First hitting distribution in different regimes: a probabilistic proof of Cooper&Frieze's First Visit Time Lemma

Elisabetta Scoppola

Dipartimento di Matematica e Fisica
Università di Roma Tre

Joint works with: Francesco Manzo and Matteo Quattropani

(MQS: "A probabilistic proof of Cooper&Frieze's First Visit Time Lemma" - ALEA - 2021)
(MS:"Exact results on the first hitting via Conditional Strong Quasi-Stationary Times and Application to metastability" - JSP - 2019)

IM - UFRJ, September 27 - 2021

Motivation

Cover time of random walks on different random graphs Cooper and Frieze CF 2007-2008-2012, C et al 2013-2014.

One of the key ingredients of Cooper and Freze's analysis is the so called First Visit Time Lemma (FVTL).

Sequence of Markov chains $(X^{(n)})_{t\geq 0}$ on a growing state space of size n: $\mathcal{X}^{(n)}$
Fix $x \in \mathcal{X}^{(n)}$ and let τ_x be the hitting time of x:

$$\tau_x = \inf\{t \geq 0 \,|\, X_t^{(n)} = x\}. \tag{1}$$

Notation: $P^{(n)}$ is the transition matrix of $(X^{(n)})_{t\geq 0}$.

$$f(n) \sim g(n) \quad \iff \quad f(n) = g(n)(1+o(1)) \text{ i.e., } \lim_{n\to\infty} \frac{f(n)}{g(n)} = 1$$

$$f(n) \lesssim g(n) \quad \iff \quad \limsup_{n\to\infty} \frac{f(n)}{g(n)} \leq 1$$

$$\mu_t^{\alpha}(y) := \sum_{x\in\mathcal{X}^{(n)}} \alpha(x)\big(P^{(n)}\big)^t(x,y), \qquad \forall y \in \mathcal{X}^{(n)}.$$

C&F FVTL - Hypotheses

Assume

- for every sufficiently large n the chain is irreducible, admitting a unique invariant measure $\pi = \pi_n$
- there exists a time $T = T_n$ such that

$$\max_{x,y} \left| P^T(x,y) - \pi(y) \right| = O\left(\frac{1}{n^3}\right), \tag{2}$$

and

$$T \max_x \pi(x) = o(1), \qquad \min_x \pi(x) = \omega(n^{-2}) \tag{3}$$

- let $a = \frac{1}{KT}$ for a suitably large constant K, fix $x \in \mathcal{X}$, the truncated probability generating function

$$\mathrm{R}(z) = \sum_{t=0}^{T-1} P^t(x,x) z^t, \qquad \forall z \in \mathbb{C}$$

satisfies

$$\min_{|z| \leq 1+a} \mathrm{R}(z) \geq \theta$$

for some constant $\theta > 0$

C&F FVTL - Claim

then, for all $y \in \mathcal{X}$ and $t \geq 0$

$$\mathbb{P}_{\mu_T^y}\big(\tau_x > t\big) = \Big(1 + O(T\pi(x))\Big)\tilde{\lambda}_x^t + o\big(e^{-at/2}\big), \tag{4}$$

where

$$\tilde{\lambda}_x = \left(1 + \frac{\pi(x)}{R_T(x)(1 + O(T\pi(x)))}\right)^{-1}$$

where $R_T(x) \geq 1$ is the expected number of returns in x within the mixing time T

$$R_T(x) = \sum_{t=0}^{T} \mu_t^x(x) \geq 1. \tag{5}$$

Equivalent to

$$\mathbb{P}_y(X_s \neq x,\ \forall s \in (T, t+T)) \sim \left(1 - \frac{\pi(x)}{R_T(x)}\right)^t$$

Remarks on C&F FVTL

C&F Hypotheses (2) and (3) on mixing and on the invariant measure are strong but are typically satisfied by random walks on many models of random graphs.

The techniques used in the proof by Cooper and Frieze rely on probability arguments but also on tools from complex analysis and an analytical expansion of probability generating functions.

The result is equivalent to

$$\left|\frac{\mathbb{P}_{\mu_y^T}\big(\tau_x > t\big)}{\tilde{\lambda}_x^t} - 1\right| = O(T\pi(x)) + \tilde{\lambda}_x^{-t} o\big(e^{-at/2}\big) = o(1)$$

and by the hypotheses since $\big(1 + \frac{\pi(x)}{R_T(x)}\big)e^{-a/2} < 1$ the control on the distribution of τ_x is really good, also for large t.

