Inequalities for Rare Events in Time-Reversible Markov Chains II

David J Aldous [*]
Department of Statistics
University of California
Berkeley CA 94720

Mark Brown [†]
Department of Mathematics
City College, CUNY
New York NY 10031

Abstract

A previous paper discussed explicit bounds in the exponential approximation for the distribution of the waiting time until a stationary reversible Markov chain first enters a "rare" subset of states. In this paper Stein's method is used to get explicit (but complicated) bounds on the Poisson approximation for the number of non-adjacent visits to a rare subset.

Stochastic Processes and their Applications **44** (1993) 15-25.

[*]Research supported by N.S.F. Grant MCS90-01710
[†]Research supported by A.F.O.S.R. Grant 89-0083

1 Introduction

The Poisson approximation for numbers of rare events which actually occur, and the exponential approximation for the waiting time until first occurrence of a rare event, are useful throughout many areas on probability – one view of this big picture is presented in Aldous [1]. This paper and [2] are a theoretical study of explicit bounds in these approximations, in the special setting of hitting times of stationary reversible Markov chains.

The following set-up and notation will be used throughout. $(X_t; t \geq 0)$ is an irreducible finite-state reversible Markov chain in continuous time. Let π be the stationary distribution. We work throughout with the *stationary* chain, occasionally writing E_π and P_π to remind ourselves. The matrix of transition rates has real eigenvalues $0 = \lambda_0 > -\lambda_1 \geq -\lambda_2 \ldots$. Call $\tau = 1/\lambda_1$ the *relaxation time* of the chain. Let A be a fixed (proper, non-empty) subset of the state space, and let $T \geq 0$ be the first hitting time on A. So $0 < E_\pi T < \infty$.

The vague general idea "if $\pi(A)$ is small then T has approximately exponential distribution" has been well understood for a long time. Here is one clean formulation of that idea in the present setting.

Theorem 1

$$|P_\pi(T/E_\pi T > t) - e^{-t}| \leq \tau/E_\pi T \quad \textit{for all } t > 0.$$

In the special case where A is a singleton this was proved by Brown [8]. That proof exploits the completely monotone property of the distribution of T, under the stationary initial distribution. In [2] it was shown that the general case could be reduced to the special case. Quite different techniques seem necessary to prove a Poisson limit theorem. Let us first say what such a theorem should look like.

For the stationary chain, excursions inside A alternate with excursions outside A. For $L > 0$ let a *L-vacation* be an excursion outside A of duration $\geq L$. Call the intervals between L-vacations *L-visits*. So a L-visit comprises one or more excursions inside A, with the intervening excursions outside A each having duration less than L. Define

$$\begin{aligned} N_L(t) &= \text{number of starts of } L\text{-visits during } [0,t] \\ &= \text{number of ends of } L\text{-vacations during } [0,t]. \end{aligned} \tag{1}$$

It is intuitively clear that, for $\tau \ll L \ll E_\pi T$, the times of starts of L-visits should form approximately a Poisson process of rate $1/E_\pi T$. The object of this paper is to use an implementation of Stein's method due to Barbour and Holst to get an explicit (but complicated) error bound for the Poisson approximation $N_L(t) \approx \text{Poisson}(t/ET)$ at fixed times. These bounds are stated in Theorem 9 and the subsequent Corollaries. Unfortunately these bounds are much less "clean" than Theorem 1. Roughly, for fixed t/ET the error is order $\frac{\tau}{ET} \log(\frac{ET}{\tau})$. In view of Theorem 1 it is natural to believe that the *log* term in unnecessary.

Our reason for doing this study is methodological. Over the last five years it has become clear that Stein's method for approximating distributions of dependent sums is useful when the summands (loosely speaking) have either "much combinatorial symmetry" or "essentially finite-range dependence". See Arratia et al [3, 4] and Barbour et al [6, 7] for Poisson approximations involving such sums. It is not yet clear whether the method has substantially wider scope. Bluntly, this paper indicates how far one can get in this particular problem without having a new idea about implementing Stein's method. We regard the problem posed here as a natural "test case" for new implementations of the method.

