Supplement to "Topological signatures of periodic-like signals"

Frédéric Chazal **Bertrand Michel** **Wojciech Reise**

A Supplement for Persistence diagrams and total truncated persistence

A.1 A formal presentation of persistence homology

Let $h \in C(\mathbb{X}, \mathbb{R})$ be a continuous function on a compact topological space $\mathbb{X}$. The persistence diagram $D(h)$ in dimension 0 associated to h is a multiset of points in $\mathbb{R}^2$, where the coordinates are the values of local extrema of h. Each point can be interpreted as a local minimum paired with a local maximum. That pairing is constructed by tracking the evolution of connected components (CCs) in sublevel sets $h^{-1}(]-\infty, t])$ as t changes.

Formally, we rely on the theory of persistence modules developed in Chazal et al. (2016). For $t \in \mathbb{R}$, we consider the t-sublevel set of h, $\mathbb{X}_t = h^{-1}(]-\infty, t])$ and we define the *terminated module*

$$V_t \coloneqq \begin{cases} H_0(\mathbb{X}_t), & \text{if } t < \max h \\ 0, & \text{otherwise,} \end{cases} \tag{A.1}$$

where H_0 is the 0-dimensional singular homology. For any $s \leq t < \max h$, the inclusion $\mathbb{X}_s \to \mathbb{X}_t$ induces a morphism between the singular homology groups $\iota_s^t : V_s \to V_t$. For $t \geq \max h$, ι_s^t is the zero morphism. We call $\mathbb{V} = ((V_t)_{t\in\mathbb{R}}, (\iota_s^t)_{s<t\in\mathbb{R}})$ the terminated persistence module associated to h.

When $\mathbb{X}$ is compact and h continuous, for any $s < t$, the rank of ι_s^t is finite. This allows to define a measure on rectangles in $\mathbb{R}^2$: for any $a < b < c < d \in \mathbb{R}$,

$$m(]a,b] \times [c,d[) = \dim\left(\frac{im(\iota_b^c) \cap \ker(\iota_c^d)}{im(\iota_a^c) \cap \ker(\iota_c^d)}\right),$$

which counts the number of connected components of $\mathbb{X}_b$ that appeared later than in $\mathbb{X}_a$ and persist in $\mathbb{X}_c$, but not until $\mathbb{X}_d$. Finally, we can define the multiplicity of a point $(s,t) \in \mathbb{R}^2$ as $m(s,t) = \lim_{\delta\to 0^+} m(]s-\delta, s+\delta] \times [t-\delta, t+\delta[)$. The diagram $D(h)$ is then the set of points for which the multiplicity is positive. By convention, $m(s,s) = \infty$, for any point $(s,s) \in \Delta \coloneqq \{(t,t) \in \mathbb{R}\}$.

Remark A.1. *Usually, the persistence module associated to a continuous function is simply* $((H_0(\mathbb{X}_t))_{t\in\mathbb{R}}, (\iota_s^t)_{s<t\in\mathbb{R}})$. *We use terminated modules to force the essential component to have a finite death value, an operation particularly convenient for persistent diagrams of periodic functions, as we will see in Lemma 2.5.*

A.2 Proof of Lemma 2.1

For any $t \in \mathbb{R}$, the homeomorphism $g \coloneqq (\gamma_1^{-1} \circ \gamma_2) : [0,T] \to [0,T]$ maps the t-sublevel set of $f \circ \gamma_2$ to $f \circ \gamma_1$. Indeed,

$$\begin{aligned}(f \circ \gamma_1)^{-1}(]-\infty, t]) &= \{y \in [0,T] \mid (f \circ \gamma_1)(y) \leq t\} \\ &= \{y = g(x) \mid (f \circ \gamma_1)(g(x)) = (f \circ \gamma_2)(y) \leq t\} \\ &= g(\{y \in [0,T] \mid (f \circ \gamma_2)(y) \leq t\}).\end{aligned}$$

Therefore, g induces an isomorphism between the two corresponding persistence modules. So the corresponding persistence diagrams are the same (as well as any invariants there–of).

A.3 Proof of Proposition 2.3

We start with the lower-bound. Since pers is translation-invariant ($\mathrm{pers}_{p,\epsilon}(f+c) = \mathrm{pers}_{p,\epsilon}(f)$, for any constant $c > 0$), we can assume that $A_g = 2\|g\|_\infty$. Let $\Gamma : D(f) \to D(f+g)$ be a matching between the diagrams and denote by $c(\Gamma)$ the associated cost. Thanks to the bottleneck stability theorem, $\inf_\Gamma c(\Gamma) \leq \|g\|_\infty$. Then, for any $(b,d) \in D(f)$ and $(b',d') = \Gamma((b,d)) \in D(f+g)$, we have $d' - b' \geq d - b - 2c(\Gamma)$ and, for any $\delta > 0$, $D(f) \cap \Delta_{2c(\Gamma)+\delta} \subset \Gamma^{-1}(D(f+g) \cap \Delta_\delta)$. Then,

$$
\begin{aligned}
\mathrm{pers}_{p,\epsilon}^p(f+g) &= \sum_{(b',d') \in D(f+g)} w_\epsilon(b',d')^p \\
&\geq \sum_{(b',d') \in D(f+g) \cap \Delta_\delta} w_\epsilon(b',d')^p \\
&\geq \sum_{(b,d) \in \Gamma^{-1}(D(f+g) \cap \Delta_\delta)} w_\epsilon((b,d) - c(\Gamma)(-1,1))^p \\
&\geq \sum_{(b,d) \in D(f) \cap \Delta_{2c(\Gamma)+\delta}} w_{\epsilon+2c(\Gamma)}(b,d)^p.
\end{aligned}
$$

For $\delta = \epsilon$, the last quantity is equal to $\mathrm{pers}_{p,\epsilon+2c(\Gamma)}^p(f)$. By taking the infimum over all matchings Γ, we obtain $\mathrm{pers}_{p,\epsilon}^p(f+g) \geq \mathrm{pers}_{p,\epsilon+2\|g\|_\infty}^p(f)$.

For the upper-bound, we first note that when $A_f \leq \epsilon$, then $\mathrm{pers}_{p,\epsilon}^p(f) = 0$. For the non-trivial case, we follow the proof of Theorem 4.13 in Perez (2022). An upper-bound of the covering number of the image of f, at radius $\tau > 0$ is $T(2\Lambda/\tau)^{1/\alpha} + 1$, so that

$$
\begin{aligned}
\mathrm{pers}_{p,\epsilon}^p(f) &\leq p \int_\epsilon^{A(f)} \left(T \left(\frac{2\Lambda}{\tau} \right)^{1/\alpha} + 1 \right) (\tau - \epsilon)^{p-1} d\tau \\
&= (A_f - \epsilon)^p + pT(2\Lambda)^{1/\alpha} \int_\epsilon^{A(f)} \frac{(\tau-\epsilon)^{p-1}}{\tau^{1/\alpha}} d\tau
\end{aligned}
$$

We recall that since $A_f/\tau \geq 1$ and $1/\alpha \leq p-1$, $(A_f/\tau)^{1/\alpha} \leq (A_f/\tau)^{p-1}$, so

$$
\frac{(\tau-\epsilon)^{p-1}}{\tau^{1/\alpha}} = \frac{1}{A_f^{1/\alpha}} \left(\frac{A_f}{\tau} \right)^{1/\alpha} (\tau - \epsilon)^{p-1} \leq A_f^{p-1-1/\alpha} \left(1 - \frac{\epsilon}{\tau} \right)^{p-1}.
$$

Finally, by recognizing that $1 - \epsilon/\tau \leq 1 - \epsilon/A_f$, we obtain

$$
\begin{aligned}
\mathrm{pers}_{p,\epsilon}^p(f) &\leq (A_f - \epsilon)^p + pT(2\Lambda)^{1/\alpha} A_f^{p-1-1/\alpha} (1 - \epsilon/A_f)^{p-1} (A_f - \epsilon) \\
&\leq (A_f - \epsilon)(1 - \epsilon/A_f)^{p-1} [A_f^{p-1} + pT(2\Lambda)^{1/\alpha} A_f^{p-1-1/\alpha}] \\
&\leq (A_f - \epsilon)^p \left(1 + pT \left(\tfrac{2\Lambda}{A_f} \right)^{1/\alpha} \right) \\
&\leq (A_f - \epsilon)^p \left(1 + pT \left(\tfrac{2\Lambda}{\epsilon} \right)^{1/\alpha} \right),
\end{aligned}
$$

where we have used that $\epsilon^{1/\alpha} \leq A_f^{1/\alpha}$.