Previous results

Exponential law of hitting times is a classic and widely studied topic in probability. Recall for instance the pioneering book by Keilson [K'79] and the papers by Aldous [A'82, AB'92-'93].

Two different regimes:

1) A single state m is frequently visited before τ_x. The hitting time is dominated by the sum of many i.i.d. excursion times and therefore it is almost exponential. See [K'79].

   Applied to study metastability: processes that are trapped for a long time in a part of their state space.
   [CGOV'84], [Sh'92],....., [OV'05], [BdH'16] and [BG'16], [FMNS'15], [FMNSS'16]

2) The chain is rapidly mixing
Analyzed in [A'82], where it is shown that

$$\sup_{t\geq 0}\left|\mathbb{P}_\pi(\tau_x > t) - e^{-\frac{t}{\mathbb{E}_\pi[\tau_x]}}\right| \leq \delta, \tag{6}$$

where

$$\delta \propto \frac{T}{\mathbb{E}_\pi[\tau_x]}\big(1+\log\frac{\mathbb{E}_\pi[\tau_x]}{T}\big)$$

By using the quasi-stationary measure introduced in the pioneering paper by Darroch-Seneta [DS'65] .
Note that Aldous' result concerns additive error bounds

$$\mathbb{P}_\pi(\tau_x > t) = e^{-\frac{t}{\mathbb{E}_\pi[\tau_x]}} + o(1)$$

and therefore it cannot provide first-order asymptotics of the exponential approximation when t is large, in contrast to the FVTL where a multiplicative bound is proved

$$\mathbb{P}_{\mu_y^T}\big(\tau_x > t\big) = \tilde{\lambda}_x^t(1+o(1))$$

Definitions and preliminary results

Let $[P]_x$ the sub-Markovian probability kernel obtained by removing the x-th row and column by the matrix P. Assume that $[P]_x$ is a primitive sub-Markovian kernel, i.e., all entries of $([P]_x)^m$ are positive for some $m \in \mathbb{N}$.

$$\mathbb{P}_y(\tau_x > t) = \sum_{z \neq x} \left([P]_x\right)^t (y, z).$$

By the Perron-Froboenius theorem there exists a unique probability distribution $\mu_x^\star$ and a leading eigenvalue $\lambda_x \in (0, 1)$ of $[P]_x$

$$\mu_x^\star [P]_x = \lambda_x \mu_x^\star, \tag{7}$$

Moreover, denoting by γ_x the corresponding right eigenvector, i.e.,

$$[P]_x \gamma_x = \lambda_x \gamma_x, \tag{8}$$

normalized by $\langle \gamma_x, \mu_x^\star \rangle = 1$ we have

$$\left([P]_x\right)^t (z, y) = \lambda_x^t \gamma_x(z) \mu_x^\star(y) + O(\beta^t) \tag{9}$$

with $\beta \in (0, \lambda_x)$.

The quasi-stationary measure $\mu_x^\star$ on $\mathcal{X} \setminus x$ is strictly related to the exponential behavior of τ_x:

$$\mathbb{P}_{\mu_x^\star}(\tau_x > t) = \sum_{z \neq x} \mu_x^\star(z) \sum_{y \neq x} \left([P]_x\right)^t (z, y) =$$

$$\lambda_x^t \sum_{y \neq x} \mu_x^\star(y) = \lambda_x^t.$$

The right eigenvector γ_x controls the dependence on the initial distribution of the probability of the event $\tau_x > t$.
Indeed by (9):

$$\lim_{t \to \infty} \frac{\mathbb{P}_y(\tau_x > t)}{\mathbb{P}_z(\tau_x > t)} = \frac{\gamma_x(y)}{\gamma_x(z)} \qquad y, z \neq x.$$

and for the eigenvectors the limits:

$$\lim_{t \to \infty} \mathbb{P}_z(X_t = y | \tau_x > t) = \lim_{t \to \infty} \frac{([P]_x)^t (z, y)}{\mathbb{P}_z(\tau_x > t)} = \mu^*(y) \quad \text{Yaglom limit}$$

$$\lim_{t \to \infty} \frac{\mathbb{P}_y(\tau_x > t)}{\lambda_x^t} = \gamma_x(y)$$

