c^2 + 2/c^2)b} > 0, \quad \forall\, b > \max\{c^2, 1\}.$$

By Corollary 2.1, no subfamily of $\mathcal{F}_c$ has total variation precutoff if $c > 0$.

When the state space is $\{0, 1, \ldots, 2n+1\}$, we choose the proper testing set $A_{2n+1} = \{0, 1, \ldots, n+1\}$ and do the same argument again. There is no total variation precutoff.

Part 2. c=0

Recall the Lemma 3.1, we have

$$\frac{H_{n,t}(x,y)}{\pi_n(y)} - 1 = \sum_{i=1}^{n} e^{-t(1-\beta_{n,i})}\psi_{n,i}(x)\psi_{n,i}(y)$$

We set $\frac{H_{n,t}(x_n,y)}{\pi_n(y)} - 1 = f_n(t,y) + g_n(t,y)$, where

$$f_n(t,y) = e^{-t(1-\beta_{n,2})}\psi_{n,2}(x_n)\psi_{n,2}(y) \quad \text{and} \quad g_n(t,y) = \sum_{i=1, i\neq 2}^{n} e^{-t(1-\beta_{n,i})}\psi_{n,i}(x)\psi_{n,i}(y)$$

By Proposition 2.1, $2D^c_{TV}(x,t) = \|\frac{H_{n,t}(x,\cdot)}{\pi_n(\cdot)} - 1\|_1 = \|f_n(t,\cdot) + g_n(t,\cdot)\|_1$.

Observe that

$$\begin{aligned} \|f_n(t,\cdot) + g_n(t,\cdot)\|_1 &= \sum_{k=0}^{n} |f_n(t,k) + g_n(t,k)|\pi_n(k) \\ &\geq \sum_{k=0}^{n} |f_n(t,k)|\pi_n(k) - \sum_{k=0}^{n} |g_n(t,k)|\pi_n(k) \\ &\geq \|f_n(t,\cdot)\|_1 - \|g_n(t,\cdot)\|_2, \end{aligned}$$

where the first inequality is given by the triangle inequality and the second inequality is given by Proposition 2.1.

Therefore, we suffice to show that, for all $c > 0$,

$$\varliminf_{n\to\infty} \left(\|f_n(t,\cdot)\|_1 - \|g_n(t,\cdot)\|_2\right) > 0.$$

Then

$$\begin{aligned} \|g_n(t,\cdot)\|_2^2 &= \sum_{k=0}^{n} \left| \sum_{i=1, i\neq 2}^{n} e^{-t(1-\beta_{n,i})}\psi_{n,i}(x)\psi_{n,i}(k) \right|^2 \pi_n(k) \\ &= \sum_{k=0}^{n} \left| (\psi_{n,1}(x_n))^2 e^{-4t/n} + \sum_{i=3}^{n} (\psi_{n,i}(x_n))^2 e^{-4ti/n} \right| \pi_n(k) \\ &= \left| \left(\frac{n-2x_n}{\sqrt{n}}\right)^2 e^{-4t/n} + \sum_{i=3}^{n} (\psi_{n,i}(x_n))^2 e^{-4ti/n} \right|. \end{aligned}$$

By the argument in section 4.1, let $\gamma = \sup_n\{|n-2x_n|/\sqrt{n}\} \vee 1 < \infty$, then we have $|\psi_{n,i}(x_n)| \leq (\gamma+1)^i$ for all $0 \leq i \leq n$.

$$\begin{aligned}
\|g_n(t,\cdot)\|_2 &\leq \left|(\frac{n-2x_n}{\sqrt{n}})^2 e^{-4t/n} + \sum_{i=3}^{n} (\gamma+1)^{2i} e^{-4ti/n}\right|^{1/2} \\
&\leq \left|(\frac{n-2x_n}{\sqrt{n}})^2 e^{-4t/n} + \sum_{i=3}^{\infty} (\gamma+1)^{2i} e^{-4ti/n}\right|^{1/2} \\
&\leq \left|(\frac{n-2x_n}{\sqrt{n}})^2 e^{-4t/n} + \frac{[(\gamma+1)^2 e^{-4t/n}]^3}{1-(\gamma+1)^2 e^{-4t/n}}\right|^{1/2}
\end{aligned}$$

