Supplementary material
Integral estimation based on Markovian design

Romain Azaïs[a], Bernard Delyon[b] and François Portier[c]

[a] *Inria Nancy – Grand Est, Team BIGS and Institut Élie Cartan de Lorraine, Nancy, France*
[b] *Institut de recherches mathématiques de Rennes, Université de Rennes 1*
[c] *LTCI, CNRS, Télécom ParisTech, Université Paris-Saclay*

A Regeneration-based bounds for expectations

We employed the Nummelin splitting technique in order to exploit the independence between the blocks B_k, $k \in \mathbb{N}^*$, as described in section 2 of the associated paper. We have however taken care of giving conditions on the moments τ_A of the original chain $(X_i)_{i\in\mathbb{N}}$ rather than on the moments θ_a of the split chain $(Z_i)_{i\in\mathbb{N}}$.

Define, for any $p > 0$,

$$\xi(p) = \sup_{x\in A} \mathbb{E}_x[\tau_A^p].$$

We start with a lemma relating moments of θ_a to moments of τ_A.

Lemma ∗A.1. *Let $(X_i)_{i\in\mathbb{N}}$ be a Markov chain satisfying (5). Then, for any $x_0 \in \mathcal{E}$, $p \geqslant 1$,*

$$\mathbb{E}_{x_0}[\theta_a^p]^{1/p} \leqslant \frac{1}{e^{\lambda_0/p} - 1}\xi(p)^{1/p} + \mathbb{E}_{x_0}[\tau_A^p]^{1/p} \tag{$*1$}$$

$$\mathbb{E}_a[\theta_a^p]^{1/p} \leqslant \lambda_0^{-1}\frac{e^{\lambda_0/p}}{e^{\lambda_0/p} - 1}\xi(p)^{1/p}. \tag{$*2$}$$

Proof. We start by showing ($*1$). Suppose that $\mathbb{E}_{x_0}[\tau_A^p] < +\infty$ and $\sup_{x\in A}\mathbb{E}_x[\tau_A^p] < +\infty$, if not, the stated inequality is obviously satisfied. By the Minkowski inequality, we have

$$\begin{aligned}\mathbb{E}_{x_0}[\theta_a^p]^{1/p} &\leqslant \mathbb{E}_{x_0}[(\theta_a - \tau_A)^p]^{1/p} + \mathbb{E}_{x_0}[\tau_A^p]^{1/p} \\ &= \mathbb{E}_{x_0}[(\theta_a - \tau_A)^p \mathbf{1}_{\{\theta_a > \tau_A\}}]^{1/p} + \mathbb{E}_{x_0}[\tau_A^p]^{1/p}.\end{aligned}$$

Let $\mathscr{F}_{\tau_A}$ denote the σ-field of the past before τ_A and note that $\{\theta_a > \tau_A\}$ is $\mathscr{F}_{\tau_A}$-measurable. By the strong Markov property, it holds

$$\mathbb{E}_{x_0}[(\theta_a - \tau_A)^p \mathbf{1}_{\{\theta_a>\tau_A\}}|\mathscr{F}_{\tau_A}] = \mathbb{E}_{x_0}[(\theta_a - \tau_A)^p|\mathscr{F}_{\tau_A}]\mathbf{1}_{\{\theta_a>\tau_A\}} \leqslant \mathbf{1}_{\{\theta_a>\tau_A\}}\sup_{x\in A}\mathbb{E}_x[\theta_a^p].$$

Hence, setting $\gamma = \sup_{x\in A}\mathbb{E}_x[\theta_a^p]^{1/p}$, and because $\lambda_0 = \mathbb{P}_{x_0}(\theta_a = \tau_A) = \mathbb{P}_{x_0}(Y_{\tau_A} = 1)$,

$$\mathbb{E}_{x_0}[\theta_a^p]^{1/p} \leqslant \gamma(1 - \lambda_0)^{1/p} + \mathbb{E}_{x_0}[\tau_A^p]^{1/p}. \tag{$*3$}$$

In particular, it follows that

$$(1-(1-\lambda_0)^{1/p})\gamma \leqslant \sup_{x\in A} \mathbb{E}_x[\tau_A^p]^{1/p}.$$

Thus, ($*3$) becomes

$$\mathbb{E}_{x_0}[\theta_a^p]^{1/p} \leqslant (1-\lambda_0)^{1/p}(1-(1-\lambda_0)^{1/p})^{-1} \sup_{x\in A} \mathbb{E}_x[\tau_A^p]^{1/p} + \mathbb{E}_{x_0}[\tau_A^p]^{1/p}, \tag{$*4$}$$

and we obtain ($*1$) by using $1-\lambda_0 \leqslant e^{-\lambda_0}$. To get ($*2$), note that for every $x_0 \in A$,

$$\mathbb{E}_{x_0}[\theta_a^p \mathbf{1}_{\{Y_0=1\}}] = \lambda_0 \mathbb{E}_a[\theta_a^p].$$

It follows that $\mathbb{E}_a[\theta_a^p] \leqslant \lambda_0^{-1}\mathbb{E}_{x_0}[\theta_a^p]$ and we get the result from ($*1$), taking the supremum over A. □

We shall need also the following extension of (10).

Lemma $*$A.2. *Let $(X_i)_{i\in\mathbb{N}}$ be a Markov chain satisfying (2), (3) and (5). For any measurable function $h : \cup_{n\geqslant 1}\mathbb{R}^n \to \mathbb{R}$ such that $\mathbb{E}_\pi[h(X_1,...X_{\theta_a})] < +\infty$, (for any n the restriction of h to $\mathbb{R}^n$ is measurable), we have*

$$\alpha_0 \mathbb{E}_\pi[h(X_1,...X_{\theta_a})] = \mathbb{E}_a\Big[\sum_{i=1}^{\theta_a} h(X_i,...X_{\theta_a})\Big]. \tag{$*5$}$$

In particular, for any $p>0$,

$$\alpha_0\mathbb{E}_\pi[\theta_a^p] \leqslant \mathbb{E}_a[\theta_a^{p+1}] \leqslant (p+1)\alpha_0\mathbb{E}_\pi[\theta_a^p]. \tag{$*6$}$$

Proof. Having (2), (3) and (5) we can use the formula (10). Define $g(x) = \mathbb{E}_x[h(X_1,...X_{\theta_a})]$ and remark that, by the Markov property and the fact that $\{i<\theta_a\}$ is $\mathscr{F}_i$-measurable,

$$\begin{aligned}
&\mathbb{E}_a(g(X_i)\mathbf{1}_{\{i<\theta_a\}}) = \mathbb{E}_a(h(X_{i+1},...X_{\theta_a})\mathbf{1}_{\{i<\theta_a\}}),\\
&g(X_{\theta_a}) = \mathbb{E}_a(h(X_1,...X_{\theta_a})).
\end{aligned}$$

Then using (10) with g, we get

$$\begin{aligned}
\alpha_0\mathbb{E}_\pi[h(X_1,...X_{\theta_a})] &= \alpha_0\pi(g)\\
&= \mathbb{E}_a\Big[\sum_{i=1}^{\theta_a} g(X_i)\Big]\\
&= \mathbb{E}_a\Big[\sum_{i=1}^{\theta_a-1} h(X_{i+1},...X_{\theta_a})\Big] + \mathbb{E}_a\Big[h(X_1,...X_{\theta_a})\Big]\\
&= \mathbb{E}_a\Big[\sum_{i=1}^{\theta_a} h(X_i,...X_{\theta_a})\Big].
\end{aligned}$$