A randomized stopping time τ_π^α is a Strong Stationary Time (SST) for the Markov chain X_t with starting distribution α and stationary measure π, if for any $t \geq 0$ and $y \in \mathcal{X}$

$$\mathbb{P}_\alpha\left(X_t = y, \tau_\pi^\alpha = t\right) = \pi(y)\mathbb{P}_\alpha\left(\tau_\pi^\alpha = t\right),$$

which is equivalent to

$$\mathbb{P}_\alpha\left(X_t = y \middle| \tau_\pi^\alpha \leq t\right) = \pi(y) \tag{10}$$

By defining the separation distance

$$\mathrm{sep}(\mu_t^\alpha, \pi) := \max_{y \in \mathcal{X}} \left[1 - \frac{\mu_t^\alpha(y)}{\pi(y)}\right],$$

if τ_π^α is a SST then

$$\mathbb{P}_\alpha(\tau_\pi^\alpha > t) \geq \mathrm{sep}(\mu_t^\alpha, \pi), \qquad \forall t \geq 0, \tag{11}$$

and when (11) holds with the equal sign for every t, the SST is minimal. Moreover, a minimal SST always exists, (see [AD'86], [AD'87], [LevPerW]).

Local chain on $\mathcal{X} \setminus x$ also called Doob's transform:

For any $y, z \neq x$, define the stochastic matrix

$$\widetilde{P}(z, y) := \frac{\gamma_x(y)}{\gamma_x(z)} \frac{P(z, y)}{\lambda_x}. \tag{12}$$

More generally

$$\widetilde{P}^t(z, y) = \frac{\gamma_x(y)}{\gamma_x(z)} \frac{\left([P]_x\right)^t(z, y)}{\lambda_x^t} \qquad \forall t \geq 0. \tag{13}$$

It is immediate to show that $\widetilde{P}$ is a primitive matrix and has invariant measure

$$\nu(y) := \gamma_x(y)\mu_x^\star(y).$$

A preliminary estimate

Consider any initial measure α on $\mathcal{X} \setminus \{x\}$ and define the transformation

$$\tilde{\alpha}(y) := \frac{\alpha(y)\gamma_x(y)}{\langle \alpha, \gamma_x \rangle}, \qquad \forall y \neq x. \tag{14}$$

and the separation distance $\mathrm{sep}(\tilde{\mu}_t^{\tilde{\alpha}}, \nu)$ for the evolution of this chain starting from $\tilde{\alpha}$

$$\mathrm{sep}(\tilde{\mu}_t^{\tilde{\alpha}}, \nu) \equiv \tilde{s}^{\tilde{\alpha}}(t) := \sup_{y \neq x} \tilde{s}^{\tilde{\alpha}}(t, y).$$

with

$$\tilde{s}^z(t, y) := 1 - \frac{\widetilde{P}^t(z, y)}{\nu(y)}; \qquad \tilde{s}^{\tilde{\alpha}}(t, y) := \sum_{z \neq x} \tilde{\alpha}(z) \tilde{s}^z(t, y) \tag{15}$$

Note that $\tilde{s}^z(t) \in [0, 1]$ and recall that separation has the sub-multiplicative property

$$\tilde{s}(t + u) \leq \tilde{s}(t)\tilde{s}(u),$$

which in particular implies an exponential decay in time of $\tilde{s}$.

We have ([MS'19])

$$1 - \tilde{s}^{\tilde{\alpha}}(t) \le \frac{\mathbb{P}_\alpha(\tau_x > t)}{\lambda_x^t \langle \alpha, \gamma_x \rangle} \le 1 + \tilde{s}^{\tilde{\alpha}}(t) \left(\frac{1}{\min_y \gamma_x(y)} - 1 \right).$$

Indeed

$$\mathbb{P}_\alpha(\tau_x > t) = \sum_{y \neq x} \sum_{z \neq x} \alpha(z) \big([P]_x\big)^t (z, y) \tag{16}$$

$$= \sum_{y \neq x} \sum_{z \neq x} \alpha(z) \gamma_x(z) \lambda_x^t \mu_x^\star(y) \frac{\widetilde{P}^t(z, y)}{\nu(y)} \tag{17}$$

$$= \lambda_x^t \sum_{z \neq x} \alpha(z) \gamma_x(z) \sum_{y \neq x} \mu_x^\star(y) (1 - \tilde{s}^z(t, y)) \tag{18}$$

$$= \lambda_x^t \langle \alpha, \gamma_x \rangle \left(1 - \sum_{y \neq x} \mu_x^\star(y) \tilde{s}^{\tilde{\alpha}}(t, y) \right) \tag{19}$$