A.4 Proof of Proposition 2.4

Let $f, g \in C([0,T])$ such that $\|f - g\|_\infty < \epsilon/4$. Let $\Gamma : D(f) \to D(g)$ be a matching. Recall that $|w_\epsilon(b,d) - w_\epsilon(\eta_b, \eta_d)| \leq |b - \eta_b| + |d - \eta_d| \leq 2\|(b,d) - (\eta_b, \eta_d)\|_\infty$. In addition, if $d - b < \epsilon/2$, then both $w_\epsilon(b,d) = 0 = w_\epsilon(\Gamma(b,d))$. Using the technique $|x_2^p - x_1^p| = |p \int_{x_1}^{x_2} t^{p-1} dt| \leq p|x_2 - x_1| \max(x_1^{p-1}, x_2^{p-1})$

from the proof of Cohen-Steiner et al. (2010, Total Persistence Stability Theorem), we have

$$\begin{aligned}\left|\sum_{(b,d)\in D(f)} w_\epsilon(b,d)^p - \sum_{(b',d')\in D(g)} w_\epsilon(b',d')^p\right| &\le p\sum_{(b,d)\in D(f)} |w_\epsilon(b,d) - w_\epsilon(\Gamma(b,d))| \max_{x\in\{(b,d),\Gamma(b,d)\}} w_\epsilon(x)^{p-1} \\ &\le 2p\|f-g\|_\infty \sum_{\substack{(b,d)\in D(f)\\ d-b\ge\epsilon/2}} \max_{x\in\{(b,d),\Gamma(b,d)\}} w_\epsilon(x)^{p-1} \\ &\le p\|f-g\|_\infty \sum_{\substack{(b,d)\in D(f)\\ d-b\ge\epsilon/2}} (w_\epsilon(b,d) + 2\epsilon/4)^{p-1} \\ &= pC_f\|f-g\|_\infty.\end{aligned}$$

Since f is continuous on a compact domain, it is uniformly continuous, so that C_f is finite and does not depend on g.

For the Lipschitz character, we follow the proof of Perez (2022, Lemma 3.20). For two α-Hölder functions f, g with constant Λ,

$$\begin{aligned}\left|\sum_{(b,d)\in D(f)} w_\epsilon(b,d)^p - \sum_{(b',d')\in D(g)} w_\epsilon(b',d')^p\right| &\le p\sum_{(b,d)\in D(f)} |w_\epsilon(b,d) - w_\epsilon(\Gamma(b,d))| \max_{x\in\{(b,d),\Gamma(b,d)\}} w_\epsilon(x)^{p-1} \\ &\le 2p\|f-g\|_\infty \left(\sum_{(b,d)\in D(f)} w_\epsilon(b,d)^{p-1} + \sum_{(b',d')\in D(g)} w_\epsilon(b',d')^{p-1}\right) \\ &= 2p(\mathrm{pers}_{p-1,\epsilon}^{p-1}(D(f)) + \mathrm{pers}_{p-1,\epsilon}^{p-1}(D(g))\|f-g\|_\infty.\end{aligned}$$

By Proposition 2.3, $\mathrm{pers}_{p-1,\epsilon}^{p-1}(D(f)) \le T^{\alpha(p-1)}\Lambda^{p-1}(1+(p-1)2^{1/\alpha})$, so that

$$|\mathrm{pers}_{p,\epsilon}^p(D(f)) - \mathrm{pers}_{p,\epsilon}^p(D(g))| \le 4pT^{\alpha(p-1)}\Lambda^{p-1}(1+(p-1)2^{1/\alpha})\|f-g\|_\infty.$$

A.5 Proof of Lemma 2.5

Let $M \coloneqq \max\phi$, $c \coloneqq \inf\{x\in[0,1] \mid \phi(x) = M\}$ and $N = \max\{n\in\mathbb{N} \mid c+n \le R\}$. Consider the persistence modules defined by (A.1) for $\phi|_{[0,c]}$, $\phi|_{[c,c+N]}$ and $\phi|_{[c+N,R]}$. For $t < M$, $\phi|_{[0,c]}{}^{-1}(]-\infty,t]) \cap \phi|_{[c,c+N]}{}^{-1}(]-\infty,t]) \subset \{c\}$ and $\phi(c) = M$, so that intersection is empty and the same holds for $\phi|_{[c+N,R]}$ and $\phi|_{[c,c+N]}$. Therefore,

$$\begin{aligned}H_0(\phi|_{[0,R]}{}^{-1}(]-\infty,t])) \simeq{}& H_0(\phi|_{[0,c]}{}^{-1}(]-\infty,t])) \oplus H_0(\phi|_{[c,c+N]}{}^{-1}(]-\infty,t])) \\ &\oplus H_0(\phi|_{[c+N,R]}{}^{-1}(]-\infty,t])).\end{aligned} \tag{A.2}$$

Since the isomorphism is induced by inclusions, it is an isomorphism between the persistence modules restricted to $t \in]-\infty, M[$. By definition (A.1), the persistence modules are all 0 for $t \ge M$, so both sides of (A.2) are trivially isomorphic for $t \ge M$. Therefore, the persistence modules (on $t \in \mathbb{R}$) are isomorphic.

By repeating the same argument as above, we can show that the persistence module of $\phi|_{[c,c+N]}$ is the direct sum of the persistence modules of $(\phi|_{[c+n,c+n+1]})_{n=0}^{N-1}$. Then, for any $n = 0,\dots,N-1$, $g_n : x \mapsto x+n$ is an isomorphism between the sub level set of $\phi|_{[c,c+1]}$ and $\phi|_{[c+n,c+n+1]}$, so the persistence module of $\phi|_{[c,c+N]}$ is isomorphic to the direct sum of N copies of $\phi|_{[c,c+1]}$. Thus, (A.2) becomes

$$\begin{aligned}H_0(\phi|_{[0,R]}{}^{-1}(]-\infty,t])) \simeq{}& \left(\bigoplus_{n=0}^{N-1} H_0(\phi|_{[c,c+1]}{}^{-1}(]-\infty,t]))\right) \oplus H_0(\phi|_{[0,c]}{}^{-1}(]-\infty,t])) \\ &\oplus H_0(\phi|_{[c+N,R]}{}^{-1}(]-\infty,t])).\end{aligned}$$

The second crucial observation is that the diagram of a direct sum of two persistence modules is the union of diagrams. The case of interval decomposable modules is treated in Chazal et al. (2016, Proposition 2.16).

The persistence modules that we consider are q-tame (Chazal et al., 2016, Theorem 3.33), so they do not necessarily admit an interval decomposition. Recall that the persistence diagram is computed via rectangle measures (Chazal et al., 2016, Section 3), defined with ranks of inclusion morphisms. For two persistence modules $\mathbb{V} = (V_t)_{t\in\mathbb{R}}$, $\mathbb{W} = (W_t)_{t\in\mathbb{R}}$ and any $s, t \in \mathbb{R}$, we have that $\mathrm{rank}((V \oplus W)_s \to (V \oplus W)_t) = \mathrm{rank}(V_s \to V_t) + \mathrm{rank}(W_s \to W_t)$. This shows that the two rectangle measures $(\mu_V + \mu_W)$ and $\mu_{V\oplus W}$ are equal and so are their persistence diagrams. If we denote by $D_1 \coloneqq D(\phi|_{[c+n,c+n+1]})$ and by D' the diagram of the sum of the rectangle measures of the $\phi|_{[0,c]}$ and $\phi|_{[c+N,R]}$, then (2) follows.

We now need to bound the p-persistence of the remainder. Denote by $\mathbb{U}$ and $\mathbb{V}$ the persistence modules associated to $\phi|_{[0,c]}$ and $\phi|_{[0,c]}$ respectively. For any $t \in \mathbb{R}$, ${\phi|_{[0,c]}}^{-1}(]-\infty,t]) \subset {\phi|_{[c-1,c]}}^{-1}(]-\infty,t])$ induces a map $U_t \to V_t$. We claim that it is an injective morphism between persistence modules. Hence, $\mathrm{rank}(U_s \to U_t) \leq \mathrm{rank}(V_s \to V_t)$ for any $s < t \in \mathbb{R}$ and both are finite. Hence, to every point $(b, d) \in D(\phi|_{[0,c]})$ with $b < d$, we can assign a point $(b', d') \in D(\phi|_{[-1+c,c]})$ in such a way that this assignment is injective (considered with multiplicity) and such that $b' \leq b < d \leq d'$. So, $\mathrm{pers}^p_{p,\epsilon}(D(\phi|_{[0,c]})) \leq \mathrm{pers}^p_{p,\epsilon}(D(\phi|_{[-1+c,c]}))$. A similar argument shows that $\mathrm{pers}^p_{p,\epsilon}(D(\phi|_{[c+N,R]})) \leq \mathrm{pers}^p_{p,\epsilon}(D(\phi|_{[c+N,c+N+1]}))$.