Concerning the second statement, we use the fact that $1+2^p+\dots\theta_a^p \geqslant \int_0^{\theta_a} x^p dx = \frac{\theta_a^{p+1}}{p+1}$ to write

$$\tfrac{1}{p+1}\mathbb{E}_a[\theta_a^{p+1}] \leqslant \mathbb{E}_a\Big[\sum_{i=1}^{\theta_a} i^p\Big] \leqslant \mathbb{E}_a[\theta_a^{p+1}].$$

We conclude by using ($*5$) with $h(x_1,\dots x_k) = k^p$, to show that the middle term is $\alpha_0\mathbb{E}_\pi[\theta_a^p]$. □

Lemma ∗A.3. *Let $(X_i)_{i\in\mathbb{N}}$ be a Markov chain satisfying (2), (3) and (5). For any $p>2$, there exists $C>0$ (depending on p, λ_0, α_0) such that for any measurable function f,*

$$\mathbb{E}_a\Big[\Big(\sum_{i=1}^{\theta_a} f(X_i)\Big)^2\Big] \leqslant C\left(\xi(p)^2\pi(f^2)+\xi(p)\mathbb{E}_\pi[f(X_0)^2\tau_A^p]\right).$$

Proof. Suppose that $f\geqslant 0$. If not, take $|f|$ instead of f. In what follows, we use the convention that empty sums equal 0. Applying Lemma ∗A.2 with

$$h(x_1,\dots x_k)=\Big(\sum_{j=1}^{k} f(x_j)\Big)^2-\Big(\sum_{j=2}^{k} f(x_j)\Big)^2=f(x_i)^2+2f(x_i)\sum_{j=2}^{k} f(x_j),$$

we get that

$$\begin{aligned}\mathbb{E}_a\Big[\Big(\sum_{i=1}^{\theta_a} f(X_i)\Big)^2\Big] &= \mathbb{E}_a\Big[\sum_{i=1}^{\theta_a} h(X_i,\dots X_{\theta_a})\Big]\\ &= \alpha_0\mathbb{E}_\pi\Big[f(X_1)\Big(f(X_1)+2\sum_{i=2}^{\theta_a} f(X_i)\Big)\Big]\\ &= \alpha_0\left(\pi(f^2)+2\mathbb{E}_\pi\Big[f(X_1)\sum_{i=2}^{\theta_a} f(X_i)\Big]\right).\end{aligned}$$

For any $p>2$, the second term is bounded as follows

$$\begin{aligned}\mathbb{E}_\pi\Big[f(X_1)\sum_{i=2}^{\theta_a} f(X_i)\Big] &= \sum_{i\geqslant 2}\mathbb{E}_\pi\Big[\mathbf{1}_{i\leqslant\theta_a} f(X_1)f(X_i)\Big]\\ &\leqslant \sum_{i\geqslant 2}\mathbb{E}_\pi\Big[i^{-p/2}\theta_a^{p/2}\, f(X_1)f(X_i)\Big]\\ &\leqslant \sum_{i\geqslant 2} i^{-p/2}\mathbb{E}_\pi\Big[f(X_1)^2\theta_a^p\Big]^{1/2}\mathbb{E}_\pi\Big[f(X_i)^2\Big]^{1/2}\\ &= \left(\sum_{i\geqslant 2} i^{-p/2}\right)\mathbb{E}_\pi\Big[f(X_1)^2\theta_a^p\Big]^{1/2}\mathbb{E}_\pi\Big[f(X_1)^2\Big]^{1/2}\\ &\leqslant \left(\frac{2}{p-2}\right)\mathbb{E}_\pi\Big[f(X_1)^2\theta_a^p\Big],\end{aligned}$$

where we have used $\sum_{i\geqslant 2} i^{-p/2}\leqslant\int_1^{+\infty} x^{-p/2}dx$. If $\widetilde{\theta}_a$ is the first time $k\geqslant 2$ such $Z_k\in a$, it holds

$$\mathbb{E}_\pi\Big[f(X_1)^2\theta_a^p\Big]\leqslant\mathbb{E}_\pi\Big[f(X_1)^2\widetilde{\theta}_a^p\Big]=\mathbb{E}_\pi\Big[f(X_0)^2(\theta_a^p+1)\Big]\leqslant 2\mathbb{E}_\pi\Big[f(X_0)^2\theta_a^p\Big].$$

Applying Lemma ∗A.1, equation (∗1), and using that for every $a,b\geqslant 0$, and $p>1$, $(a+b)^p\leqslant 2^{p-1}(a^p+b^p)$, we get

$$\mathbb{E}_\pi\Big[f(X_0)^2\theta_a^p\Big]\leqslant 2^{p-1}\left(\frac{1}{(e^{\lambda_0/p}-1)^p}\xi(p)\pi(f^2)+\mathbb{E}_\pi\left[f(X_0)^2\tau_A^p\right]\right).$$

Bringing everything together, we get

$$\mathbb{E}_a\Big[\Big(\sum_{i=1}^{\theta_a} f(X_i)\Big)^2\Big] \leqslant \alpha_0 \left(\pi(f^2) + \frac{2^{p+2}}{p-2}\left(\frac{1}{(e^{\lambda_0/p}-1)^p}\xi(p)\pi(f^2) + \mathbb{E}_\pi\left[f(X_0)^2\tau_A^p\right]\right)\right).$$

This leads to the stated result. □

Lemma ∗A.4. *Let $(X_i)_{i\in\mathbb{N}}$ be a Markov chain satisfying (2), (3) and (5). There exists $C > 0$ (depending on p, λ_0, α_0) such that, for any measurable function g with $\pi(g) = 0$, any $n \geqslant 1$ and $p > 2$,*

$$\mathbb{E}_\pi\Big[\Big(\sum_{i=1}^{n} g(X_i)\Big)^2\Big] \leqslant nC\left(\xi(p)^2\pi(g^2) + \xi(p)\mathbb{E}_\pi[g(X_0)^2\tau_A^p]\right).$$

Proof. Defining the blocks sums as (see equation (9))

$$G_k = \sum_{i=\theta_a(k)+1}^{\theta_a(k+1)} g(X_i),$$

(in this whole section we set $\sum_a^b = 0$ if $b < a$) G_k is an i.i.d. sequence and one has

$$\sum_{i=1}^{n} g(X_i) = \sum_{i=1}^{\theta_a\wedge n} g(X_i) + \sum_{k=1}^{l_n-1} G_k + \mathbf{1}_{\theta_a\leqslant n}\sum_{i=\theta_a(l_n)+1}^{n} g(X_i)$$

where l_n is the number of times Z_i visits a before n, i.e.,

$$l_n = \sum_{i=1}^{n} \mathbf{1}_{\{Z_i\in a\}}. \tag{∗7}$$

As the chain has been split into independent blocks, the process $L \mapsto \sum_{k=1}^{L} G_k$ is a martingale. The sequence (l_n) is random and is expected to be of order n. Since $l_n \leqslant n$, following Bertail and Clémençon (2011), page 21, we have

$$\Big|\sum_{i=1}^{n} g(X_i)\Big| \leqslant \sum_{i=1}^{\theta_a\wedge n} f(X_i) + \max_{1\leqslant L\leqslant n}\Big|\sum_{k=1}^{L} G_k\Big| + \mathbf{1}_{\theta_a\leqslant n}\sum_{i=\theta_a(l_n)+1}^{n} f(X_i),$$

where $f = |g|$ (considering f instead of g will help later for the treatment of the concerned terms). By the Minkowski inequality, denoting by $\|\cdot\|_2$ the $L_2(\mathbb{P}_\pi)$ norm, we have

$$\Big\|\sum_{i=1}^{n} g(X_i)\Big\|_2 \leqslant \Big\|\sum_{i=1}^{\theta_a\wedge n} f(X_i)\Big\|_2 + \Big\|\max_{1\leqslant L\leqslant n}\Big|\sum_{k=1}^{L} G_k\Big|\Big\|_2 + \Big\|\sum_{i=\theta_a(l_n)+1}^{n} f(X_i)\Big\|_2. \tag{∗8}$$