C&F regime in [MQS'21]

We consider the following asymptotic assumption for the sequence of Markov chains: There exist

- A real number $c > 2$.
- A diverging sequence $T = T(n)$.

such that

(HP1) Fast mixing:

$$\max_{x,y\in\mathcal{X}} |\mu_T^x(y) - \pi(y)| = o(n^{-c}).$$

(HP2) Small $\pi_{\max}$:

$$T \max_{x\in\mathcal{X}} \pi(x) = o(1).$$

(HP3) Large $\pi_{\min}$:

$$\min_{x\in\mathcal{X}} \pi(x) = \omega(n^{-2}).$$

First Visit Time Lemma in [MQS'21]

Theorem

Under the assumptions (HP1), (HP2) *and* (HP3) *for all* $x \in \mathcal{X}$, *it holds*

$$\sup_{t \geq 0} \left| \frac{\mathbb{P}_\pi(\tau_x > t)}{\lambda_x^t} - 1 \right| \longrightarrow 0, \tag{20}$$

and

$$\left| \frac{\lambda_x}{\left(1 - \frac{\pi(x)}{R_T(x)}\right)} - 1 \right| \longrightarrow 0. \tag{21}$$

Corollary

Under the same assumptions, for all $x \in \mathcal{X}$

$$\sum_{y \in \mathcal{X} \setminus \{x\}} \pi(y)\gamma_x(y) \to 1. \tag{22}$$

For all $t \geq 0$ and $x \in \mathcal{X}^{(n)}$, $\zeta_t(x)$ denotes the random time spent by the process in the state x within time t, i.e., $\zeta_t(x) := \sum_{s=0}^{t-1} 1_{X_s = x}$.

Theorem

Under the same set of assumptions, for every $x \in \mathcal{X}$ *and for all* $y \in \mathcal{X} \setminus \{x\}$*:*

$$\gamma_x(y) \lesssim 1,$$
$$\gamma_x(y) \gtrsim \left[1 - \mathbb{E}_y\left[\zeta_T(x)\right]\right]_+$$

Remarks on the results in [MQS'21]

Comparison C&F - MQS

- no hypotheses on generating functions, only probabilistic arguments
- starting distribution: π or μ_T^y are essentially equal by fast mixing (see definition of T)
- again multiplicative estimates but without precise magnitude of second order corrections
- estimates on $\gamma_x(y)$ for any y and $\tilde{s}^{\tilde{\pi}}(t) = o(1)$ can be proved, stationary and quasi-stationary distributions coincide in the thermodynamic limit
- in MQS a direct and simple proof.

Simple extension to first visit to a set G instead of x by a collapsing strategy $G \longrightarrow g$

On $\tilde{\mathcal{X}} = (\mathcal{X} \setminus G) \cup \{g\}$ define the transition matrix

$$\tilde{P}(x,y) = \begin{cases} P(x,y) & \text{if } x, y \neq g \\ \sum_{z \in G} P(x,z) & \text{if } x \neq g,\ y = g \\ \sum_{z \in G} \frac{\pi(z)}{\pi(G)} P(z,y) & \text{if } x = g,\ y \neq g \\ \sum_{y \in G} \sum_{z \in G} \frac{\pi(z)}{\pi(G)} P(z,y) & \text{if } x = y = g \end{cases} \tag{23}$$

we have immediately

$$\tilde{\pi} = \tilde{\pi}\tilde{P} \tag{24}$$

with $\tilde{\pi}(x)$ the projection of π on $\tilde{\mathcal{X}}$

$$\tilde{\pi}(x) = \begin{cases} \pi(x) & \text{if } x \neq g \\ \pi(G) & \text{if } x = g \end{cases}$$

The first hitting time to the set G under the original chain $(\mathcal{X}, P)$, starting at π, has the same law of the first hitting time to state g under the chain $(\tilde{\mathcal{X}}, \tilde{P})$ starting at $\tilde{\pi}$. Notice further that $[P]_G = [\tilde{P}]_g$. Hence, it is enough to check that the modified Markov chain satisfies assumptions (HP1)-(HP2)-(HP3), so G "small".