B Complements for the normalized functionals of persistence

B.1 Lipschitz constant for k^{pi} and $k^{pi,t}$

First, $(x, y) \mapsto \exp(-(x^2 + y^2))$ is $2\sqrt{2}/e-$Lipschitz with respect to the Euclidean norm, so $4/3-$Lipschitz for the Minkowski norm. Consider the kernel

$$k^{pi,t}(b, d)(x, y) = \tfrac{1}{2\pi\sigma^2}\left(2 - \tfrac{\|(b,d)-(x,y)\|_\infty}{\sigma}\right)_+^r \exp\left(-\tfrac{(b-x)^2+(d-y)^2}{2\sigma^2}\right).$$

Then, for $r > 1$, the first term is $r2^r/\sigma$-Lipschitz,

$$\begin{aligned}
&\left|\left(2 - \tfrac{\|(b,d)-(x,y)\|_\infty}{\sigma}\right)_+^r - \left(2 - \tfrac{\|(b',d')-(x,y)\|_\infty}{\sigma}\right)_+^r\right| \\
=&\left|\int_0^1 \tfrac{d}{dt}\left(2 - \tfrac{\|(b,d)+(b'-b,d'-d)t-(x,y)\|_\infty}{\sigma}\right)_+^r dt\right| \\
\leq&\int_0^1 \left| r\left(2 - \tfrac{|b+(b'-b)t-x|}{\sigma}\right)_+^{r-1} (-1)^{b-x>b'-bt}\,\tfrac{(b'-b)}{\sigma} 1_{|b+(b'-b)t-x|\geq|d+(d-d')t-y|}\right. \\
&\left.\quad + r\left(2 - \tfrac{|d+(d'-d)t-y|}{\sigma}\right)_+^{r-1} (-1)^{d-y>d'-dt}\,\tfrac{(d'-d)}{\sigma} 1_{|b+(b'-b)t-x|\leq|d+(d-d')t-y|}\right| dt. \\
\leq&\int_0^1 \tfrac{r}{\sigma}\left(\left(2 - \tfrac{|b+(b'-b)t-x|}{\sigma}\right)_+^{r-1} |b-b'| + r\left(2 - \tfrac{|d+(d'-d)t-y|}{\sigma}\right)_+^{r-1} |d-d'|\right) dt \\
\leq&\tfrac{r}{\sigma}\left((2 - \tfrac{\min(|b-x|,|b'-x|)}{\sigma})_+^{r-1}|b-b'| + r(2 - \tfrac{\min(|d-y|,|d'-y|)}{\sigma})_+^{r-1}|d-d'|\right) \\
\leq&\tfrac{2r}{\sigma}(2 - \tfrac{\min(\|(b,d)-(x,y)\|_\infty,\|(b',d')-(x,y)\|_\infty)}{\sigma})_+^{r-1}\|(b,d)-(b',d')\|_\infty \\
\leq&\tfrac{2^r r}{\sigma}\|(b,d)-(b',d')\|_\infty.
\end{aligned}$$

Then, we obtain

$$\begin{aligned}
&|k^{pi,t}(b,d)(x,y) - k^{pi,t}(b',d')(x,y)| \\
\leq&\tfrac{1}{2\pi\sigma^2}\left|\left(2 - \tfrac{\|(b,d)-(x,y)\|_\infty}{\sigma}\right)_+^r - \left(2 - \tfrac{\|(b',d')-(x,y)\|_\infty}{\sigma}\right)_+^r\right| \exp\left(-\tfrac{(b-x)^2+(d-y)^2}{2\sigma^2}\right) \\
&+ \tfrac{1}{2\pi\sigma^2}\left(2 - \tfrac{\|(b',d')-(x,y)\|_\infty}{\sigma}\right)_+^r \left|\exp\left(-\tfrac{(b-x)^2+(d-y)^2}{2\sigma^2}\right) - \exp\left(-\tfrac{(b'-x)^2+(d'-y)^2}{2\sigma^2}\right)\right| \\
\leq&\tfrac{1}{2\pi\sigma^2}\tfrac{2^r r}{\sigma}\|(b,d)-(b',d')\|_\infty + \tfrac{1}{2\pi\sigma^2}2^r\tfrac{4}{e}\left\|\left(\tfrac{b-x}{\sigma},\tfrac{d-y}{\sigma}\right) - \left(\tfrac{b'-x}{\sigma},\tfrac{d'-y}{\sigma}\right)\right\|_\infty \\
\leq&\tfrac{2^{r-1}}{\pi\sigma^3}(r+2)\|(b,d)-(b'd')\|_\infty.
\end{aligned}$$

B.2 Proof of Proposition 2.10

Let $\Gamma : D_1 \to D_2$ be a matching between the two diagrams, we have

$$\begin{aligned}\|\rho(D_1) - \rho(D_2)\|_{\mathcal{H}} \leq & \sum_{x\in D_1} w_\epsilon(x)^p \|k(x) - k(\Gamma(x))\|_{\mathcal{H}} + \|k(\Gamma(x))\|_{\mathcal{H}} |w_\epsilon(x)^p - w_\epsilon(\Gamma(x))^p| \\ \leq & \sup_{x\in D_1} \|k(x) - k(\Gamma(x))\|_{\mathcal{H}} \sum_{x\in D_1} w_\epsilon(x)^p \\ & + \sup_{x\in D_1} \|k(\Gamma(x))\|_{\mathcal{H}} \sum_{x\in D_1} |w_\epsilon(x)^p - w_\epsilon(\Gamma(x))^p| \\ \leq & L_k d_B(D_1, D_2) \mathrm{pers}^p_{p,\epsilon}(D_1) \\ & + p(L_k U + C) \sum_{x\in D_1} |w_\epsilon(x) - w_\epsilon(\Gamma(x))| (w_\epsilon(x)^{p-1} + w_\epsilon(\Gamma(x))^{p-1}),\end{aligned}$$

where in the last inequality, we used that

$$\|k(\Gamma(x))\|_{\mathcal{H}} \leq L_k \|k(x_1, x_2) - k(\tfrac{x_1+x_2}{2}, \tfrac{x_1+x_2}{2})\|_{\mathcal{H}} + \|k\left(\tfrac{x_1+x_2}{2}, \tfrac{x_1+x_2}{2}\right)\|_{\mathcal{H}} = L_k \tfrac{x_2 - x_1}{2} + C.$$

The sum in the second term can be bounded from above by $d_B(D_1, D_2)(\mathrm{pers}^{p-1}_{p-1,\epsilon}(D_1) + \mathrm{pers}^{p-1}_{p-1,\epsilon}(D_2))$.

Consider now the normalized version.

$$\begin{aligned}\|\bar\rho(D_1) - \bar\rho(D_2)\|_{\mathcal{H}} \leq & \frac{\|\rho(D_1) - \rho(D_2)\|_{\mathcal{H}}}{\sum_{x\in D_1} w_\epsilon(x)^p} + \|\bar\rho(D_2)\|_{\mathcal{H}} \frac{|\sum_{x\in D_1} w_\epsilon(x)^p - \sum_{y \in D_2} w_\epsilon(y)^p|}{\sum_{x\in D_1} w_\epsilon(x)^p} \\ \leq & d_B(D_1, D_2) \left(L_k + p(L_k U + C) \frac{\mathrm{pers}^{p-1}_{p-1,\epsilon}(D_1) + \mathrm{pers}^{p-1}_{p-1,\epsilon}(D_2)}{\mathrm{pers}^p_{p,\epsilon}(D_1)} \right) \\ & + p(L_k U + C) d_B(D_1, D_2) \frac{\mathrm{pers}^{p-1}_{p-1,\epsilon}(D_1) + \mathrm{pers}^{p-1}_{p-1,\epsilon}(D_2)}{\mathrm{pers}^p_{p,\epsilon}(D_1)} \\ \leq & \left(L_k + 2p(L_k U + C) \frac{\mathrm{pers}^{p-1}_{p-1,\epsilon}(D_1) + \mathrm{pers}^{p-1}_{p-1,\epsilon}(D_2)}{\mathrm{pers}^p_{p,\epsilon}(D_1)} \right) d_B(D_1, D_2).\end{aligned}$$

Combine $\mathrm{pers}^{p-1}_{p-1,\epsilon}(D_1) + \mathrm{pers}^{p-1}_{p-1,\epsilon}(D_2) \leq 2 \max_{k=1,2} \mathrm{pers}^{p-1}_{p-1,\epsilon}(D_k)$ with the observation that the bound is symmetric so that we can have $\mathrm{pers}^p_{p,\epsilon}(D_2)$ in the denominator.