Using Doob's inequality, we have

$$\mathbb{E}_\pi \max_{1\leqslant L\leqslant n}\Big|\sum_{k=1}^{L} G_k\Big|^2 \leqslant 4n\mathbb{E}_\pi[|G_1|^2] = 4n\mathbb{E}_a\Big[\Big(\sum_{i=1}^{\theta_a} g(X_i)\Big)^2\Big],$$

then, from Lemma ∗A.3, we get for every $p > 2$ that there exist $\widetilde{C}$ such that

$$\mathbb{E}_\pi \max_{1 \leqslant L \leqslant n} \Big| \sum_{k=1}^{L} G_k \Big|^2 \leqslant 4n\widetilde{C} \left(\xi(p)^2 \pi(g^2) + \xi(p) \mathbb{E}_\pi [g(X_0)^2 \tau_A^p] \right)$$

This is also a crude bound for the third term in (∗8) since

$$\mathbb{E}_\pi \Big[\Big(\sum_{i=\theta_a(l_n)+1}^{n} f(X_i) \Big)^2 \Big] \leqslant \mathbb{E}_\pi \Big[\Big(\sum_{i=\theta_a(l_n)+1}^{\theta_a(l_n+1)} f(X_i) \Big)^2 \Big] = \mathbb{E}_a \Big[\Big(\sum_{i=1}^{\theta_a} f(X_i) \Big)^2 \Big].$$

Now we consider the first term in (∗8). Using Lemma ∗A.2 with

$$h(x_1, \ldots x_k) = \Big(\sum_{j=1}^{k \wedge n} f(x_j) \Big)^2,$$

we get

$$\begin{aligned}
\mathbb{E}_\pi \Big[\Big(\sum_{j=1}^{\theta_a \wedge n} f(X_j) \Big)^2 \Big] &= \alpha_0^{-1} \mathbb{E}_a \Big[\sum_{i=1}^{\theta_a} \Big(\sum_{j=i}^{\theta_a \wedge n} f(X_j) \Big)^2 \Big] \\
&= \alpha_0^{-1} \mathbb{E}_a \Big[\sum_{i=1}^{\theta_a \wedge n} \Big(\sum_{j=i}^{\theta_a \wedge n} f(X_j) \Big)^2 \Big] \\
&\leqslant n \alpha_0^{-1} \mathbb{E}_a \Big[\Big(\sum_{j=1}^{\theta_a} f(X_j) \Big)^2 \Big] \\
&\leqslant n \mathbb{E}_a \Big[\Big(\sum_{j=1}^{\theta_a} f(X_j) \Big)^2 \Big].
\end{aligned}$$

We conclude again with Lemma ∗A.3. □

B Proofs of section 3

B.1 Proof of Lemma 1

We start by proving (11). Define $k = \lfloor s \rfloor$. From the Taylor formula with integral remainder applied to $g(t) = \psi(x - tu)$, we get

$$\begin{aligned}
\psi(x - hu) - \psi(x) &= \sum_{j=1}^{k-1} \frac{h^j}{j!} g^{(j)}(0) + \int_0^h g^{(k)}(t) \frac{(h-t)^{k-1}}{(k-1)!} dt \\
&= \sum_{j=1}^{k} \frac{h^j}{j!} g^{(j)}(0) + \int_0^h (g^{(k)}(t) - g^{(k)}(0)) \frac{(h-t)^{k-1}}{(k-1)!} dt.
\end{aligned}$$

The first term is a polynomial in u which vanishes after integration with respect to K as by assumption, K is orthogonal to the first non-constant polynomial of degree $j \leq \lfloor s \rfloor$. Using the chain rule to compute $g^{(k)}$ and using basic inequalities with some combinatorics, we obtain that there exists a constant C (depending only on k and d) such that for every $t \in \mathbb{R}$,

$$|g^{(k)}(t) - g^{(k)}(0)| \leqslant C|u|_1^k \sum_{l\in\mathcal{P}_k} |\psi^{(l)}(x-tu) - \psi^{(l)}(x)|,$$

where $\mathcal{P}_k = \{(l_1,\dots l_d) \in \mathbb{N}^d : \sum_{i=1}^d l_i = k\}$. It follows that

$$\Big|\int_0^h (g^{(k)}(t) - g^{(k)}(0))\frac{(h-t)^{k-1}}{(k-1)!}dt\Big| \leqslant \frac{h^{k-1}C}{(k-1)!}\sum_{l\in\mathcal{P}_k}\int_0^h |\psi^{(l)}(x-tu) - \psi^{(l)}(x)|\,|u|_1^k dt.$$

Hence

$$\Big|\int (\psi(x-hu) - \psi(x))\,K(u)du\Big| \leqslant \frac{h^{k-1}C}{(k-1)!}\sum_{l\in\mathcal{P}_k}\int_0^h\int |\psi^{(l)}(x-tu) - \psi^{(l)}(x)|\,|u|_1^k\,|K(u)|\,du\,dt$$

and by the generalized Minkowski inequality (Folland, 1999, page 194)[1],

$$\begin{aligned}
\big\|\psi - \psi_h\big\|_{L_q(\pi)} &\leqslant \frac{h^{k-1}C}{(k-1)!}\sum_{l\in\mathcal{P}_k}\int\int\left(\int |\psi^{(l)}(x-tu) - \psi^{(l)}(x)|^q|u|_1^{qk}|K(u)|^q\mathbf{1}_{0\leqslant t\leqslant h}\pi(x)dx\right)^{1/q}du\,dt\\
&\leqslant \frac{h^{k-1}C}{(k-1)!}M_1\pi_\infty^{1/q}\sum_{l\in\mathcal{P}_k}\int\left(|tu|_1^{q(s-k)}|u|_1^{qk}|K(u)|^q\right)^{1/q}\mathbf{1}_{0\leqslant t\leqslant h}du\,dt\\
&= \frac{h^sC}{(k-1)!(s-k+1)}M_1\pi_\infty^{1/q}\,\#\{\mathcal{P}_k\}\int |u|_1^s|K(u)|du.
\end{aligned}$$

This implies (11).