Let us emphasize that our results are "absolute" inequalities, i.e. do not involve any unspecified constants depending on the Markov chain under consideration. There are two reasons for this emphasis. First, without going into details we assert that our results extend unchanged to *continuous*-space stationary reversible processes (under weak regularity conditions – say strong Markov and cadlag paths). For such processes we may have $\tau = \infty$, but the results are only interesting when $\tau < \infty$. One explanation is that the proofs extend, unchanged except for terminology. Another explanation is that, given a continuous-space process, one may express it as a weak limit of finite-state processes in such a way that relaxation times, first hitting times and the other parameters of interest converge.

The second reason for our emphasis is that, if explicit bounds are not required, then exponential, Poisson and compound Poisson limit theorems can be derived in very great generality by "soft" arguments based either upon regeneration (e.g. for Harris positive-recurrent chains) or upon characterizations of the limit process via independent increments (e.g. for non-Markovian stationary processes under mixing hypotheses). So there is no point in proving such results by *ad hoc* arguments depending on special structure. Conversely, Smith [9] observed that certain "mixing hypothesis" results could be quantified via Stein's method, giving bounds in terms of the mixing coefficients. This seems unsatisfactory as a general procedure, since to estimate such coefficients one needs to use the structure of the process, and it seems preferable to give bounds directly in terms of natural parameters of the process.

2 Tools

Recall that we work with the *stationary* chain (X_t). It is convenient to standardize time by setting

$$ET = 1$$

which of course involves no loss of generality for theoretical results. In practice, estimating ET well can be hard: see [2] for some bounds. Note that. after standardizing time, we are interested in the case $\tau \ll 1$.

Our argument has four ingredients of rather different flavors, which we lay out in this section. We shall need a technical result from [2] Lemma 13.

Lemma 2 *The density function $f(t)$ of T is completely monotone and satisfies*

(a) $f(t) \leq 1 + \tau/(2t),\ t > 0.$

(b) $f(t) \geq 1 - 2\tau - t,\ t > 0.$

(c) $0 \leq f(t_1) - f(t_2) \leq 2(t_2 - t_1)$ *for* $t_2 \geq t_1 \geq \tau(5 + 2\log(1/\tau))$*, provided* $\tau \leq 1$.

Associate with the stationary chain a marked point process, defined as follows. For each excursion (t_1, t_2) of the chain outside A, put a "point" at t_1 and mark it with the duration $l = t_2 - t_1$ of the excursion. This point process has rate $r(l)$ defined by

$$P(\text{ excursion of length } \in (l, l + dl) \text{ starts during } (s, s + ds)) = r(l)\ dl\ ds.$$

The next lemma relates $r(l)$ to the density $f(t)$ of T for the stationary chain.

Lemma 3 *For* $L > 0$

$$f(L) = \int_L^\infty r(l) dl = \textit{ rate of starts of } L\textit{-vacations } .$$

Proof. This is a standard renewal-theory relationship. Fix L. Consider the asymptotic proportion of time s such that

$$X_{s+u} \in A^c \text{ for all } 0 \leq u \leq L, \quad X_{s+L+dL} \in A. \tag{2}$$

Each L-vacation contains an interval of length dL of such s, and so the asymptotic proportion equals

$$(\text{ rate of starts of } L\text{-vacations }) \times (dL).$$

But computing the stationary probability of the event (2) gives $f(L)dL$.

Recall that our goal is to study $N_L(t)$ defined at (1). Lemma 3 implies

$$EN_L(t) = tf(L). \tag{3}$$

Because $N_L(t)$ is decreasing in L, this enables us to bound the effect of varying L, as follows:

$$E|N_{L_1}(t) - N_{L_2}(t)| \leq t(f(L_1) - f(L_2));\; L_1 < L_2. \tag{4}$$

So in the sequel it is enough to get a bound for *some* L in the range $\tau \ll L \ll 1$, and then (4) and Lemma 2(c) imply a bound for all L in that range.

Another ingredient is the following well-known "maximal correlation" property of stationary reversible Markov chains.