B.3 Stability of the functional via bias: proof of Proposition 2.14

We start by proving a lemma.

Lemma B.1 (Perturbed, pathwise version)**.** *Consider $g \in C([0,T], \mathbb{R})$ an α-Hölder function with constant Λ and set $\delta := \|g\|_\infty$. If $2\delta \leq \max \phi - \min \phi$, then*

$$\|\bar\rho(\phi + g) - \bar\rho(\phi)\|_{\mathcal{H}} \leq L_k(P_1 \delta + P_2 \delta^2 + P_3 \delta^3) =: L_k P(\delta),$$

where

$$\begin{aligned}P_1 =& 1 + 4A_\phi C_T C^\epsilon_{p-1,p}(\phi), \\ P_2 =& 8C_T C^\epsilon_{p-1,p}(\phi) + 4pA_\phi (C_T C^\epsilon_{p-2,p}(\phi) + \tfrac{C_{p-3,\Lambda,\alpha,T}}{\mathrm{pers}^p_{p,\epsilon}(\phi)}), \\ P_3 =& 4p \left(C_T C^\epsilon_{p-2,p}(\phi) + \tfrac{C_{p-3,\Lambda,\alpha,T}}{\mathrm{pers}^p_{p,\epsilon}(\phi)} \right),\end{aligned}$$

and

$$C_T = \frac{\lceil T \rceil}{\lfloor T \rfloor - 2}, \qquad C^\epsilon_{p,p'}(\phi) = \frac{\mathrm{pers}^p_{p,\epsilon}(\phi)}{\mathrm{pers}^{p'}_{p',\epsilon}(\phi)}, \qquad A_\phi = \|\phi\|_\infty.$$

Proof. By the diagram stability theorem, $d_B(D(\phi + g), D(\phi)) \leq \|g\|_\infty \leq \delta$. The persistence of a point in $D(\phi)$ and $D(\phi + g)$ is bounded by $2A_\phi$ and $2A_{\phi+g} \leq 2(A_\phi + \delta)$ respectively. Using Proposition 2.4, we also bound $\mathrm{pers}^{p}_{p-1,\epsilon}(\phi+g) \leq \mathrm{pers}^{p-1}_{p-1,\epsilon}(\phi) + p\delta(\mathrm{pers}^{p-2}_{p-2,\epsilon}(\phi) + \mathrm{pers}^{p-2}_{p-2,\epsilon}(g))$. Using the uniform bound on persistence from Proposition 2.3, $\mathrm{pers}^{p-2}_{p-2,\epsilon}(g) \leq C_{p-3,\Lambda,\alpha,T}$. Finally, putting these together with Proposition 2.10, we obtain:

$$
\begin{aligned}
\|\bar\rho(\phi) - \bar\rho(\phi+g)\|_{\mathcal{H}} &\leq L_k \left(1 + 2pU \frac{\mathrm{pers}^{p-1}_{p-1,\epsilon}(\phi) + \mathrm{pers}^{p-1}_{p-1,\epsilon}(\phi+g)}{\mathrm{pers}^{p}_{p,\epsilon}(\phi)}\right) d_B(D(\phi), D(\phi+g)) \\
&\leq \delta L_k \left(1 + 4p(\|\phi\|_\infty + \delta) \tfrac{2\lceil T\rceil \mathrm{pers}^{p-1}_{p-1,\epsilon}(\phi|_{[c,c+1]}) + p\delta(\mathrm{pers}^{p-2}_{p-2,\epsilon}(\phi) + C_{p-3,\Lambda,\alpha,T})}{(\lfloor T\rfloor - 2)\mathrm{pers}^{p}_{p,\epsilon}(\phi)}\right) \\
&\leq \delta L_k \Big[\big(1 + 4A_\phi C_T C^{\epsilon}_{p-1,p}(\phi)\big) + \\
&\qquad \left(8C_T C^{\epsilon}_{p-1,p}(\phi) + 4pA_\phi(C_T C^{\epsilon}_{p-2,p}(\phi) + \tfrac{C_{p-3,\Lambda,\alpha,T}}{\mathrm{pers}^{p}_{p,\epsilon}(\phi)})\right)\delta^1 + \\
&\qquad 4p\left(C_T C^{\epsilon}_{p-2,p}(\phi) + \tfrac{C_{p-3,\Lambda,\alpha,T}}{\mathrm{pers}^{p}_{p,\epsilon}(\phi)}\right)\delta^2\Big].
\end{aligned}
$$

□

Proof of Proposition 2.14. Combining lemma B.1 and theorem 2.13,

$$
\begin{aligned}
\|\bar\rho(\phi\circ\gamma_1 + g_1) - \bar\rho(\phi\circ\gamma_2 + g_2)\|_{\mathcal{H}} \leq\;& \|\bar\rho(\phi + (g_1)_{\gamma_1^{-1}}) - \bar\rho(\phi|_{[0,R_1]})\|_{\mathcal{H}} \\
&+ \|\bar\rho(\phi|_{[0,R_1]}) - \bar\rho(\phi|_{[0,R_2]})\|_{\mathcal{H}} \\
&+ \|\bar\rho(\phi|_{[0,R_2]}) - \bar\rho(\phi + (g_2)_{\gamma_2^{-2}})\|_{\mathcal{H}} \\
\leq\;& L_k(P(\delta_1) + P(\delta_2) + 2\tfrac{4}{\min(R_1,R_2)}\|\bar\rho(\phi|_{[c,c+1]})\|_{\mathcal{H}}) \\
\leq\;& L_k\left(P(\delta_1) + P(\delta_2) + \tfrac{8}{\min(R_1,R_2)-2}\tfrac{A_\phi}{2}\right) \\
\leq\;& L_k\left(P(\max(\delta_1,\delta_2)) + \tfrac{4A_\phi}{\min(R_1,R_2)-2}\right).
\end{aligned}
$$

□

C Measurability of functionals: proof of Proposition 3.2

Since $\bar\rho_u$ is applied pathwise, it is not obvious under what conditions $\bar\rho_u(S)$ is a $\mathbb{R}$-valued random variable and it is even less whether $\bar\rho(S)$ is a $C(\mathbb{U},\mathbb{R})$-valued random variable. Such considerations could be circumvented by using outer probabilities (Radulović, 1996; Kosorok, 2008), but we address them in Proposition 3.2. For the proof, we will need the following lemma.

Lemma C.1 (Pettis' measurability theorem)**.** *Consider $h:\Omega\to E$, where (E,d_E) is a Banach space. If E is separable as a metric space and h is weakly-measurable, then h is measurable with respect to the Borel σ-algebra induced by d_E.*

Proof (proposition 3.2). First, assume that S is weakly-measurable on $E = C([0,T],\mathbb{R})$ and that $(C([0,T],\mathbb{R}), \|\cdot\|_\infty)$ is separable. Using lemma C.1, we get that S is $\sigma(\|\cdot\|_\infty)$-measurable. Because $\bar\rho : C_{\epsilon+q}([0,T],\mathbb{R}) \to C(\mathbb{U},\mathbb{R})$ is continuous, it is measurable for the two σ-algebra on the domain and co-domain. This allows us to conclude that $\bar\rho(S)$ is $(C(\mathbb{U},\mathbb{R}), \sigma(\|\cdot\|_\infty))$-measurable.