To show (12), it suffices to provide an upper-bound proportional to h^s and another one proportional to h^r. Because $|\pi(\psi - \psi_h)| \leqslant \pi(|\psi - \psi_h|)$, applying (11) with $q = 1$, we obtain the upper-bound $C_1M_1\pi_\infty h^s$. By Fubini's theorem and using the symmetry about 0 of K, it holds

$$\int \pi(x)\psi_h(x)dx = \int \psi(x)\pi_h(x)dx. \qquad (*9)$$

Hence, introducing the probability density $\widetilde{\psi}(y) = \left(\int |\psi(x)|dx\right)^{-1}|\psi(y)|$, $y \in \mathbb{R}^d$, we find

$$\begin{aligned}
\left|\int \pi(x)(\psi(x) - \psi_h(x))dx\right| &= \left|\int \psi(x)(\pi(x) - \pi_h(x))dx\right|\\
&\leqslant \left(\int |\psi(x)|dx\right)\int \widetilde{\psi}(x)\,|\pi(x) - \pi_h(x)|\,dx\\
&= \left(\int |\psi(x)|dx\right)\|\pi - \pi_h\|_{L_1(\widetilde{\psi})}.
\end{aligned}$$

Applying (11) with $\widetilde{\psi}$ and π in place of π and ψ respectively, we get the bound $\widetilde{C}_1M_2\psi_\infty h^r$, for some $\widetilde{C}_1 > 0$ depending on K and r. Equation (12) is then deduced from these two bounds.

[1]For any nonegative measurable function $g(.,.)$ on $\mathbb{R}^{k+d}$, any σ-finite measures μ and ν, and any $q \geqslant 1$,

$$\left(\int\left(\int g(y,x)d\mu(y)\right)^q d\nu(x)\right)^{1/q} \leqslant \int\left(\int g(y,x)^q d\nu(x)\right)^{1/q}d\mu(y).$$

B.2 Proof of Proposition 3

For any f and $\widetilde{f}$ in $\mathcal{F}_1 \times \ldots \times \mathcal{F}_d$, we have

$$|\Psi(f) - \Psi(\widetilde{f}))| \leqslant \sum_{j=1}^{d} C_j(F)|f_j - \widetilde{f}_j|. \tag{$*10$}$$

Let us first prove that G is an envelope for $\mathcal{G}$. Applying ($*10$) with f_0 in place of $\widetilde{f}$, we get that $2\sum_{j=1}^{d} C_j(F)F_j$ is an envelope for the class $\mathcal{G} - \Psi(f_0)$. As a result G is an envelope for the class $\mathcal{G}$. The envelope property is proved.

Let Q be such that $Q(G^2) < +\infty$. Define the following probability measures on $\mathcal{X}$,

$$dQ_j = q_j^{-2} C_j(F)^2 \, dQ, \qquad \text{with } q_j^2 = \int C_j(F)^2 \, dQ.$$

Note that $q_j < +\infty$ is implied by $Q(G^2) < +\infty$. Let $\mathcal{C}_j$ denote a set of functions forming an $\epsilon\|F_j\|_{L_2(Q_j)}$-covering of the metric space $(\mathcal{F}_j, L_2(Q_j))$. For $f = (f_1, \ldots f_d) \in \mathcal{F}_1 \times \ldots \times \mathcal{F}_d$, there exists $\widetilde{f} = (\widetilde{f}_1, \ldots \widetilde{f}_d) \in \mathcal{C}_1 \times \ldots \times \mathcal{C}_d$ such that, using ($*10$) and the Minkowski inequality,

$$\begin{aligned}
\|\Psi(f) - \Psi(\widetilde{f})\|_{L_2(Q)} &\leqslant \sum_{j=1}^{d} \|(f_j - \widetilde{f}_j) C_j(F)\|_{L_2(Q)} \\
&\leqslant \sum_{j=1}^{d} q_j \|f_j - \widetilde{f}_j\|_{L_2(Q_j)} \\
&\leqslant \epsilon \sum_{j=1}^{d} q_j \|F_j\|_{L_2(Q_j)}.
\end{aligned}$$

The number of possible d-uplets $(\widetilde{f}_1, \ldots \widetilde{f}_d)$ is at most $\prod_{j=1}^{d} \#\{\mathcal{C}_j\}$, thus

$$\mathcal{N}\Big(\mathcal{G},\, L_2(Q), \epsilon \sum_{j=1}^{d} q_j \|F_j\|_{L_2(Q_j)}\Big) \leqslant \prod_{j=1}^{d} \mathcal{N}\left(\mathcal{F}_j,\, L_2(Q_j),\, \epsilon\|F_j\|_{L_2(Q_j)}\right).$$

We have

$$\begin{aligned}
\int G(x)^2 dQ &\geqslant \int |\Psi(f_0)|^2 dQ + 4\sum_{j=1}^{d} \int F_j^2 C_j(F)^2 \, dQ \\
&\geqslant \sum_{j=1}^{d} \int F_j^2 C_j(F)^2 \, dQ \\
&= \sum_{j=1}^{d} q_j^2 \|F_j\|_{L_2(Q_j)}^2.
\end{aligned}$$

Combining this with the Schwartz inequality gives

$$\sum_{j=1}^{d} q_j \|F_j\|_{L_2(Q_j)} \leqslant d^{1/2} \left(\sum_{j=1}^{d} q_j^2 \|F_j\|_{L_2(Q_j)}^2 \right)^{1/2} \leqslant d^{1/2} \|G\|_{L_2(Q)}.$$

Hence

$$\mathcal{N}\left(\mathcal{G},\, L_2(Q), \epsilon d^{1/2}\|G\|_{L_2(Q)}\right) \leqslant \prod_{j=1}^{d} \mathcal{N}\left(\mathcal{F}_j,\, L_2(Q_j),\, \epsilon\|F_j\|_{L_2(Q_j)}\right).$$

The VC class assumption on $\mathcal{F}_j$, with characteristics (A_j, v_j), implies that the right hand side is smaller than $\varepsilon^{-(v_1+\cdots+v_d)}A_1^{v_1}\ldots A_d^{v_d}$. This concludes the proof.

B.3 Proof of Proposition 4

The first statement is proved in van der Vaart and Wellner (1996), Example 2.5.4. The second statement, under (16a), is given by Lemma 22, (i), in Nolan and Pollard (1987) (the definitions are different than the ones we use; as stated page 789, their "Euclideanity" implies VC). Under (16b), invoking Lemma 22, (ii), in Nolan and Pollard (1987), the class of real valued functions $\{x \mapsto K^{(0)}(h^{-1}(y_1 - x_1))\,:\, y_1 \in \mathbb{R},\, h > 0\}$ is a uniformly bounded VC class of function. Then, since $\Psi(z) = z_1 \ldots z_d$ satisfies (15), Proposition 3 implies the conclusion.

B.4 Proof of Proposition 5

We begin by applying Proposition 3 to $\mathcal{F}_1 = \{(t,x) \mapsto \mathbf{1}_{t\leqslant M} \,:\, M \in \mathbb{R}\}$ and $\mathcal{F}_2 = \{(t,x) \mapsto K(h^{-1}(y-x))\,:\, y \in \mathbb{R}^d,\, h > 0\}$ (both classes are VC by Proposition 4), with $\Psi(z_1, z_2) = z_1 z_2$ which satisfies (15). The resulting class

$$\{(t,x) \mapsto \mathbf{1}_{t\leqslant M} K(h^{-1}(y-x))\,:\, y \in \mathbb{R}^d,\, h > 0,\, M \in \mathbb{R}\}$$

is uniformly bounded VC. Then we can consider the product of $\{(t,x) \mapsto t\}$ and $\mathcal{F}_3$. As for every $z_1, \tilde{z}_1 \in [-A_1, A_1]$ and $z_2, \tilde{z}_2 \in [-A_2, A_2]$, we have

$$|z_1 z_2 - \tilde{z}_1 \tilde{z}_2| \leqslant A_2|z_1 - \tilde{z}_1| + A_1|z_2 - \tilde{z}_2|,$$

this yields a VC class with envelope $(t,x) \mapsto 2((1 \vee K_\infty)|t| + (1 \vee |t|)K_\infty)$.