Then

$$\|g_n(cn,\cdot)\|_2 \leq \left|(\frac{n-2x_n}{\sqrt{n}})^2 e^{-4c} + \frac{[(\gamma+1)^2 e^{-4c}]^3}{1-(\gamma+1)^2 e^{-4c}}\right|^{1/2}$$

if $(\gamma+1)^2 < e^4 c$. Therefore, for $(\gamma+1)^2 < e^4 c$,

$$\varlimsup_{n\to\infty} \|g_n(cn,\cdot)\|_2 = \left|\frac{(\gamma+1)^3 e^{-12c}}{1-(\gamma+1)^2 e^{-4c}}\right|^{1/2}. \tag{4.5}$$

Note that $|\psi_{n,2}(x_n)| = |\sqrt{\frac{1}{2(1-1/n)}}[(\frac{n-2x_n}{\sqrt{n}})^2 - 1]|$, and since $\frac{n-2x_n}{\sqrt{n}} \to 0$, we can choose $N > 0$ s.t. for all $n \geq N$

$$|\psi_{n,2}(x_n)| \geq \sqrt{\frac{1}{2}}(1-\frac{1}{n}).$$

Then, for $n \geq N$,

$$\begin{aligned}
\|f_n(t,\cdot)\|_1 &\geq \frac{1}{\sqrt{2}}(1-\frac{1}{n})e^{-4t/n}\left(\sum_{k=0}^{n} |\psi_{n,2}(k)|\pi_n(k)\right) \\
&\geq \frac{1}{\sqrt{2}}(1-\frac{1}{n})e^{-4t/n}\left\{\sum_{k=0}^{n} |\frac{1}{\sqrt{2}}[(\frac{n-2k}{\sqrt{n}})^2 - 1]|\pi_n(k)\right\} \\
&\geq \frac{1}{\sqrt{2}}(1-\frac{1}{n})e^{-4t/n}\left\{\sum_{\frac{n}{2}-\frac{\sqrt{n}}{4}<k<\frac{n}{2}+\frac{\sqrt{n}}{4}}^{n} |\frac{1}{\sqrt{2}}[(\frac{n-2k}{\sqrt{n}})^2 - 1]|\pi_n(k)\right\} \\
&\geq \frac{3}{8}(1-\frac{1}{n})e^{-4t/n}\pi_n(\{k : |\frac{n}{2}-k| < \frac{\sqrt{n}}{4}\})
\end{aligned}$$

By the central limit theorem,

$$\lim_{n\to\infty} \pi_n(|\frac{Y_n - n/2}{\sqrt{n}}| < \frac{1}{4}) = \lim_{n\to\infty} \pi_n(|\frac{Y_n - n/2}{\sqrt{n}/2}| < \frac{1}{2}) = \frac{2}{\sqrt{2\pi}}\int_0^{\frac{1}{2}} e^{-u^2/2}du > 1/3.$$

Hence

$$\begin{aligned}
\varliminf_{n\to\infty} (\|f_n(cn,\cdot)\|_1 - \|g_n(cn,\cdot)\|_2) &\geq \frac{1}{8}e^{-4c} - \frac{(\gamma+1)^3/2e^{-6c}}{\sqrt{1-(\gamma+1)^2 e^{-4c}}} \\
&= e^{-4c}\left(\frac{1}{8} - \frac{(\gamma+1)^3/2e^{-2c}}{\sqrt{1-(\gamma+1)^2 e^{-4c}}}\right) > 0
\end{aligned}$$

if $c > \frac{1}{4}\ln\{(\gamma+1)[64(\gamma+1)^2+1]\}$. By the total variation is non-increasing, we have

$$\varlimsup_{n\to\infty} D^c_{TV}(x_n, cn) > 0, \quad \forall\, c > 0.$$

By Corollary 2.1, no subfamily has a total variation precutoff if $c = 0$.

A Appendix

The following proposition is the Proposition 4.1 in section 4.