Let us now verify the assumptions of Lemma C.1. We introduce the notation for the measurable spaces on which γ and W are defined,

$$
\gamma : (\Omega_r, \mathcal{A}_r) \to (\Gamma_{v_{\min}}, \sigma(\|\cdot\|_\infty)), \qquad W : (\Omega_n, \mathcal{A}_n) \to \left(C_{A_\phi - (\epsilon+q)}([0,T],\mathbb{R}), \sigma(\mathbb{R}^{[0,T]})\right),
$$

By continuity of ϕ, the composition $\phi\circ\gamma$ is $\sigma(\mathbb{R}^{[0,T]})$-measurable. As a sum of two (independent) random variables, $S = \phi\circ\gamma + W$ is $\sigma(\mathbb{R}^{[0,T]})$-measurable for $(\Omega,\mathcal{A})$, where $\Omega = \Omega_r\times\Omega_n$ and $\mathcal{A} = \mathcal{A}_r\otimes\mathcal{A}_n$. The product σ-algebra $\sigma(\mathbb{R}^{[0,T]})$ coincides with that of weak measurability on $\mathbb{R}^{[0,T]}$. The space $C([0,T],\mathbb{R})$ with the topology induced by $\|f\|_\infty := \sup_{x\in[0,T]}|f(x)|$ is a Banach, separable space. Any subspace of a separable metric space is separable, so $S(\Omega)$ is also separable. □

Remark C.2. *Lemma C.1 and its application to prove the measurability of the process were taken from the course Steinwart (2022).*

D Complements for the robustness of the signature in the additive noise setting

D.1 Proof of Theorem 3.4

We start by treating S path–wise. Using Proposition 2.10 and the bottleneck stability of persistence diagrams,

$$\begin{aligned}\|\bar{\rho}(\phi\circ\gamma_1+W)-\bar{\rho}(\phi\circ\gamma_2+W)\|_{\mathcal{H}} &= \|\bar{\rho}(\phi+W_{\gamma_1^{-1}})-\bar{\rho}(\phi+W_{\gamma_1^{-1}})\|_{\mathcal{H}}\\ &\leq L_k\left(1+4pU\max_{k=1,2}\frac{\mathrm{pers}_{p-1,\epsilon}^{p-1}(\phi+W_{\gamma_k^{-1}})}{\mathrm{pers}_{p,\epsilon}^{p}(\phi+W_{\gamma_k^{-1}})}\right)\|W_{\gamma_1^{-1}}-W_{\gamma_2^{-1}}\|_\infty,\end{aligned} \tag{D.3}$$

where L_k is a regularity constant of the kernel and U is an upper-bound on the persistence of any point in both diagrams. The persistence of any point in the diagram $D(h)$ of a function h is bounded by A_h. Hence, the persistence of a point in $D(\phi+W)$ is bounded by $U=A_{\phi+W}\leq A_\phi+A_W\leq A_\phi+(A_\phi-\epsilon-q)\leq 2A_\phi$.

Next, we obtain an upper–bound of $\max_{k=1,2}\mathrm{pers}_{p-1,\epsilon}^{p-1}(\phi+W_{\gamma_k^{-1}})/\mathrm{pers}_{p,\epsilon}^{p}(\phi+W_{\gamma_k^{-1}})$. By Proposition J.1, we can assume that W has α-Hölder paths with a (random) constant Λ_W, for $\alpha:=\min(1,r_1-1)/r_2$. This implies that $1+1/\alpha<p$ and we use the continuity of truncated persistence from Proposition 2.4 to obtain

$$\mathrm{pers}_{p-1,\epsilon}^{p-1}(\phi+W_{\gamma_k^{-1}})\leq\mathrm{pers}_{p-1,\epsilon}^{p-1}(\phi|_{[0,T]})+(p-1)\|W\|_\infty(\mathrm{pers}_{p-2,\epsilon}^{p-2}(\phi|_{[0,T]})+\mathrm{pers}_{p-2,\epsilon}^{p-2}(W_{\gamma_k^{-1}})). \tag{D.4}$$

For any $x\in[0,1]$ and $p\geq 0$, the function $p\mapsto x^p$ is decreasing, so that

$$\begin{aligned}\mathrm{pers}_{p-1,\epsilon}^{p-1}(\phi|_{[0,T]}) &= (A_\phi-\epsilon)^{p-1}\sum_{(b,d)\in D}\max\left(\tfrac{d-b-\epsilon}{A_\phi-\epsilon},0\right)^{p-1}\\ &\leq (A_\phi-\epsilon)^{p-1}\sum_{(b,d)\in D}\max\left(\tfrac{d-b-\epsilon}{A_\phi-\epsilon},0\right)^{p-2}\\ &= (A_\phi-\epsilon)\mathrm{pers}_{p-2,\epsilon}^{p-2}(\phi).\end{aligned}$$

Since $\|W\|_\infty<(A_\phi-\epsilon)/2$ and the persistence does not depend on the parametrisation , equation (D.4) becomes

$$\begin{aligned}\mathrm{pers}_{p-1,\epsilon}^{p-1}(\phi+W_{\gamma_k^{-1}}) &\leq (A_\phi-\epsilon)\mathrm{pers}_{p-2,\epsilon}^{p-2}(\phi)\left(1+\tfrac{p-1}{2}\left(1+\frac{\mathrm{pers}_{p-2,\epsilon}^{p-2}(W)}{\mathrm{pers}_{p-2,\epsilon}^{p-2}(\phi)}\right)\right)\\ &\leq p(A_\phi-\epsilon)\mathrm{pers}_{p-2,\epsilon}^{p-2}(\phi)\left(1+\tfrac{1}{2}\frac{\mathrm{pers}_{p-2,\epsilon}^{p-2}(W)}{\mathrm{pers}_{p-2,\epsilon}^{p-2}(\phi)}\right).\end{aligned}$$

An upper–bound for the persistence of W is given in Proposition 2.3

$$\mathrm{pers}_{p,\epsilon}^{p}(W)\leq(A_W-\epsilon)^p\left(1+pT\left(\tfrac{2\Lambda_W}{\epsilon}\right)^{1/\alpha}\right),$$

where Λ_W is the path–wise Hölder constant of W. The amplitude A_ϕ upper–bounds the persistence of a point and it is also realized as the persistence of a pair of a global minimum and a global maximum, so $\mathrm{pers}_{p-2,\epsilon}^{p-2}(\phi|_{[0,R]})\geq(R-2)(A_\phi-\epsilon)^{p-2}$ and hence

$$\frac{\mathrm{pers}_{p,\epsilon}^{p}(W)}{\mathrm{pers}_{p-2,\epsilon}^{p-2}(\phi)}\leq\left(\tfrac{A_W-\epsilon}{A_\phi-\epsilon}\right)^{p-2}(A_W-\epsilon)^2\tfrac{T}{R-2}\left(1+p\left(\tfrac{2\Lambda_W}{\epsilon}\right)^{1/\alpha}\right).$$

Putting the above together, with $p\geq 2$,

$$\begin{aligned}\mathrm{pers}_{p-1,\epsilon}^{p-1}(\phi+W_{\gamma_k^{-1}}) &\leq p(A_\phi-\epsilon)\mathrm{pers}_{p-2,\epsilon}^{p-2}(\phi)\\ &\quad\times\left(1+\left(\tfrac{A_W-\epsilon}{A_\phi-\epsilon}\right)^{p-2}(A_W-\epsilon)^2\tfrac{T}{R-2}\max\left(1,p\left(\tfrac{2\Lambda_W}{\epsilon}\right)^{1/\alpha}\right)\right).\end{aligned}$$

We have therefore an upper–bound for the numerator. To lower–bound the denominator, we use Proposition 2.3:

$$\begin{aligned}\mathrm{pers}_{p,\epsilon}^{p}(\phi+W_{\gamma_k^{-1}}) &\geq \mathrm{pers}_{p,\epsilon+A_W}^{p}(\phi)\\ &\geq (R-2)(A_\phi-(A_W+\epsilon))^p\\ &\geq (R-2)(A_\phi-(A_\phi-\epsilon+q+\epsilon))^p=(R-2)q^p.\end{aligned}$$

We conclude that C_{Λ_W} upper-bounds $\max_k \mathrm{pers}_{p-1,\epsilon}^{p-1}(\phi + W_{\gamma_k^{-1}})/\mathrm{pers}_{p,\epsilon}^{p}(\phi + W_{\gamma_k^{-1}})$,

$$C_{\Lambda_W} \coloneqq L_k \left(1 + \tfrac{8p^2 A_\phi}{(R-2)q^p}(A_\phi - \epsilon)\mathrm{pers}_{p-2,\epsilon}^{p-2}(\phi)\right.$$
$$\left. \times \left(1 + \left(\tfrac{A_W-\epsilon}{A_\phi-\epsilon}\right)^{p-2} (A_W - \epsilon)^2 \tfrac{T}{R-2} \max\left(1, p\left(\tfrac{2\Lambda_W}{\epsilon}\right)^{1/\alpha}\right)\right)\right).$$

As $A_W \leq A_\phi - \epsilon - q$, the only remaining stochastic term in C_{Λ_W} is $\Lambda_W^{1/\alpha}$. Also, the bound only depends on R (which is fixed), but not on γ itself.