Lemma 4

$$\begin{aligned}\exp(-t/\tau) &= \max\{\mathrm{cor}(Z_1, Z_2) : Z_1 \in \mathcal{F}(X_s, s \leq 0), Z_2 \in \mathcal{F}(X_s, s \geq t)\} \\ &= \max_{h,g} \mathrm{cor}(h(X_0), g(X_t)).\end{aligned}$$

The final ingredient is the following implementation of Stein's method given by Barbour and Holst [6] Theorem 2.1. Write $||\cdot||_{TV}$ for total variation distance between distributions:

$$||\mu - \mu^*||_{TV} = \sup_D |\mu(D) - \mu^*(D)|.$$

Proposition 5 *Let $(B_i; 1 \leq i \leq k)$ be events. Let $N = \sum_{i=1}^k 1_{B_i}$ count how many of these events occur. Let $\lambda = EN = \sum_i P(B_i)$. For each i let N_i have the conditional distribution of $N-1$ given B_i, and for each i let $(N^{(i)}, N_i)$ be some joint distribution with $N^{(i)} \stackrel{d}{=} N$. Then*

$$||\mathrm{dist}(N) - \mathrm{Poisson}(\lambda)||_{TV} \leq \min(1, \lambda^{-1}) \sum_i P(B_i) E|N^{(i)} - N_i|.$$

3 Proofs

We start with a coupling construction.

Lemma 6 *Consider a triple (X,B,Y), where X and Y are non-negative integer-valued r.v.'s, and B is an event. Suppose*
(a1) The maximal correlation between $\mathcal{F}(X)$ and $\mathcal{F}(B,Y)$ is at most ρ.
(a2) The maximal correlation between $\mathcal{F}(X,B)$ and $\mathcal{F}(Y)$ is at most ρ.
(b) $EX^2 \le \sigma^2$ and $EY^2 \le \sigma^2$.

Define (X^*,Y^*) to have the conditional distribution of (X,Y) given B. Then there exists a joint distribution for $((X,Y),(X^*,Y^*))$ such that

$$E|(X+Y)-(X^*+Y^*)| \le 12\max(\sigma^{3/2}P^{-5/8}(B)\rho^{1/4}, \sigma P^{-3/4}(B)\rho^{1/2}).$$

Proof. If events F,G have correlation c then

$$P(F\cap G)-P(F)P(G) = c\sqrt{P(F)P(F^c)P(G)P(G^c)}$$

and so

$$|P(F\cap G)-P(F)P(G)| \le |c/2|\sqrt{P(F)}. \tag{5}$$

Applying (5) to $F=\{X=i,B\}$ and $G=\{Y=j\}$ gives, using (a2),

$$|P(X=i,B,Y=j)-P(X=i,B)P(Y=j)| \le (\rho/2)\sqrt{P(B)}.$$

Applying (5) to $F=B$ and $G=\{X=i\}$ gives, using (a1),

$$|P(X=i,B)-P(X=i)P(B)| \le (\rho/2)\sqrt{P(B)}.$$

Combining these,

$$|P(X=i,B,Y=j)-P(X=i)P(B)P(Y=j)| \le \rho\sqrt{P(B)}.$$

In other words

$$|P(X^*=i,Y^*=j)-P(X=i)P(Y=j)| \le \rho P^{-1/2}(B).$$

Let $K\ge 1$ be an integer. Then

$$\sum_{i=0}^{K-1}\sum_{j=0}^{K-1}|P(X^*=i,Y^*=j)-P(X=i)P(Y=j)| \le K^2\rho P^{-1/2}(B).$$

The same argument gives a better bound for $P(X=i, Y=j)$. Combining,

$$\sum_{i=0}^{K-1}\sum_{j=0}^{K-1} |P(X^*=i, Y^*=j) - P(X=i, Y=j)| \leq 2K^2 \rho P^{-1/2}(B). \quad (6)$$