Proposition A.1. *Let K_n is the transition matrix of the Ehrenfest chain on $\mathcal{S}_n$. Suppose $A = \{0, 1, \ldots, \lceil n/2 \rceil\}$. Then $K_n^m(x, A) \geq 1/2$ for all $x \in A$, $m \geq 0$.*

Proof. Let $\mathcal{S}_n = \{0, 1, \ldots, 2n\}$. Obviously, A is more elements than A^c. By the symmetry of the Ehrenfest chain, $K_n^m(x, A) \geq \frac{1}{2}$ for all $x \in A$, $m \geq 0$. Consider the state space is $\{0, 1, , \ldots, 2n+1\}$. Note $K_n(i, i+1) \geq K_n(i, i-1)$ if $i \leq n$, and $K_n(i, i+1) \leq K_n(i, i-1)$ if $i > n$, and $K_n(n, n+1) = K_n(n+1, n) = (n+1)/(2n+1)$. That implies $K_n^m(x, A) \geq K_n^m(n+1, A)$ for all $x \in A$, $m \geq 0$. We show that $K_n^{2m}(n, n-2i) \geq K_n^{2m}(n, n+2i+2) \geq K_n^{2m}(n, n-2i-2)$ for all $i \geq 0$, $m > 0$.

Let $m = 1$. $K^2(n, n) \geq K^2(n, n+2) \geq K^2(n, n-2)$.

Let $m = w$, for all $i \geq 0$,

$$K^{2w}(n, n-2i) \geq K^{2w}(n, n+2i+2) \geq K^{2w}(n, n-2i-2).$$

Case 1: show $K_n^{2(w+1)}(n, n) \geq K_n^{2(w+1)}(n, n+2)$.

Expand $K_n^{2(w+1)}(n, n) = \sum_i K^{2w}(n, i)K^2(i, n)$ and $K_n^{2(w+1)}(n, n+2) = \sum_i K^{2w}(n, i)K^2(i, n+2)$. Since we know

a $K^2(n, n) - K^2(n, n+2) = \frac{n^2+3n+1}{(2n+1)^2}$

b $K^2(n+2, n) - K^2(n+2, n+2) = \frac{-n^2-n+5}{(2n+1)^2}$

c $K^2(n-2, n) - K^2(n+4, n+2) = \frac{-2n-6}{(2n+1)^2}$

And a+b+c=0. Comparison entries of $\sum_i K^{2w}(n, i)K^2(i, n)$ and $\sum_i K^{2w}(n, i)K^2(i, n+2)$, then we obtain $K_n^{2(w+1)}(n, n) \geq K_n^{2(w+1)}(n, n+2)$.

Case 2: show $K^{2(w+1)}(n,n+2i) \geq K^{2(w+1)}(n,n-2i) \geq K^{2(w+1)}(n,n+2i+2)$, $\forall i \geq 1$. Observe that, for $i \geq 1$,

$$\begin{aligned} K^{2(w+1)}(n,n+2i) &= K^{2w}(n,n+2(i-1))K^2(n+2(i-1),n+2i) \\ &+ K^{2w}(n,n+2i)K^2(n+2i,n+2i) \\ &+ K^{2w}(n,n+2(i+1))K^2(n+2(i+1),n+2i), \end{aligned}$$

$$\begin{aligned} K^{2(w+1)}(n,n-2i) &= K^{2w}(n,n-2(i-1))K^2(n-2(i-1),n-2i) \\ &+ K^{2w}(n,n-2i)K^2(n-2i,n-2i) \\ &+ K^{2w}(n,n-2(i+1))K^2(n-2(i+1),n-2i), \end{aligned}$$

$$\begin{aligned} K^{2(w+1)}(n,n+2i+2) &= K^{2w}(n,n+2i)K^2(n+2i,n+2i+2) \\ &+ K^{2w}(n,n+2i+2)K^2(n+2i+2,n+2i+2) \\ &+ K^{2w}(n,n+2i+4)K^2(n+2i+4,n+2i+2). \end{aligned}$$