Let $\pi : \mathcal{A}_{r,1} \times \mathcal{A}_{r,2} \to \mathbb{R}$ be a coupling of μ_1 and μ_2. Specifically, π is a measure on the product space $(\mathcal{G} \times \mathcal{G}, \mathcal{A}_{r,1} \otimes \mathcal{A}_{r,2})$, such that $\pi(A, \mathcal{G}) = \mu_1(A)$ and $\pi(\mathcal{G}, A) = \mu_2(A)$, for all $A \in \mathcal{A}$. Then, $\pi \otimes \nu : ((A_1, B_1), (A_2, B_2)) \mapsto \pi(A_1, A_2)\nu(B_1 \cap B_2)$ is a coupling of $\mu_1 \otimes \nu$ and $\mu_2 \otimes \nu$. Using the coupling and (D.3),

$$\begin{aligned}
\|\mathbb{E}[\bar\rho(\phi \circ \gamma_1 + W) \mid W] - \mathbb{E}[\bar\rho(\phi \circ \gamma_2 + W) \mid W]\|_{\mathcal{H}} &= \|\mathbb{E}_\pi[\bar\rho(\phi \circ \gamma_1 + W) - \bar\rho(\phi \circ \gamma_2 + W) \mid W]\|_{\mathcal{H}} \\
&\leq \mathbb{E}_\pi\left[\|\bar\rho(\phi \circ \gamma_1 + W) - \bar\rho(\phi \circ \gamma_2 + W)\|_{\mathcal{H}} \mid W\right] \\
&\leq C_{\Lambda_W}\mathbb{E}[\|W_{\gamma_1^{-1}} - W_{\gamma_2^{-1}}\|_\infty \mid W], \\
&\leq C_{\Lambda_W}\Lambda_W \mathbb{E}[\|\gamma_1^{-1} - \gamma_2^{-1}\|_\infty^\alpha].
\end{aligned}$$

We have thus completely separated the bound into a product, with terms depending on ν and (μ_1, μ_2).

On one hand, it remains to take the expectation with respect to W. We bound the moments of Λ_W using Theorem J.2, obtaining

$$\mathbb{E}[\Lambda_W] \leq 16\tfrac{\alpha+1}{\alpha}(K_{r_2,r_1})^{1/r_2}$$
$$\mathbb{E}[\Lambda_W^{1+1/\alpha}] \leq 6^{r_2+2} K_{r_2,r_1}^{(1/r_2+1/(r_1-1))}.$$

On the other hand, by Jensens' inequality, $\mathbb{E}[\|\gamma_1^{-1} - \gamma_2^{-1}\|_\infty^\alpha] \leq \mathbb{E}[\|\gamma_1^{-1} - \gamma_2^{-1}\|_\infty]^\alpha$. Using the lower–bound on the modulus of continuity,

$$\sup_{r \in [0,R]} |\gamma_1^{-1}(r) - \gamma_2^{-1}(r)| = \sup_{t \in [0,T]} |t - \gamma_2^{-1}(\gamma_1(t))| \leq \sup_{t \in [0,T]} \frac{1}{v_{\min}} |\gamma_2(t) - \gamma_1(t)|.$$

Taking the infimum over couplings, we obtain the $1-$Wasserstein distance $W_1(\mu_1, \mu_2)$.

E Stationary regimes for the time series models

E.1 Model 1 is stationary

In this section, we check that $(\mathrm{frac}(\gamma_n))_{n\in\mathbb{N}}$ is stationary. It is sufficient to show that for any $K \geq 1$, $(\mathrm{frac}(\gamma_0), \ldots, \mathrm{frac}(\gamma_K)) \sim (\mathrm{frac}(\gamma_n), \ldots \mathrm{frac}(\gamma_{n+K}))$, for any $n \geq 0$. We write

$$(\mathrm{frac}(\gamma_n), \ldots \mathrm{frac}(\gamma_{n+K})) = \mathrm{frac}(\mathrm{frac}(\gamma_0 + \sum_{r=0}^{n-1} V_r) + \mathrm{frac}(0, V_n, \ldots, \sum_{r=n}^{n+K-1} V_r)),$$

and we analyze the two terms separately. Here, frac is applied component–wise. First, because $(V_n)_{n\in\mathbb{N}}$ are i.i.d, $\left(\sum_{r=0}^{k} V_r\right) \sim \left(\sum_{r=n}^{n+k} V_r\right)$, for any $n, k \in \mathbb{N}$. Therefore, $(0, V_0, \ldots, \sum_{r=0}^{n-1} V_r) \sim (0, V_n, \ldots, \sum_{r=n}^{n+K-1} V_r)$. It also remains true when we apply frac component–wise, because it is a measurable mapping $\mathbb{R}^{K+1} \to \mathbb{R}^{K+1}$. Second, we claim the following lemma on the sum of two random variables, one of which is uniform.

Lemma E.1. *If $U \sim \mathcal{U}([0,1])$ and Z is a real–valued random variable independent of U, then* $\mathrm{frac}(U + Z) \sim \mathrm{frac}(U) \sim U$.

Before showing Lemma E.1, we conclude the proof by applying it to $U = \gamma_0$ and $Z = \sum_{r=0}^{n-1} V_r$. Indeed, γ_0 is independent from $(V_r)_{r=0}^{n-1}$, so we obtain that $\mathrm{frac}(\gamma_0) \sim \mathrm{frac}(\gamma_0 + \sum_{r=0}^{n-1} V_r)$. Finally, combining the above with $\mathrm{frac}((0, V_0, \ldots, \sum_{r=0}^{n-1} V_r)) \sim \mathrm{frac}((0, V_n, \ldots, \sum_{r=n}^{n+K-1} V_r))$, we have that $\mathrm{frac}(\gamma_0, \ldots, \gamma_K) \sim \mathrm{frac}(\gamma_n, \ldots, \gamma_{n+K})$.

Proof of Lemma E.1. First, it is clear that for $s \leq 0$, $P(\text{frac}(U+Z) < s) = 0$ and that for $s > 1$, $1 \geq P(\text{frac}(U+Z) < s) \geq P(\text{frac}(U+Z) \leq 1) = 1$. For $0 < s < 1$,

$$P(\text{frac}(U+Z) \leq s) = P\left(U + Z \in \bigcup_{k\in\mathbb{Z}} [k, k+s]\right) = \sum_{k\in\mathbb{Z}} P(U+Z \in [k, k+s]). \tag{E.5}$$

Because U and Z are independent, $P(U + Z \in [k, k+s]) = (\mu_U \star \mu_Z)([k, k+s])$, where μ_U and μ_Z are the probability measures of U and Z respectively and $\star$ denotes their convolution. Note that since μ is translation–invariant,

$$\begin{aligned}
(\mu_U \star \mu_Z)([k, k+s]) &= \int_{\mathbb{R}} \int_0^1 1_{[k,k+s]}(z+u) du d\mu_Z(z) \\
&= \int_{\mathbb{R}} \mu([0,1] \cap [k-z, k+s-z]) d\mu_Z(z) \\
&= \int_{\mathbb{R}} \mu([-k, -k+1] \cap [-z, -z+s]) d\mu_Z(z) \\
&= \int_{\mathbb{R}} \mu([-k, -k+1[\cap [-z, -z+s]) d\mu_Z(z)
\end{aligned}$$

Going back to (E.5),

$$\begin{aligned}
P(\text{frac}(U+Z) \leq s) &= \sum_{k\in\mathbb{Z}} \int_{\mathbb{R}} \mu([-k, -k+1[\cap [-z, -z+s]) d\mu_Z(z) \\
&= \int_{\mathbb{R}} \sum_{k\in\mathbb{Z}} \mu([-k, -k+1[\cap [-z, -z+s]) d\mu_Z(z) \\
&= \int_{\mathbb{R}} \mu([-z, -z+s]) d\mu_Z(z) \\
&= \mu([0, s]) \int_{\mathbb{R}} d\mu_Z(z). \\
&= s.
\end{aligned}$$

Therefore, the distribution function of $\text{frac}(U+Z)$ is uniform on $[0, 1]$ and therefore also equal to that of $\text{frac}(U)$. □

E.2 Stationary distribution for Model 2

The process $(\text{frac}\gamma_n)_{n\in\mathbb{N}}$ is defined in (18), via the Markov chain $(V_n)_{n\in\mathbb{N}}$. Recall that this Markov chain has a transition probability kernel P, with support included in $I = [v_{\min}, v_{\max}]$. The time series $(\text{frac}(\gamma_n))_{n\in\mathbb{N}}$ is not itself a Markov Chain but it can be easily checked that $(\text{frac}(\gamma_n), V_n)_{n\geq 1}$ is. Let $\tilde{P}$ be the transition kernel of this Markov Chain and let $\mathcal{U}$ be the uniform measure on $[0, 1] \times I$. By applying a similar proof as the proof of Doeblin’s Condition for Model 2, see Section F.2 step 8, it can be shown that for any $x \in [0, 1] \times I$, whenever $A \in \mathcal{B}([0, 1] \times I)$ is such that $\mathcal{U}(A) > 0$, then $P^n(x, A) > 0$ for some n sufficiently big. In other terms, the Markov Chain $(\text{frac}\gamma_n, V_n)_{n\geq 0}$ is $\mathcal{U}$-irreducible.