Main steps of the proof of FVTL

1) For all $x, y \in \mathcal{X}$ and $t > 0$ it holds

$$\mathbb{P}_{\mu_T^y}(\tau_x > t) \sim \mathbb{P}_\pi(\tau_x > t).$$

indeed for any y, z we have $\mu_T^y(z) \sim \pi(z)$:

$$\max_{y,z\in\mathcal{X}} \left| \frac{\mu_T^y(z)}{\pi(z)} - 1 \right| = \max_{y,z\in\mathcal{X}} \frac{1}{\pi(z)} \left|\mu_T^x(z) - \pi(z)\right|$$

$$\leq \frac{1}{\min_{z\in\mathcal{X}} \pi(z)} \max_{y,z\in\mathcal{X}} \left|\mu_T^y(z) - \pi(z)\right| \leq \frac{o(n^{-c})}{Cn^{-2}} = o(n^{-c+2})$$

so that

$$\mathbb{P}_{\mu_T^y}(\tau_x > t) = \sum_z \mu_T^y(z)\mathbb{P}_x(\tau_x > t) = (1 + o(1)) \sum_z \pi(z)\mathbb{P}_x(\tau_x > t)$$

$$= (1 + o(1))\mathbb{P}_\pi(\tau_x > t).$$

2) For all $x \in \mathcal{X}$ and for all $t > T$ it holds

$$\max_{y \in \mathcal{X}} \mathbb{P}_y(\tau_x > t) \lesssim \mathbb{P}_\pi(\tau_x > t - T).$$

indeed

$$\mathbb{P}_y(\tau_x > t) = \sum_{z \in \mathcal{X}} \mathbb{P}_y(X_T = z;\ \tau_x > T)\mathbb{P}_z(\tau_x > t - T)$$

$$\leq \sum_{z \in \mathcal{X}} \mathbb{P}_y(X_T = z)\mathbb{P}_z(\tau_x > t - T) \sim \sum_{z \in \mathcal{X}} \pi(z)\mathbb{P}_z(\tau_x > t - T) = \mathbb{P}_\pi(\tau_x > t - T)$$

3) for all $t > T$ it holds

$$\mathbb{P}_\pi(\tau_x > t - T) \sim \mathbb{P}_\pi(\tau_x > t)$$

Similar result in [FMNS'15] in a metastable regime.

From 2) +3)

$$\max_{y \in \mathcal{X}} \mathbb{P}_y(\tau_x > t) \lesssim \mathbb{P}_\pi(\tau_x > t).$$

4) For all $x \in \mathcal{X}$

$$\lambda_x^T \sim 1$$

and

$$\mathbb{E}_\pi[\tau_x] \sim \mathbb{E}_{\mu_x^\star}[\tau_x] = \frac{1}{1 - \lambda_x}$$

5) Abdullah result:

$$\mathbb{E}_\pi[\tau_x] \sim \frac{R_T(x)}{\pi(x)}$$

From 4)+5) we get

$$1 - \lambda_x \sim \frac{\pi(x)}{R_T(x)}$$

Proof of the FVTL

$$\mu_T^{\mu_x^\star}(y) = \mathbb{P}_{\mu_x^\star}(X_T = y) = \mathbb{P}_{\mu_x^\star}(X_T = y, \tau_x > T) + \mathbb{P}_{\mu_x^\star}(X_T = y, \tau_x \le T)$$
$$= \lambda_x^T \mu_x^\star(y) + (1 - \lambda_x) \sum_{s=1}^{T} \lambda_x^s \mu_{T-s}^x(y)$$

so that

$$\lambda_x^T \mu_x^\star(y) \le \mu_T^{\mu_x^\star}(y) \le \lambda_x^T \mu_x^\star(y) + (1 - \lambda_x)\mathbb{E}_x[\zeta_T(y)]$$

and for

$$\mathbb{P}_\pi(\tau_x > t) \sim^{1)} \mathbb{P}_{\mu_T^{\mu_x^\star}}(\tau_x > t)$$

we obtain

$$\mathbb{P}_\pi(\tau_x > t) \gtrsim \lambda_x^{t+T} \sim^{4)} \lambda_x^t$$

on the other hand

$$\mathbb{P}_\pi(\tau_x > t) \lesssim \sum_{y\in\mathcal{X}} \lambda_x^T \mu_x^\star(y) \mathbb{P}_y(\tau_x > t) + (1-\lambda_x) \sum_{y\in\mathcal{X}} \mathbb{E}_x[\zeta_T(y)] \mathbb{P}_y(\tau_x > t)$$
$$\le \lambda_x^{t+T} + (1 - \lambda_x) T \max_y \mathbb{P}_y(\tau_x > t) =^{2)+3)+4)+5)} \lambda_x^t + o\left(\mathbb{P}_\pi(\tau_x > t)\right)$$