Now apply the general inequality

$$||\mu - \mu^*||_{TV} \leq \sum_{i \in C} |\mu(i) - \mu^*(i)| \ + \mu(C^c)$$

to $C = \{X \geq K \text{ or } Y \geq K\}$, to get

$$||\text{dist}(X^*, Y^*) - \text{dist}(X, Y)||_{TV} \leq 2K^2 \rho P^{-1/2}(B) + P(X \geq K \text{ or } Y \geq K)$$

$$\leq 2\rho P^{-1/2}(B) K^2 + 2\sigma^2 K^{-2} \equiv P(\Omega_K), \text{ say.} \quad (7)$$

Now given two distributions whose total variation distance is d, we can construct random variables Z and Z^* with those distributions such that $Z = Z^*$ outside an event Ω of probability d. So we can construct a joint distribution such that $(X, Y) = (X^*, Y^*)$ outside an event Ω_K with probability (7). Then

$$E|(X+Y) - (X^* + Y^*)| \leq EX1_{\Omega_K} + EX^*1_{\Omega_K} + EY1_{\Omega_K} + EY^*1_{\Omega_K}.$$

But

$$EX^*1_{\Omega_K} \leq \sqrt{EX^{*2}P(\Omega_K)} \leq \sqrt{\frac{EX^2}{P(B)}P(\Omega_K)} = \sigma\sqrt{\frac{P(\Omega_K}{P(B)}}.$$

The same bound holds for Y^*, and this bound (without the $P(B)$ term) holds for X and Y. So

$$E|(X+Y) - (X^* + Y^*)| \leq 4\sigma\sqrt{\frac{P(\Omega_K}{P(B)}}. \quad (8)$$

The final step is to choose an integer K to minimize $P(\Omega_K)$. By considering $K = \max(\lfloor a^{1/4} \rfloor, 1)$ one can verify that for $a > 0$

$$\min_{K \geq 1} K^2 + aK^{-2} \leq \frac{9}{2}\max(a^{1/2}, 1).$$

Then by scaling, for $a, b > 0$

$$\min_{K \geq 1} bK^2 + aK^{-2} \leq \frac{9}{2}\max((ba)^{1/2}, b).$$

Applying this to (7) and (8) leads to the bound stated in the Lemma.

Here is the set-up where we shall apply Proposition 5. Fix L and t_0, with t_0/L an integer. Divide the time interval $[0, t_0]$ into successive intervals of length L. Let B_i be the event "some L-vacation ends during $[(i-1)L, iL]$". Note that at most one such ending can occur during $[(i-1)L, iL]$. Thus $N_L(t_0)$ equals the number of events $(B_i : 1 \leq i \leq t_0/L)$ which occur. We shall use Proposition 5 to approximate $N_L(t_0)$ by the Poisson distribution. Note that from Lemma 3

$$P(B_i) = Lf(L). \tag{9}$$

Note also that the occurrence of B_i depends only on the behavior of the chain during time $[(i-2)L, iL]$.

Lemma 7

$$\begin{aligned} P(B_1 \cap B_j) &\leq L(f(L) - f(3L)), \ j = 2, 3 && (10) \\ &\leq P^2(B_1) + e^{-(j-3)L/\tau} P(B_1), \ j \geq 4 && (11) \end{aligned}$$

Proof. For $j \geq 4$ we have $B_j \in \mathcal{F}(X_s : s \geq (j-2)L)$ and $B_1 \in \mathcal{F}(X_s : s \leq L)$, and so Lemma 4 implies $\mathrm{cor}(B_1, B_j) \leq e^{-(j-3)/\tau}$. This establishes (11). For (10), consider the event D defined by:

there is some L-vacation containing $2L$ and contained in $(0, 3L)$.

Clearly $B_3 \cap (B_1 \cup B_2) \subseteq D$, and so

$$\max(P(B_1 \cap B_2), P(B_1 \cap B_3)) = \max(P(B_1 \cap B_3), P(B_2 \cap B_3)) \leq P(D).$$

By Lemma 3 the chance that an excursion outside A starts during $(s, s+ds)$ and ends during $(t, t+dt)$ is $-f'(t-s)dsdt$. So

$$\begin{aligned} P(D) &= \int\int_{0 \leq s \leq 2L \leq t \leq 3L; t-s > L} -f'(t-s)dtds \\ &= \int_0^L (f(2L-s) - f(3L-s))ds + \int_L^{2L} (f(L) - f(3L-s))ds \\ &= Lf(L) - P(2L < T < 3L). \end{aligned}$$

Because f is decreasing, $P(2L < T < 3L) \geq Lf(3L)$, establishing (10).