We claim that suppose $a+b+c = d+e+f$, where $a \geq d > 0$, $b \geq e > 0$, $f \geq c > 0$ and $A \geq D \geq B \geq E \geq C \geq F > 0$, then $Aa+Bb+Cc \geq Dd+Ee+Ff$. It's easy to prove if let $a-d = \varepsilon_1 > 0$ and $b-e = \varepsilon_2 > 0$. Thus, it implies $K^{2(w+1)}(n,n+2i) \geq K^{2(w+1)}(n,n-2i) \geq K^{2(w+1)}(n,n+2i+2)$, $\forall i \geq 1$. By an induction, $K_n^{2m}(n,n-2i) \geq K_n^{2m}(n,n+2i+2) \geq K_n^{2m}(n,n-2i-2)$, $\forall i \geq 0$, $m > 0$. Then $K_n^{2m+1}(n+1,n-2i) \geq K_n^{2m+1}(n+1,n+2i+2) \geq K_n^{2m+1}(n+1,n-2i-2)$, $\forall i \geq 0$, $m > 0$. Since the symmetry of Ehrenfest chains, $K^{2m}(n+1,n+1-2i) \geq K^{2m}(n+1,n+3+2i) \geq K^{2m}(n+1,n-1-2i)$, $\forall i \geq 0$, $m > 0$. □

For a discrete time simple random walk $(X_n)_{n=0}^{\infty}$ on $\mathbb{Z}$, the first passage time to $\{\pm m\}$ is defined by

$$T_m = \inf\{n \geq 0 | X_n = m \text{ or } X_n = -m\}$$

For a continuous time simple random walk on $\mathbb{Z}$, let $N(t)$ be a Poisson process with parameter 1 and independent of X_n and set $Y_t = X_{N(t)}$. Thus, the first passage time to $\{\pm m\}$ is defined by

$$\widetilde{T}_m = \inf\{t \geq 0 | Y_t = m \text{ or } Y_t = -m\}.$$

Lemma A.1. *[4, Theorem 3.1] Let T_m, $\widetilde{T}_m$ be defined above and P_0 be the conditional probability given the initial state is 0. Then, for any $b > 1$ and $m \geq 5$,*

$$\min\{P_0(T_m > bm^2), P_0(\widetilde{T}_m > bm^2)\} \geq e^{-2b}.$$

Lemma A.2. *[1, Lemma A.1] For* $n > 0$*, let* $a_n \in \mathbb{R}^+$*,* $b_n \in \mathbb{Z}^+$*,* $c_n = \frac{b_n - a_n}{\sqrt{a_n}}$ *and* $d_n = e^{-a_n} \sum_{i=0}^{b_n} \frac{{a_n}^i}{i!}$*. Assume that* $a_n + b_n \to \infty$*. Then*

$$\overline{\lim_{n\to\infty}}\, d_n = \Phi\left(\overline{\lim_{n\to\infty}}\, c_n\right), \quad \underline{\lim_{n\to\infty}}\, d_n = \Phi\left(\underline{\lim_{n\to\infty}}\, c_n\right),$$

where $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt$*.*

In particular, if c_n *converges (the limit can be* $+\infty$ *and* $-\infty$*), then* $\lim_{n\to\infty} d_n = \Phi\left(\lim_{n\to\infty} c_n\right)$*.*

References

[1] Guan-Yu Chen and L. Saloff-Coste. The cutoff phenomenon for ergodic Markov processes. Electron. J. Probab., 13 (2008) 26-78.

[2] Guan-Yu Chen and L. Saloff-Coste. The L^2-cutoff for reversible Markov processes. Journal of Functional Analysis, 258 (2010) 2246-2315.

[3] Laurent Saloff-Coste. Lectures on finite Markov chains.In *Lectures on probability theory and statistics*, volume 1665 of *Lecture Notes in Mathematics*(1996) 301-413.

[4] Guan-Yu Chen, Yang-Jen Fan and Yuan-Chung Sheu. The cutoff phenomenon for Ehrenfest processes. Submitted.

[5] R. Koekoek and R. Swarttouw. The askey-scheme of hypergeometric orthogonal polynomials and its q-analog. http://math.nist.gov/opsf/projects/koekoek.html, 1998.