B.5 Proof of Theorem 6

We have to study

$$\widehat{\pi}(y) = n^{-1} \sum_{i=1}^{n} K_i(y),$$

where

$$K_i(y) = K_{h_n}(y - X_i).$$

As in the proof of Lemma ∗A.4, we will use the split chain defined in section 2, $\theta_a(k)$ will stand for the time of the k-th return to the set a ($\theta_a(1) > 0$), and l_n, defined in (∗7), is the number of such returns before n.

Recall that $\alpha_0 = \mathbb{E}_a[\theta_a]$. Using the stationarity and equation (10), its expectation under π can be computed as

$$\mathbb{E}_\pi[l_n] = \sum_{k=1}^{n} \mathbb{E}_\pi[\mathbf{1}_{Z_k \in a}] = n\mathbb{E}_\pi[\mathbf{1}_{\{Z_0 \in a\}}] = \frac{n}{\alpha_0}.$$

Let us now evaluate the variance of l_n. From Lemma ∗A.4 with with $g(z) = (\mathbf{1}_{\{z\in a\}} - \alpha_0^{-1})/n$, there exists $C > 0$ such that, for any $n \geqslant 1$,

$$\mathbb{E}_\pi\Big[\Big(\sum_{i=1}^{n} g(X_i)\Big)^2\Big] \leqslant nC\left(\pi(g^2) + \mathbb{E}_\pi[g(X_0)^2 \tau_A^{p_0}]\right).$$

Because

$$\mathbb{E}_\pi[\mathbf{1}_{\{Z_0\in a\}}\tau_A^{p_0}] = \int \mathbb{E}_z[\tau_A^{p_0}]\mathbf{1}_{\{z\in a\}} d\pi(z) = \mathbb{E}_a[\tau_A^{p_0}]\pi(a) < +\infty,$$

we conclude that there exists some constant $\widetilde{C} > 0$ such that

$$\mathbb{E}_\pi[(l_n/n - \alpha_0^{-1})^2] \leqslant \widetilde{C} n^{-1}. \tag{∗11}$$

Consequently,

$$\sup_{y\in\mathbb{R}^d}\left|\left(1 - \frac{\alpha_0(l_n-1)}{n}\right)\pi_{h_n}(y)\right| \leqslant \left|1 - \frac{\alpha_0(l_n-1)}{n}\right| \sup_{y\in\mathbb{R}^d}|\pi(y)| \longrightarrow 0, \quad \text{in } \mathbb{P}_\pi\text{-probability.}$$

Hence, in place of $\widehat{\pi}(y) - \pi_{h_n}(y)$, we can rather study

$$\widehat{T}(y) = \widehat{\pi}(y) - \frac{\alpha_0(l_n-1)}{n}\pi_{h_n}(y)$$

which will have a simpler expansion. The idea of the proof is to use the results available for the independent case. Since terms inside one block are not independent, the trick is to notice that we can consider the case when only one term in each block is picked at random. More precisely if $\Delta_k = \theta_a(k+1) - \theta_a(k)$ and I_k is a uniformly chosen point among $\{\theta_a(k)+1, \ldots, \theta_a(k+1)\}$, the variables

$$\widetilde{K}_k(y) = K_{I_k}(y), \qquad k = 1, \ldots l_n - 1,$$

satisfy

$$\mathbb{E}[\widetilde{K}_k(y)|\mathscr{F}_\infty] = \Delta_k^{-1} \sum_{i=\theta_a(k)+1}^{\theta_a(k+1)} K_i(y),$$

where $\mathscr{F}_\infty$ denote the σ-field generated by the whole chain. We can rewrite

$$\begin{aligned}
\widehat{T}(y) &= n^{-1}\sum_{i=1}^{\theta_a(1)} K_i(y) + n^{-1}\sum_{k=1}^{l_n-1}\left(\left(\sum_{i=\theta_a(k)+1}^{\theta_a(k+1)} K_i(y)\right) - \alpha_0\pi_{h_n}(y)\right) + n^{-1}\sum_{i=\theta_a(l_n)+1}^{n} K_i(y) \\
&= n^{-1}\sum_{i=1}^{\theta_a(1)} K_i(y) + \mathbb{E}\left\{n^{-1}\sum_{k=1}^{l_n-1}\left(\Delta_k\widetilde{K}_k(y) - \alpha_0\pi_{h_n}(y)\right)\Big|\mathscr{F}_\infty\right\} + n^{-1}\sum_{i=\theta_a(l_n)+1}^{n} K_i(y) \\
&= \widehat{T}_1(y) + \mathbb{E}[Z_n(y)|\mathscr{F}_\infty] + \widehat{T}_2(y).
\end{aligned}$$

Concerning the boundary terms $\widehat{T}_1$ and $\widehat{T}_2$, we have

$$\mathbb{E}_\pi\Big[\sup_{y\in\mathbb{R}^d}|\widehat{T}_1(y)|\Big] \leqslant n^{-1}\mathbb{E}_\pi\Big[\sup_{y\in\mathbb{R}^d}\sum_{i=1}^{\theta_a}|K_{h_n}(y-X_i)|\Big] \leqslant n^{-1}h_n^{-d}K_\infty\mathbb{E}_\pi[\theta_a],$$

and similarly,

$$\mathbb{E}_\pi\Big[\sup_{y\in\mathbb{R}^d}|\widehat{T}_2(y)|\Big] \leqslant n^{-1}\mathbb{E}_a\Big[\sup_{y\in\mathbb{R}^d}\sum_{i=1}^{\theta_a}|K_{h_n}(y-X_i)|\Big] = n^{-1}h_n^{-d}K_\infty\mathbb{E}_a[\theta_a].$$

We now consider the term $\mathbb{E}[Z_n(y)|\mathscr{F}_\infty]$. From the definition of I_1 and using (10), for any measurable function g with $\pi(g) < +\infty$, we have

$$\mathbb{E}_a[\Delta_1 g(X_{I_1})] = \mathbb{E}_a\Big[\theta_a\frac{1}{\theta_a}\sum_{i=1}^{\theta_a}g(X_i)\Big] = \alpha_0\pi(g). \tag{$*12$}$$

In particular, $\alpha_0\pi_{h_n}(y) = \mathbb{E}_a[\Delta_1\widetilde{K}_1(y)]$. It follows that

$$Z_n(y) = n^{-1}\sum_{k=1}^{l_n-1}\Big(\Delta_k\widetilde{K}_k(y) - \mathbb{E}_a[\Delta_1\widetilde{K}_1(y)]\Big).$$

We are planning to apply Theorem 2, but the problems for now are that l_n is random and Δ_k is not bounded. Define

$$m_n = (nh_n^{-d}/\log(n))^{1/(2p_0-1)}. \tag{$*13$}$$

We shall analyse the terms when $\Delta_k \leqslant m_n$ and $\Delta_k > m_n$ independently. The reason why such a value of m_n is considered shall be made clear in the next few lines (below equation ($*22$)). We have

$$\begin{aligned} Z_n(y) &= n^{-1}\sum_{k=1}^{l_n-1}\Big(\mu_k\widetilde{K}_k(y) - \mathbb{E}_a[\mu_1\widetilde{K}_1(y)]\Big) + n^{-1}\sum_{k=1}^{l_n-1}\Big(\nu_k\widetilde{K}_k(y) - \mathbb{E}_a[\nu_1\widetilde{K}_1(y)]\Big) \qquad (*14)\\ \mu_k &= \Delta_k\mathbf{1}_{\Delta_k\leqslant m_n}\\ \nu_k &= \Delta_k\mathbf{1}_{\Delta_k> m_n}. \end{aligned}$$