By Proposition 4.2.2 Meyn and Tweedie (2012), there is a maximal measure, Q, for which $(\text{frac}\gamma_n, V_n)_{n\geq 0}$ is Q-irreducible. By Theorem 8.3.4 in Meyn and Tweedie (2012), such a maximally-irreducible chain is either transient or recurrent. The fact that we can reach any element from the cover from any other one , as well as the compacity of the domain $[0, 1]$ prevents transience, so we can conclude it is recurrent. A recurrent chain admits an invariant measure, by Theorem 10.0.1 Meyn and Tweedie (2012).

F Exponential mixing of the reparametrisation process: proof of Proposition 3.10

In this section, we show that the reparametrisation processes defined in Section 3.3.1 is exponentially mixing.

The proof of Proposition 3.10 relies on the continuity of the transition kernel with respect to the Lebesgue measure and the use of the following sufficient condition.

Theorem F.1 (Doukhan (1995, Section 2.4, Theorem 1))**.** *Let $(Z_n)_n$ be a Markov chain and ν a non-negative and non-zero measure. If there exists $r \in \mathbb{N}^*$ such that*

$$P(Z_r \in A \mid Z_0 = z) \geq \nu(A), \qquad \textit{for any } z, \textit{ and } A \textit{ any P-measurable set,} \tag{F.6}$$

then $(Z_n)_n$ is ϕ-mixing and the coefficients decay exponentially fast, whatever the initial measure.

Condition (F.6) is called a Doeblin condition. It consists in providing a non-trivial lower–bound on the family of measures $(P^r(z, A))_z$. We first treat the case where $(V_n)_{n\in\mathbb{N}}$ are all i.i.d. The case where $(V_n)_{n\in\mathbb{N}}$ is a Markov Chain is similar, but technically more difficult.

F.1 Proof of Proposition 3.10 for Model 1

Recall that $\gamma_n = \gamma_{n-1} + V_{n-1}$. In Model 1, V_n is independent from $(V_k)_{k<n}$ and γ_0, so $(\gamma_n)_{n\in\mathbb{N}}$ is a Markov chain. We will now verify (F.6). Let $r := \lceil 2/(b-a) \rceil$ and $\epsilon = \lfloor (b-a)/r \rfloor$.

Lemma F.2. *Consider two measures μ_1, μ_2 such that $\mu_k(A) \geq c_k\mu(A)$, for $A \in \mathcal{B}([a_k, b_k])$. Then, for any $0 < \epsilon < \min(b_1 - a_1, b_2 - a_2)$, we have that $(\mu_1 \star \mu_2)(A) \geq c_1c_2\epsilon\mu(A)$, for any $A \in \mathcal{B}([a_1 + a_2 + \epsilon, b_1 + b_2 - \epsilon])$.*

We now apply this Lemma F.2 inductively to μ_1 and μ_2 the measures of $\sum_{n=1}^{r_1} V_{r_1}$ and V_{r_1+1} respectively, for $1 \leq r_1 \leq r-1$. We conclude that $P(\sum_{n=1}^{r} V_n \in A) \geq c\mu(A)$ for all $A \in \mathcal{B}(B)$, where $B := [r(a+\epsilon) - \epsilon, r(b-\epsilon) + \epsilon]$ and $c := c_1c_2\epsilon^{r-1}$. Thanks to our choice of r and ϵ, B is an interval of length at least 1.

Let $x_0 \in [0, 1[$ and $A \in \mathcal{B}([0, 1])$. We write $\mathrm{frac}^{-1}(A) = \cup_{k\in\mathbb{Z}} A + k$, where $A + k = \{a + k \mid a \in A\}$. Then,

$$\begin{aligned}
P(\mathrm{frac}(\gamma_r) \in A \mid \gamma_0 = x_0) &= P\left(x_0 + \sum_{n=0}^{r} V_n \in \mathrm{frac}^{-1}(A)\right) \\
&= P\left(\sum_{n=0}^{r} V_n \in \bigcup_{k\in\mathbb{Z}} (A+k) - x_0\right) \\
&\geq P\left(\sum_{n=0}^{r} V_n \in \bigcup_{k\in\mathbb{Z}} (A+k-x_0) \cap B\right) \\
&\geq c\mu\left(\bigcup_{k\in\mathbb{Z}} (A+k-x_0) \cap B\right) \\
&= c\sum_k \mu(A+k-x_0 \cap B),
\end{aligned}$$

where the last equality follows from the fact that $\mu(A \cap (A+1)) = 0$, because $A \subset [0, 1]$. Notice that for any set $A + z \cap B = (A \cap (B - z)) + z$ and that $\mu(A + z) = \mu(A)$, for any $z \in \mathbb{R}$. Hence, for any $k \in \mathbb{Z}$,

$$\mu\left(A + k - x_0 \cap B\right) = \mu(k - x_0 + (A \cap (B - k + x_0))) = \mu(A \cap (B - k - x_0)).$$

Recall that B is an interval of length greater than 1, so $(B - k - x_0)_{k\in\mathbb{Z}}$ is a cover of $\mathbb{R}$. Hence,

$$\begin{aligned}
P(\mathrm{frac}(\gamma_r) \in A \mid \gamma_0 = x_0) &\geq c\sum_k \mu(A \cap (B-k-x_0)) \\
&\geq c\mu\left(A \cap \bigcup_k (B-k-x_0)\right) \\
&= c\mu(A).
\end{aligned}$$

We can therefore set $\mu := c\mu$. The measure does not depend on x_0 and it has total mass $c > 0$.

F.2 Proof of Proposition 3.10 for Model 2

In this section we verify the Doeblin condition (F.6) for the Markov Chain $(\mathrm{frac}\gamma_n, V_n)_{n\in\mathbb{N}}$.

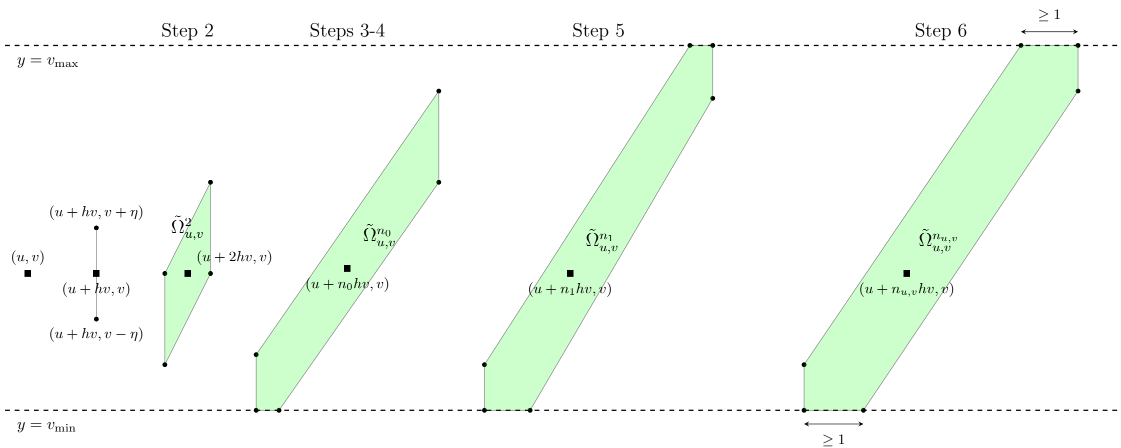

Figure 1: A schematic illustration of the form of a densitys' support. The density lower–bounds $\tilde{P}^n((u, v), \cdot))$.