Idea of the proof of step 3)

$$\frac{\mathbb{P}_\pi(\tau_x > t+T)}{\mathbb{P}_\pi(\tau_x > t)} \geq 1 - o(1), \quad \forall t > 0$$

Prove that for any t

$$\frac{\mathbb{P}_\pi(\tau_x > t+T)}{\mathbb{P}_\pi(\tau_x > t)} \geq 1 - \varepsilon \cdot \frac{\mathbb{P}_\pi(\tau_x > t-T)}{\mathbb{P}_\pi(\tau_x > t)}$$

with $\varepsilon = o(1)$ and consider the sequence $(y_i)_{i\geq 1}$

$$y_i := \frac{\mathbb{P}_\pi(\tau_x > (i+1)T)}{\mathbb{P}_\pi(\tau_x > iT)}.$$

We have

$$y_{i+1} \geq 1 - \frac{\varepsilon}{y_i}.$$

Being $\varepsilon < 1/4$, we can define

$$\bar{\varepsilon} := \frac{1}{2} - \sqrt{\frac{1}{4} - \varepsilon}$$

and get by induction

$$y_i \geq 1 - \bar{\varepsilon}, \qquad \forall i \geq 1.$$

Indeed, note that $\varepsilon = \bar{\varepsilon}(1 - \bar{\varepsilon}) < \bar{\varepsilon}$

$$y_1 = \frac{\mathbb{P}_\pi(\tau_x > 2T)}{\mathbb{P}_\pi(\tau_x > T)} = 1 - \frac{\mathbb{P}_\pi(\tau_x \in [T, 2T])}{\mathbb{P}_\pi(\tau_x > T)} \geq 1 - \frac{(T+1)\pi(x)}{1 - (T+1)\pi(x)}$$

$$\geq 1 - \frac{\varepsilon}{1 - \varepsilon} \geq 1 - \bar{\varepsilon}$$

and

$$y_{i+1} \geq 1 - \frac{\varepsilon}{y_i} \geq 1 - \frac{\varepsilon}{1 - \bar{\varepsilon}} \geq 1 - \bar{\varepsilon}.$$

The result of the induction can be immediately extended from times iT to general times $t = iT + t_0$ with $t_0 < T$ by noting that again we get

$$1 - \frac{(T + t_0)\pi(x)}{1 - t_0\pi(x)} \geq 1 - \frac{\varepsilon}{1 - \varepsilon}.$$

Idea of the proof of step 4)

$$\lambda_x^{2T} = \mathbb{P}_{\mu_x^\star}(\tau_x > 2T) = \sum_{z \neq x} \mathbb{P}_{\mu_x^\star}(X_T = z,\ \tau_x > T)\, \mathbb{P}_z(\tau_x > T)$$

$$= \sum_{z \neq x} \left[\mathbb{P}_{\mu_x^\star}(X_T = z) - \mathbb{P}_{\mu_x^\star}(X_T = z,\ \tau_x \leq T)\right] \mathbb{P}_z(\tau_x > T)$$

$$\sim^{1)} \mathbb{P}_\pi(\tau_x > T) - \sum_{z \neq x} \mathbb{P}_{\mu_x^\star}(X_T = z,\ \tau_x \leq T)\, \mathbb{P}_z(\tau_x > T)$$

$$\geq \mathbb{P}_\pi(\tau_x > T) - \max_z \mathbb{P}_z(\tau_x > T)\, \mathbb{P}_{\mu_x^\star}(\tau_x \leq T)$$

$$\sim^{2)} \mathbb{P}_\pi(\tau_x > T)\left(1 - \mathbb{P}_{\mu_x^\star}(\tau_x \leq T)\right) = \mathbb{P}_\pi(\tau_x > T)\left(1 - (1 - \lambda_x^T)\right).$$