Lemma 8

$$EN_L^2(t_0) \leq \frac{1 + e^{-L/\tau}}{1 - e^{-L/\tau}} t_0 f(L) + t_0^2 f^2(L) + 4t_0(f(L) - f(3L)) \equiv \sigma^2 \ \textit{say}.$$

Proof. Set $K = t_0/L$. Then

$$EN_L^2(t_0) = Kr_0 + 2\sum_{j=1}^{K-1}(K-j)r_j \tag{12}$$

where $r_j = P(B_1 \cap B_{j+1})$. So $r_0 = Lf(L)$, implying $Kr_0 = t_0 f(L)$. By (10)

$$Kr_1 + Kr_2 \leq 2KL(f(L) - f(3L)) = 2t_0(f(L) - f(3L)).$$

By (11), for $j \geq 3$

$$r_j \leq (Lf(L))^2 + Lf(L)s^{j-2}, \quad \text{where } s = e^{-L/\tau}.$$

So

$$\sum_{j=3}^{K-1}(K-j)r_j \leq \frac{1}{2}K^2(Lf(L))^2 + KLf(L)\sum_{i=1}^{\infty} s^i = \frac{1}{2}t_0^2 f^2(L) + t_0 f(L)\frac{s}{1-s}.$$

Combining these bounds establishes the lemma.

We can now give the Poisson approximation result. Here is the raw version.

Theorem 9 *Let $ET = 1$. Let t_0/L be an integer. Then*

$$||\mathrm{dist}(N_L(t_0)) - \mathrm{Poisson}(t_0 f(L))||_{TV} \leq (\Delta_1 + \Delta_2)\ \min(1, t_0 f(L))$$

where Δ_1, Δ_2 are specified by (13,14) below.

Proof. Fix i. Let $N^{*(i)}$ and N_i^* be as in Proposition 5, but only counting events B_j with $|j-i| \geq 3$. We want to apply Lemma 6 to the events $B_i =$ "some L-vacation ends during $[(i-1)L, iL]$" and to the r.v.'s

$$X = \sum_{j=1}^{i-3} 1_{B_j}, \quad Y = \sum_{j=i+3}^{t_0/L} 1_{B_j}.$$

The hypotheses of Lemma 6 holds with $\rho = e^{-L/\tau}$, for σ as in Lemma 8. The conclusion of Lemma 6 is that we can arrange a joint distribution of $N^{*(i)}$ and N_i^* such that

$$E|N^{*(i)} - N_i^*| \leq 12\max(\sigma^{3/2}(Lf(L))^{-5/8}e^{-L/4\tau}, \sigma(Lf(L))^{-3/4}e^{-L/2\tau})$$
$$= \Delta_1 \text{ say.} \tag{13}$$

Taking into account the extra j's, we can write

$$E|N^{(i)} - N_i| \leq \Delta_1 + \Delta_2^*$$

where

$$\begin{aligned} \Delta_2^* &= 5P(B_1) + 2P(B_2|B_1) + 2P(B_3|B_1) \\ &\leq 5Lf(L) + 4\frac{f(L) - f(3L)}{f(L)} \equiv \Delta_2 \text{ say, by Lemma 7.} \end{aligned} \tag{14}$$

The Theorem now follows from the conclusion of Proposition 5.

Remarks. Theorem 9 gives a bound involving t_0, τ, L and $f(L), f(3L)$. But we can use Lemma 2 to bound $f(L), f(3L)$ in terms of the other quantities, and thus obtain a bound on the Poisson approximation involving only t_0, τ, L. This is done in the Corollaries below – although the results could be presented in many different ways.