Choose $\eta_n = \sqrt{\log(n)/n}$, and set $l_n^0 = \lfloor n\alpha_0^{-1}\rfloor$, $l_n^- = \lfloor n(\alpha_0^{-1}-\eta_n)\rfloor$, $l_n^+ = \lfloor n(\alpha_0^{-1}+\eta_n)\rfloor$. By construction, as $n\to+\infty$,

$$n^{1/2}(l_n^+ - \alpha_0^{-1}) \to +\infty, \quad n^{1/2}(l_n^- - \alpha_0^{-1}) \to -\infty.$$

Therefore, from ($*11$), we obtain that the event $l_n^- \leqslant l_n - 1 \leqslant l_n^+$ has probability going to 1. Suppose from now on this event is realized. The number

$$l_n' = \big((l_n-1)\wedge l_n^+)\big)\vee l_n^-$$

is equal to $l_n - 1$. Since l_n' and l_n^0 both belong to $[l_n^-, l_n^+]$, for every sequence A_k, $k = 1, 2, \ldots$, it holds that

$$\Big|n^{-1}\sum_{k=1}^{l_n'}A_k\Big| \leqslant n^{-1}\Big|\sum_{k=1}^{l_n^0}A_k\Big| + n^{-1}\sum_{k=l_n^-}^{l_n^+}|A_k|.$$

Taking $A_k = \mu_k \widetilde{K}_k(y) - \mathbb{E}_a[\mu_k \widetilde{K}_k(y)]$, this gives

$$n^{-1} \sum_{k=1}^{l_n'} \Big(\mu_k \widetilde{K}_k(y) - \mathbb{E}_a[\mu_k \widetilde{K}_k(y)]\Big)$$

$$\leqslant n^{-1} \Big| \sum_{k=1}^{l_n^0} \big(\mu_k \widetilde{K}_k(y) - \mathbb{E}_a[\mu_k \widetilde{K}_k(y)]\big) \Big| + n^{-1} \sum_{k=l_n^-}^{l_n^+} |\mu_k \widetilde{K}_k(y) - \mathbb{E}_a[\mu_k \widetilde{K}_k(y)]|. \qquad (*15)$$

We treat the first term of (*15) by applying Theorem 2 with $\xi_i = (\Delta_i, X_{I_i})$, $i = 1, 2, \ldots$, and the class of functions $\{(t,x) \mapsto t\mathbf{1}_{\{t \leqslant m_n\}} K(h_n^{-1}(x-y)) \,:\, y \in \mathbb{R}^d\}$. This class being a subclass of (17) which is VC with envelope $F(t,x) = 2((1 \vee K_\infty)|t| + (1 \vee |t|)K_\infty)$ and characteristic (A, v) (in virtue of Proposition 5). Hence we can apply Theorem 2. We have to estimate the various quantities involved in (13). On the first hand,

$$\begin{aligned}
\sup_{f \in \mathcal{F}} \mathbb{E}[f(\xi_1)^2]) &= \sup_{y \in \mathbb{R}^d} \mathbb{E}_\pi[\Delta_1^2 \mathbf{1}_{\Delta_1 \leqslant m_n} K(h_n^{-1}(X_{I_1} - y))^2] \\
&\leqslant m_n \sup_{y \in \mathbb{R}^d} \mathbb{E}_\pi[\Delta_1 K(h_n^{-1}(X_{I_1} - y))^2] \\
&= m_n \sup_{y \in \mathbb{R}^d} \mathbb{E}_a\Big[\sum_{i=1}^{\theta_a} K(h_n^{-1}(X_i - y))^2\Big] \qquad \text{(cf. (*12))} \\
&= m_n \alpha_0 \sup_{y \in \mathbb{R}^d} \mathbb{E}_\pi[K(h_n^{-1}(X_1 - y))^2] \qquad \text{(cf. (*5))} \\
&\leqslant m_n \alpha_0 h_n^d \pi_\infty \int K(x)^2 dx \\
&= c^2 m_n h_n^d, \qquad\qquad c^2 = \alpha_0 \|\pi\|_\infty \int K(x)^2 dx.
\end{aligned}$$

On the other hand, using $(1 \vee |t|) \leqslant 1 + |t|$ and then (*2), we find

$$\mathbb{E}[F(\xi_1)^2] \leqslant 2((1 + K_\infty)\mathbb{E}|\Delta_1| + K_\infty) \leqslant C(1 + \sup_{x \in A} \mathbb{E}_x[\tau_A^2]),$$

for some $C > 0$. We choose

$$\sigma^2 = c^2 m_n h_n^d.$$

With this choice of σ, equation (13) will be satisfied if

$$c^2 m_n h_n^d \geqslant \frac{16 v n^{-1}}{2} \log\Big(A^2 \max\big(1, \mathbb{E}[F(\xi_1)^2]/c^2 m_n h_n^d\big)\Big) m_n^2 K_\infty^2.$$

Since $h_n \to 0$ and $m_n \to +\infty$, this condition will be met for n large enough if, as $n \to \infty$,

$$m_n \leqslant \frac{n h_n^d}{\log(h_n^{-1})}.$$

This is equivalent to

$$\frac{n h_n^{-d}}{\log(n)} \ll \Big(\frac{n h_n^d}{\log(h_n^{-1})}\Big)^{2p_0 - 1} \qquad (*16)$$

which is

$$1 \ll \Big(\frac{nh_n^{dp_0/(p_0-1)}}{\log(n)}\Big)^{2(p_0-1)}\Big(\frac{\log(n)}{\log(h_n^{-1})}\Big)^{2p_0-1}. \tag{$*17$}$$

This is satisfied indeed since the first term tends to infinity by assumption, and the fact that $nh_n^{dp_0/(p_0-1)} \to +\infty$ implies that the second one is bounded from below.

Computing the bound given in Theorem 2, multiplying by $(nh_n^d)^{-1}$, we obtain that

$$\mathbb{E}_\pi \sup_{y\in\mathbb{R}^d}\Big|n^{-1}\sum_{k=1}^{l_n^0}\mu_k\widetilde{K}_k(y) - \mathbb{E}_a[\mu_k\widetilde{K}_k(y)]\Big| \leqslant (nh_n^d)^{-1}C_0\sqrt{vl_n^0c^2m_nh_n^d\log\Big(A\big(1\vee\tfrac{\beta}{cm_n^{1/2}h_n^{d/2}}\big)\Big)}$$