Hence

$$\lambda_x^T \gtrsim \mathbb{P}_\pi(\tau_x > T) \geq 1 - (T+1)\pi(x)$$

Proof of step 5) (Abdullah): By [AF02]:

$$\mathbb{E}_\pi[\tau_x] = \frac{Z(x,x)}{\pi(x)},$$

where Z is the fundamental matrix,

$$Z(x,x) := \sum_{t=0}^{\infty} \mu_t^x(x) - \pi(x).$$

By the submultiplicativity of the sequence

$$D(t) := \max_{x,y} \left|\mu_t^x(y) - \pi(y)\right|,$$

i.e., $D(t+s) \leq 2D(t)D(s)$, $\forall t, s > 0$, and by fast mixing we get

$$\max_{x,y} \left|\mu_{kT}^x(y) - \pi(y)\right| \leq \left(\frac{2}{n^c}\right)^k, \qquad \forall k \in \mathbb{N}.$$

Noting that $R_T(x) \geq 1$, and $Tn^{-c} \leq T\pi_{\max}$

$$Z(x,x) = \sum_{t \leq T} \left(\mu_T^x(x) - \pi(x)\right) + T \sum_{k \geq 1} \left(\frac{2}{n^c}\right)^k$$

$$= R_T(x) + O(T\pi(x)) + O(Tn^{-c}) = R_T(x)(1 + o(1))$$

General remark

In both the regimes, (1) metastability and (2) hitting to a rare state for rapidly mixing chains, the exponential decay is due to the existence of two asymptotically different time scales:

$$\mathbb{E}\tau_x \gg T \quad \text{(local equilibrium)} \begin{cases} \text{fast recurrence to } m & \text{in case(1)} \\ \text{fast recurrence to } \pi & \text{in case(2)} \end{cases}$$

The two regimes can be understood in a common framework:

recurrence to a point $\Longleftrightarrow$ recurrence to a measure

The quasi-stationary measure plays the role of a recurrent measure before the hitting.

The hitting to the measure can be studied by extending the theory of strong stationary times to quasi-stationarity by means of conditional strong quasi-stationary time (CSQST).

Conditional Strong Quasi-Stationary Times (CSQST) and results in [MS'19]

Given an initial distribution α on $\mathcal{X} \setminus \{x\}$, a randomized stopping time $\tau_{\star,x}^{\alpha}$ is a CSQST if for any $y \neq x$ and $t \geq 0$

$$\mathbb{P}_\alpha(X_t = y, \tau_{\star,x}^{\alpha} = t) = \mu_x^\star(y)\mathbb{P}_\alpha(\tau_{\star,x}^{\alpha} = t < \tau_x)$$

or equivalently

$$\mathbb{P}_\alpha(X_{\tau_{\star,x}^{\alpha}} = y \big| \tau_{\star,x}^{\alpha} < \tau_x) = \mu_x^\star(y)$$

We prove in [MS'19] that for any α there exists a CSQST $\tau_{\star,x}^{\alpha}$ s.t. for any $t > 0$

$$\mathbb{P}_\alpha(\tau_x > t) = \lambda_x^t \langle \alpha, \gamma_x \rangle (1 - \tilde{s}^{\tilde{\alpha}}(t)) + \mathbb{P}_\alpha(\tau_{\star,x}^{\alpha} \wedge \tau_x > t)$$

with $\mathbb{P}_\alpha(\tau_{\star,x}^{\alpha} \wedge \tau_x > t)$ exponentially decreasing to 0 in t.

Remarks on the results in [MS'19]

Hypotheses on the chain are very general

- finite (or countable) state space, n is fixed
- the target x can be a set of states
- the chain is ergodic
- the starting distribution α is arbitrary

explicit dependence on α

$$\mathbb{P}_\alpha(\tau_x > t) = \lambda_x^t \langle \alpha, \gamma_x \rangle (1 - \tilde{s}^{\tilde{\alpha}}(t)) + \mathbb{P}_\alpha(\tau_{\star,x}^\alpha \wedge \tau_x > t) \qquad (25)$$

We can obtain an exponential decay estimate of $\mathbb{P}_\alpha(\tau_x > t)$ only if the local relaxation time

$$T := \sup_\alpha \mathbb{E}\left(\tau_{\star,x}^\alpha \wedge \tau_x\right)$$

is such that $T(1 - \lambda_x) \ll 1$.

However can be difficult to control some of the quantities involved in (25), in particular:

- the right eigenvector γ_x
- the decay of the separation distance $\tilde{s}^{\tilde{\alpha}}$

[MQS'21]+[MS'19]

In the regime of rare target and rapidly mixing chain :

- estimates on $\gamma_x(y)$ for all y
- starting from π, estimates on $\tilde{s}^{\tilde{\pi}}$

Stationary and quasi-stationary distributions coincide in the limit $n \to \infty$