Corollary 10 *Let $ET = 1$. Suppose $\tau \leq \frac{1}{1000}$, and define $L_0 = \frac{13}{2}\tau \log \frac{1}{\tau}$. Then for all $t > 0$*

$$||\text{dist}(N_{L_0}(t)) - \text{Poisson}(tf(L_0))||_{TV} \leq (25 + \frac{t + t^2}{17})L_0.$$

Remarks. For $\tau > \frac{1}{1000}$ the bound is greater than 1 and so the result is vacuous. Also, we could replace the given Poisson distribution by the Poisson distribution with mean t, by adding an extra term $t|1 - f(L_0)|$ to the bound, but then bounding this extra term via Lemma 2 would create an (asymptotically) much larger bound than that stated.

Proof. We want to apply Theorem 9 to the largest integer multiple t_0 of L_0 with $t_0 \leq t$. Arithmetic gives

$$L_0 \leq 0.04491, \quad \frac{\tau}{L_0} \leq \frac{1}{44} \tag{15}$$

and then from Lemma 2 (a,b)

$$0.953 \leq f(L_0) \leq \frac{89}{88}. \tag{16}$$

Lemma 2 (c) gives

$$f(L_0) - f(3L_0) \leq 4L_0 \tag{17}$$

and then

$$\begin{aligned}\Delta_2 &\leq 5L_0 f(L_0) + 16L_0/f(L_0) \text{ by (14)} \\ &\leq 22L_0 \text{ using (16).}\end{aligned}$$

Next, from (13) and (16)

$$\Delta_1 \leq 13 \max(\sigma^{3/2} L_0^{-5/8} \tau^{13/8}, \sigma L_0^{-3/4} \tau^{13/4}).$$

For $\sigma \geq 1$ the first term is larger, so

$$\Delta_1 \leq 13 L_0^{-5/8} \tau^{13/8} \max(1, \sigma^{3/2}).$$

Using (15) to bound $(\tau/L_0)^{5/8}$ we get

$$\Delta_1 \leq \frac{5\tau}{4} \max(1, \sigma^{3/2}) \leq \frac{5\tau}{4} \max(1, \sigma^2).$$

Because $\sigma^2 \geq t_0 f(L_0)$ we have by Theorem 9

$$\begin{aligned}||\text{dist}(N_{L_0}(t_0) - \text{Poisson}(t_0 f(L_0))||_{TV} &\leq (\Delta_1 + \Delta_2) \min(1, \sigma^2) \\ &\leq 22L_0 + \frac{5\tau}{4} \max(1, \sigma^2) \min(1, \sigma^2) \\ &\leq 22L_0 + \frac{L_0}{34} \sigma^2 \text{ by (15).} \qquad (18)\end{aligned}$$

The definition of σ^2 in Lemma 8 and (15,17) give a crude bound

$$\sigma^2 \leq 2(t_0 + t_0^2).$$

Finally, for integer-valued r.v.'s

$$||\text{dist}(X) - \text{dist}(Y)||_{TV} \leq P(X \neq Y) \leq E|X - Y|.$$

Because $0 \leq t - t_0 \leq L_0$ we have using (4)

$$||\text{dist}(N_{L_0}(t)) - \text{dist}(N_{L_0}(t_0))||_{TV} \leq L_0 f(L_0) \qquad (19)$$

$$||\text{Poisson}(t f(L_0)) - \text{Poisson}(t_0 f(L_0))||_{TV} \leq L_0 f(L_0). \qquad (20)$$

Combining (18,19,20) and using the crude bound $2L_0 f(L_0) \leq 3L_0$, we establish the Corollary.

To get a similar result for general L, one could repeat the argument above with L in place of L_0. But as mentioned earlier, it seems simpler to apply (4) and Lemma 2, which leads to the following result.