But since

$$m_nh_n^d = \left(\frac{n}{\log(n)}\right)^{1/(2p_0-1)} h_n^{2d(p_0-1)/(2p_0-1)},$$

this quantity is larger than some negative power of n (cf. (18)) and using this for bounding the logarithm, we get

$$\mathbb{E}_\pi \sup_{y\in\mathbb{R}^d}\Big|n^{-1}\sum_{k=1}^{l_n^0}\mu_k\widetilde{K}_k(y) - \mathbb{E}_a[\mu_k\widetilde{K}_k(y)]\Big| \leqslant C'B(n,h_n,m_n) \tag{$*18$}$$

for some $C' > 0$ and where

$$B(n,h,m) = \sqrt{\frac{m\log(n)}{nh^d}}.$$

The second term of ($*15$) is smaller than

$$\Big|n^{-1}\sum_{k=l_n^-}^{l_n^+}|\mu_k\widetilde{K}_k(y) - \mathbb{E}_a[\mu_k\widetilde{K}_k(y)]| - \mathbb{E}_a|\mu_1\widetilde{K}_1(y) - \mathbb{E}_a[\mu_1\widetilde{K}_1(y)]|\Big|$$
$$+ n^{-1}(l_n^+ - l_n^-)\mathbb{E}_a(|\mu_1\widetilde{K}_1(y) - \mathbb{E}_a[\mu_1\widetilde{K}_1(y)]|).$$

Consider the class

$$\Big\{(\beta,x)\mapsto \big|\beta\mathbf{1}_{\{\beta\leqslant m_n\}}K(h^{-1}(x-y)) - \mathbb{E}_a[\mu_1K(h^{-1}(X_1-y))]\big| \,:\, y\in\mathbb{R}^d,\ h>0\Big\}.$$

This class is included in the larger class of functions $z\mapsto |f(z)-w|$, where f describes the VC class (17), and $w\in\mathbb{R}$ is ranging over the segment $A = [-\alpha_0K_\infty, \alpha_0K_\infty]$. This larger class is VC because, (i) the class $\{f(z)-w\}$ remains VC and (ii) the transformation $x\mapsto|x|$ being Lipschitz, we can apply Proposition 3. This is basically the same as before, with the only difference that now $l_n^+ - l_n^- \leqslant 3\eta_nn$, we obtain that there exists a constant $C>0$ such that

$$\mathbb{E}_\pi \sup_{y\in\mathbb{R}^d}\Big|n^{-1}\sum_{k=l_n^-}^{l_n^+}|\mu_k\widetilde{K}_k(y) - \mathbb{E}_\pi[\mu_1\widetilde{K}_1(y)]|\Big| \leqslant C\left(\sqrt{\eta}_nB(n,h_n,m_n) + \eta_n\mathbb{E}_\pi|\mu_1\widetilde{K}_1(y)|\right). \tag{$*19$}$$

From (∗12), we know that

$$\mathbb{E}_\pi|\mu_1\widetilde{K}_1(y)| \leqslant \mathbb{E}_a\left[\Delta_1|\widetilde{K}_1(y)|\right] = \alpha_0 \int |K_{h_n}(y-x)|\pi(x)dx \leqslant \alpha_0\pi_\infty \int |K(u)|du.$$

Then, bringing together (∗15), (∗18) and (∗19) gives that, for some $C > 0$,

$$\mathbb{E}_\pi \sup_{y\in\mathbb{R}^d} \left| n^{-1}\sum_{k=1}^{l'_n} \left(\mu_k\widetilde{K}_k(y) - \mathbb{E}_a[\mu_k\widetilde{K}_k(y)]\right)\right| \leqslant CB(n,h_n,m_n) \tag{∗20}$$

because $\eta_n \ll B(n,h_n,m_n)$ and $\eta_n \ll 1$. Concerning the second term in (∗14), since $l'_n \leqslant n$ and by Lemma ∗A.1, we have

$$\begin{aligned}
\mathbb{E}_\pi\Big[\sup_{y\in\mathbb{R}^d}\big|n^{-1}\sum_{k=1}^{l'_n}\nu_k\widetilde{K}_k(y)\big|\Big] &\leqslant K_\infty h_n^{-d}\mathbb{E}_\pi\Big[n^{-1}\sum_{k=1}^{n}\nu_k\Big]\\
&= K_\infty h_n^{-d}\mathbb{E}\pi\Big[\theta_a\mathbf{1}_{\theta_a>m_n}\Big]\\
&\leqslant K_\infty h_n^{-d}m_n^{-(p_0-1)}\mathbb{E}_\pi\Big[\theta_a{}^{p_0}\Big]\\
&\leqslant K_\infty h_n^{-d}m_n^{-(p_0-1)}\lambda_0^{-p_0}\frac{e^{\lambda_0}}{(e^{\lambda_0/p_0}-1)^{p_0}}\sup_{x\in A}\mathbb{E}_x[\tau_A^{p_0}].
\end{aligned}\tag{∗21}$$

Bringing together (∗14), (∗20), (∗21), we finally get, for some $C > 0$,

$$\mathbb{E}_\pi\Big[\sup_{y\in\mathbb{R}^d}\big|n^{-1}\sum_{k=1}^{l'_n}\Delta_k\widetilde{K}_k(y)\big|\Big] \leqslant C\left(B(n,h_n,m_n) + h_n^{-d}m_n^{-(p_0-1)}\right). \tag{∗22}$$

The value of m_n that balances these terms together is given by (∗13) and we obtain that there exists $C > 0$ such that

$$\mathbb{E}_\pi\Big[\sup_{y\in\mathbb{R}^d}\big|n^{-1}\sum_{k=1}^{l'_n}\Delta_k\widetilde{K}_k(y)\big|\Big] \leqslant C\left(\frac{\log(n)}{nh_n^{dp_0/(p_0-1)}}\right)^{(p_0-1)/(2p_0-1)}.$$

By assumption, this term goes to 0 as $n \to +\infty$. Let $\epsilon > 0$, we have that

$$\begin{aligned}
\mathbb{P}_\pi\Big(\sup_{y\in\mathbb{R}^d}\big|\mathbb{E}[Z_n(y)|\mathscr{F}_\infty]\big| \geqslant \epsilon\Big) &\leqslant \mathbb{P}_\pi\Big(\mathbb{E}[\sup_{y\in\mathbb{R}^d}|Z_n(y)|\ |\mathscr{F}_\infty] \geqslant \epsilon\Big)\\
&\leqslant \mathbb{P}_\pi\Big(\mathbb{E}[\sup_{y\in\mathbb{R}^d}|Z_n(y)|\,|\mathscr{F}_\infty] \geqslant \epsilon,\, l_n-1 = l'_n\Big) + \mathbb{P}_\pi(l_n-1\neq l'_n)\\
&\leqslant \epsilon^{-1}\mathbb{E}_\pi\Big[\mathbb{E}[\sup_{y\in\mathbb{R}^d}|Z_n(y)|\,|\mathscr{F}_\infty]\mathbf{1}_{\{l_n-1=l'_n\}}\Big] + \mathbb{P}_\pi(l_n-1\neq l'_n)\\
&= \epsilon^{-1}\mathbb{E}_\pi\Big[\sup_{y\in\mathbb{R}^d}|Z_n(y)|\mathbf{1}_{\{l_n-1=l'_n\}}\Big] + \mathbb{P}_\pi(l_n-1\neq l'_n)\\
&\leqslant \epsilon^{-1}\mathbb{E}_\pi\Big[\sup_{y\in\mathbb{R}^d}\big|n^{-1}\sum_{k=1}^{l'_n}\Delta_k\widetilde{K}_k(y)\big|\Big] + \mathbb{P}_\pi(l_n-1\neq l'_n).
\end{aligned}$$

Then we finish the proof by recalling that $l_n - 1 = l'_n$ whenever $l_n^- \leqslant l_n - 1 \leqslant l_n^+$, which has probability going to 1.