Corollary 11 *Let* $ET = 1$*. Suppose* $\tau \leq \frac{1}{1000}$*, and define* $L_0 = \frac{13}{2}\tau \log \frac{1}{\tau}$*. Then for all* $t > 0$ *and all* $L \geq \tau(5 + 2\log(1/\tau))$*,*

$$||\mathrm{dist}(N_L(t)) - \mathrm{Poisson}(tf(L))||_{TV} \leq (25 + \frac{t + t^2}{17})L_0 + 4t|L - L_0|.$$

Remarks. We haven't taken much care to keep the numerical constants small. The presence of the "t^2" terms in the bounds is an artifact of our crude argument around (6). On the other hand, getting errors of at least order $\tau \log(1/\tau)$ seems intrinsic to any technique of proof involving breaking time into blocks and using the fact that blocks are almost 1-dependent. Theorem 1 gave bounds on the exponential approximation by completely different methods, but it is not clear what different method could be used for Poisson approximation.

4 Final remarks.

1. We defined L-visits to a subset A in terms of "time". One could instead work in terms of "space" by considering B with $A \subset B$, defining a B-visit to be the period of time spent on A during one excursion inside B, and then defining $N_B(t)$ to be the number of B-visits before time t. In concrete examples, this might be more natural. Charles Stein, lecturing at Stanford in June 1991, outlined how to use this idea to obtain an exponential approximation for hitting times for random walk on the torus.

2. Let us briefly indicate what one can do by "soft" arguments, if one does not require explicit error bounds. Take a sequence of stationary reversible chains and subsets, and standardize time so that $E_\pi T = 1$. A subsequential weak limit of stationary point processes with "mixing times" $\to 0$ must be a point process with stationary independent increments and hence a compound Poisson process. To check the limit is a simple Poisson process, it is enough to check that $\mathrm{var}(N_L(t)) \to t$, and one can use our Lemma 8 or easier cruder estimates. Such an argument shows

Proposition 12 *Consider a sequence of reversible chains and subsets, with $E_\pi T = 1$. Suppose $\tau \to 0$. Let $L \to 0$ and $\frac{L}{\tau \log(1/\tau)} \to \infty$. Then*

$$(N_L(t); t \geq 0) \xrightarrow{d} (N(t); t \geq 0)$$

where $N(t)$ is a Poisson counting process of rate 1.

3. We also expect a compound Poisson approximation for the total sojourn time in A during $[0,t]$. One can certainly give soft limit theorems analogous to Proposition 12, but no results with explicit error bounds are known. It would be interesting if Stein's method could be developed to handle such compound Poisson approximations. Neither the simple trick in Arratia et al. [4] sec. 3.1 nor the more sophisticated ideas in Barbour et al [5] seem helpful.

References

[1] D.J. Aldous. *Probability Approximations via the Poisson Clumping Heuristic*. Springer-Verlag, 1989.

[2] D.J. Aldous and M. Brown. Inequalities for rare events in time-reversible Markov chains I. In M. Shaked and Y.L. Tong, editors, *Stochastic Inequalities*, volume 22 of *Lecture Notes*, pages 1–16. Institute of Mathematical Statistics, 1992.

[3] R. Arratia, L. Goldstein, and L. Gordon. Two moments suffice for Poisson approximations: the Chen-Stein method. *Ann. Probab.*, 17:9–25, 1989.

[4] R. Arratia, L. Goldstein, and L. Gordon. Poisson approximation and the Chen-Stein method. *Statist. Sci.*, 5:403–434, 1990.

[5] A.D. Barbour, L.H.Y. Chen, and W.-L. Loh. Compound Poisson approximation for nonnegative random variables via Stein's method. *Ann. Probab.*, 20:1843–1866, 1992.

[6] A.D. Barbour and L. Holst. Some applications of the Stein-Chen method for proving Poisson convergence. *Adv. in Appl. Probab.*, 21:74–90, 1989.

[7] A.D. Barbour, L. Holst, and S. Janson. *Poisson Approximation.* Oxford University Press, 1992.

[8] M. Brown. Consequences of monotonicity for Markov transition functions. Technical report, City College, CUNY, 1990.

[9] R.L. Smith. Extreme value theory for dependent sequences via the Stein-Chen method of Poisson approximation. *Stochastic Process. Appl.*, 30:317–327, 1988.