B.6 Proof of Corollary 7

Without loss of generality, because $h_n \to 0$, we can assume that $K(u) = 0$ for every $|u| \geqslant 1$. Theorem 6 implies that

$$\inf_{y\in Q} \widehat{\pi}(y) \geqslant \inf_{y\in Q} \pi_{h_n}(y) - \epsilon_n,$$

where $\epsilon_n = \sup_{y\in\mathbb{R}^d} |\widehat{\pi}(y) - \pi_{h_n}(y)| \to 0$, in $\mathbb{P}_\pi$-probability. Define, for any $x \in Q$ and $h > 0$,

$$\begin{aligned} b(x,h) &= \inf_{y\in Q,\, |y-x|\leqslant h} \pi(y), \\ M(x,h) &= \sup_{y\in Q,\, |y-x|\leqslant h} \pi(y). \end{aligned}$$

Let $K = K_+ + K_-$ be the decomposition of K with respect to the non-negative part and the negative part. Let $x \in Q$, for every $h > 0$, we have

$$\begin{aligned} \pi_h(x) &= \int \pi(x - hu)K(u)du \\ &\geqslant b(x,h)\int \mathbf{1}_{\{x-hu\in Q\}}K_+(u)du + M(x,h)\int \mathbf{1}_{\{x-hu\in Q\}}K_-(u)du \\ &= b(x,h)\int \mathbf{1}_{\{x-hu\in Q\}}K(u)du + (M(x,h) - b(x,h))\int \mathbf{1}_{\{x-hu\in Q\}}K_-(u)du \\ &\geqslant b\int \mathbf{1}_{\{x-hu\in Q\}}K(u)du - \sup_{x\in Q}|M(x,h) - b(x,h)|, \end{aligned}$$

By virtue of Heine's theorem, π is uniformly continuous on Q, hence $\sup_{x\in Q}|M(x,h) - b(x,h)| \to 0$ as $h \to 0$. Consequently, as $h_n \to 0$, we have for every $\epsilon > 0$, that $\inf_{x\in Q} \pi_{h_n}(x) \geqslant bc - \epsilon$. Choosing ϵ small enough and using that $\epsilon_n \to 0$, in $\mathbb{P}_\pi$-probability, gives the statement.

C Changing the initial measure

Appendix A focuses on Markov chains that either starts from their atom a, e.g., Lemma ∗A.3, or from their invariant measure π, e.g., Lemma ∗A.4. Some link between the underlying probabilities $\mathbb{P}_a$ and $\mathbb{P}_\pi$ is provided in Lemma ∗A.2. The following lemma turns out to be a useful ingredient to extend convergences in $\mathbb{P}_\pi$-probability to convergences in $\mathbb{P}_\nu$, ν being any measure absolutely continuous with respect to π.

Lemma ∗C.5. *Let $(X_i)_{i\in\mathbb{N}}$ be a Markov chain and let ν be a probability measure absolutely continuous with respect to π. Suppose that $f : \cup_{n\geqslant 1}\mathbb{R}^n \to \mathbb{R}^+$ is a bounded measurable function such that $\mathbb{E}_\pi f(X_1, \dots X_n) \to 0$ as $n \to +\infty$, then*

$$\mathbb{E}_\nu f(X_1, \dots X_n) \to 0.$$

Proof. Denote by q the Radon–Nikodym derivative of ν with respect to π. Let

$$g_n(x) = \mathbb{E}_x[f(X_1, \dots X_n)],$$

and $M > 0$ be such that $\sup_{n\geqslant 1} f(x_1, \dots x_n) < M$ for every sequence $(x_n)_{n\in\mathbb{N}^*}$. We have

$$\begin{aligned}
\mathbb{E}_\nu f(X_1, \dots X_n) &= \int g_n(x) d\nu(x) \\
&= \int g_n(x) q(x) d\pi(x) \\
&\leqslant A \int g_n(x) d\pi(x) + \int g_n(x) q(x) \mathbf{1}_{q(x)>A} d\pi(x) \\
&= A\mathbb{E}_\pi f(X_1, \dots X_n) + \mathbb{E}_\nu[g_n(X_0)\mathbf{1}_{q(X_0)>A}] \\
&\leqslant A\mathbb{E}_\pi f(X_1, \dots X_n) + M\mathbb{P}_\nu(q(X_0) > A),
\end{aligned}$$

for any $A > 0$. In the previous display, the term on the right-hand side can be made arbitrarily small by taking A large and for any such A, the term on the left-hand side goes to 0 by assumption. □

For application purposes, this simple lemma is fine. Notice however that by Corollary 6.9 of Nummelin (1984), under an additional aperiodicity assumption, the distribution of our Harris chain converges in total variation to π as soon as $\mathbb{E}_\pi[\tau_A] < \infty$ (see also Definition 5.5 and Proposition 5.15). In view of the equations ($*1$) and ($*6$), this means that $\sup_{x\in A} \mathbb{E}_x[\tau_A^2] < \infty$. The control of the bound in Lemma $*$A.4 already requires this. Given this, it is not difficult to check that the conclusion of Lemma $*$C.5 holds true even if ν is a Dirac measure δ_x, under the additional assumption that for all $k \in \{1, \dots n\}$

$$\sup_{(x_1,\dots x_n,y)\in\mathbb{R}^{n+1}} |f(x_1, \dots x_n) - f(x_1, \dots x_{k-1}, y, x_{k+1}, \dots x_n)| = \varepsilon_n \to 0.$$

This is obviously satisfied when f is an empirical mean over uniformly bounded terms. We have indeed for any fixed x_0

$$\begin{aligned}
\mathbb{E}_x f(X_1, \dots X_n) &= \mathbb{E}_x[f(x_0, \dots x_0, X_{k+1}, \dots X_n)] + kO(\varepsilon_n) \\
&= \int \mathbb{E}_y[f(x_0, \dots x_0, X_1, \dots X_{n-k})] P^k(x, dy) + kO(\varepsilon_n) \\
&= \mathbb{E}_\pi[f(x_0, \dots x_0, X_{k+1}, \dots X_n)] + O(\|\pi - P^k(x, .)\|) f_\infty + kO(\varepsilon_n) \\
&= \mathbb{E}_\pi[f(X_1, \dots X_n)] + O(\|\pi - P^k(x, .)\|) f_\infty + 2kO(\varepsilon_n).
\end{aligned}$$

This remark is of course not new, and is related to the coupling properties of the Harris chains, e.g., Proposition 29 in Roberts and Rosenthal (2004).

References

Bertail, P. and S. Clémençon (2011). A renewal approach to markovian u-statistics. *Mathematical Methods of Statistics 20*(2), 79–105.

Folland, G. B. (1999). *Real analysis* (Second ed.). Pure and Applied Mathematics (New York). John Wiley & Sons, Inc., New York. Modern techniques and their applications, A Wiley-Interscience Publication.

Nolan, D. and D. Pollard (1987). U-processes: rates of convergence. *The Annals of Statistics 15*(2), 780–799.

Nummelin, E. (1984). *General irreducible Markov chains and nonnegative operators*, Volume 83 of *Cambridge Tracts in Mathematics*. Cambridge University Press.

Roberts, G. O. and J. S. Rosenthal (2004). General state space Markov chains and MCMC algorithms. *Probab. Surv. 1*, 20–71.

van der Vaart, A. W. and J. A. Wellner (1996). *Weak convergence and empirical processes. With applications to statistics*. Springer Series in Statistics. New York: Springer-Verlag.