Consider now $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ and let $(x, A) \mapsto 1_A(x)$, which is also a transition probability kernel. We define a product kernel on $R := \mathbb{R} \times I$, where $I = [v_{\min}, v_{\max}]$. It is characterised by the following measure on rectangles

$$((y, v), (A \times B)) \mapsto 1_A(y)P(v, B).$$

More generally, it extends to any set $A \in \mathcal{B}(R)$ as $((y, v), (A \times B)) \mapsto P(v, A_y)$, where

$$A_y = \{v \in I \mid (y, v) \in A\} \tag{F.7}$$

is the projection of $A \cap \{x = y\}$ onto the second coordinate. We define the map T

$$\begin{array}{rccc} T: & \mathbb{R}^2 & \to & \mathbb{R}^2 \\ & (x, v) & \mapsto & (x + hv, v), \end{array}$$

and we let $\tilde{P}$ be the pull-back of the product kernel P by this map. Explicitly, for $A \in \mathcal{B}(R)$,

$$\tilde{P}((u, v), A) = P(v, A_{u+hv}). \tag{F.8}$$

In what follows, we show (F.6) for the Markov chain $((\mathrm{frac}\gamma_n, V_n))_{n\in\mathbb{N}}$, which has transition probability kernel $\mathrm{frac}_\star\tilde{P}$. Figure 1 illustrates the proof. For $(u, v) \in R$, we show that $\tilde{P}^n((u, v), \cdot))$ is lower–bounded by a uniform measure of which we carefully characterise the support, $\Omega^n_{u,v}$. In Steps 1-6, we show that for a certain $n_{u,v} \in \mathbb{N}$, the support of this uniform measure, $\Omega^{n_{u,v}}_{u,v}$, is large enough. In 7, we show that $n_{u,v} \leq N \in \mathbb{N}$, for all $(u, v) \in R$. We conclude in 8 by showing (F.6). Compared with the i.i.d case treated in Section F.1, 1 is the analogue of Lemma F.2, except that the iteration requires the additional Steps 2-5.

Step 1 (lower–bound for $\tilde{P}^2((u, v), \cdot))$**.** *For* $(u, v) \in R$ *and* $(z_1, z_2) \in R$*, according to* (F.7)*,*

$$([0, z_1] \times [v_{\min}, z_2])_{u+h(v+y)} = \begin{cases} [v_{\min}, z_2], & \text{if } u + h(v + y) \in [0, z_1], \\ \emptyset, & \text{otherwise.} \end{cases}$$

In (F.8)*, we observe that integrating with respect to* $\tilde{P}^2$ *amounts to integrating* P *along a vertical strip, so marginalizing with respect to* (γ_1, V_1)*,*

$$\begin{aligned} \tilde{P}^2((u, v),]-\infty, z_1] \times [v_{\min}, z_2]) &= \int_R \tilde{P}((u, v), dxdy))\tilde{P}((x, y),]-\infty, z_1] \times [v_{\min}, z_2]) \\ &= \int_I P(v, dy)P(y, (]-\infty, z_1] \times [v_{\min}, z_2])_{u+h(v+y)}) \\ &= \int_{v_{\min}}^{\max((z_1-u)/h-v, v_{\max})} P(v, dy)P(y, [v_{\min}, z_2]) \end{aligned}$$

Differentiating the above expression with respect to z_1 and then z_2, for $z_1 \le u + h(v + v_{\max})$, we get

$$\frac{\partial \tilde{P}^2((u,v),]-\infty, z_1] \times [v_{\min}, z_2])}{\partial z_1} = f_v\left(\tfrac{z_1-u}{h} - v\right) P\left(\tfrac{z_1-u}{h} - v,]v_{\min}, z_2]\right)$$

$$f^{\star 2}_{(u,v)}(z_1, z_2) = \frac{\partial^2 \tilde{P}^2((u,v),]-\infty, z_1] \times [0, z_2])}{\partial z_1 \partial z_2} = f_v\left(\tfrac{z_1-u}{h} - v\right) f_{\frac{z_1-u}{h}-v}(z_2).$$

As $f_v(y) \ge \mu_0 1_{[v-\eta, v+\eta]}(y)$, we have $f^{\star 2}_{(u,v)}(z_1, z_2) \ge \mu_0^2$, if

$$\begin{cases} \dfrac{z_1 - u}{h} - v \in [\max(v_{\min}, v - \eta), \min(v_{\max}, v + \eta)], \\ \qquad z_2 \in \left[\max\left(v_{\min}, \dfrac{z_1-u}{h} - v - \eta\right), \min\left(v_{\max}, \dfrac{z_1-u}{h} - v + \eta\right)\right]. \end{cases}$$

The above is equivalent to

$$\begin{cases} z_1 = u + 2hv + kh\eta \\ z_2 = v + (k+l)\eta, \end{cases} \qquad \text{for some } l \in [-1,1], k \in [-1,1] \cap \left[\tfrac{v_{\min}-v}{\eta}, \tfrac{v_{\max}-v}{\eta}\right]. \tag{F.9}$$

So, $\tilde{P}^2((u,v), \cdot)$ has a density $f^{\star 2}_{(u,v)}$ with respect to the Borel measure on $\mathbb{R}^2$. That density is lower–bounded: for $(z_1, z_2) \in R \cap \Omega^2_{(u,v)}$, we have $f^{\star 2}_{(u,v)}(z_1, z_2) \ge \mu_0^2$, where

$$\Omega^2_{(u,v)} \coloneqq \{(u + 2hv, v) + k(h\eta, \eta) + l(0, \eta) \mid k, l \in [-1,1]\}. \tag{F.10}$$

When $(v_{\min} - v)/\eta < -1$ and $1 < (v_{\max} - v)/\eta$, then $\Omega^2_{(u,v)} \subset R$ and we carry on with the induction to 2. Otherwise, we go directly to 3 as $\Omega^{n+1} \cap R^c \neq \emptyset$.

Step 2 (Lower–bound for $n \ge 3$, while $\Omega^n_{(u,v)} \cap R^c = \emptyset$). *We start by defining the parallelograms $\Omega^n_{(u,v)}$ and showing some properties of the vectors that generate them. Then, by induction that for $n \ge 2$, we will show the following statement:*

> *For $0 < \epsilon < \min(1/4, \eta(v_{\max} - v_{\min})/2)$ and $\eta < (v_{\max} - v_{\min})/2$,*
> *$\tilde{P}^n$ has a density $f^{\star n}$ lower–bounded by $(\eta\epsilon/2)^{n-2} \mu_0^n$ on $\Omega^n_{(u,v)}$.* (F.11)

Our induction is valid while $\Omega^n_{u,v} \subset R$ and 3 shows how to modify it when it ceases to be the case. Our arguments become progressively more geometric, for what we find the illustration of the proof in Figure 2 helpful.

To define $\Omega^n_{(u,v)}$, let $v_2 \coloneqq T(0, \eta) = (h\eta, \eta)$ and for $n \ge 3$,

$$v_n = (1 - \epsilon)(T(0, \eta) + T(v_{n-1})) \in \mathbb{R}^2. \tag{F.12}$$

For $n \ge 3$, we define

$$\Omega^n_{(u,v)} \coloneqq \{T^n(u,v) + l(0,\eta) + kv_n \mid l, k \in [-1,1]\}. \tag{F.13}$$

Notice that if we take $n = 2$ in (F.13), *we get $\Omega^2_{(u,v)}$ from as defined in* (F.10).

While one can obtain an explicit expression of v_n, it is of little pratical interest: we only need to ensure that the horizontal component of v_n remains sufficiently large. This is detailed in the proof of Lemma F.3.

Since we have shown the statement (F.11) *for $n = 2$, we proceed with the induction step. For $(z_1, z_2) \in \Omega^{n+1}_{(u,v)} \cap R$, we calculate*

$$\begin{aligned} \tilde{P}^{n+1}((u,v),]-\infty, z_1] \times [v_{\min}, z_2]) &= \int_{R \cap \{x + yh \le z_1\}} \tilde{P}^n((u,v), dxdy) P(y, [v_{\min}, z_2]) \\ &= \int_{R \cap \{x + yh \le z_1\}} f^{\star n}_{(u,v)}(x, y) P(y, [v_{\min}, z_2]) dxdy. \end{aligned}$$

We can rewrite $R \cap \{x + yh \le z_1\} = \{(x,y) \in \mathbb{R} \mid y \in I,\ x \le z_1 - yh\}$. Differentiating with respect to z_1, we obtain

$$\frac{\partial \tilde{P}^{n+1}((u,v),]-\infty, z_1] \times [v_{\min}, z_2])}{\partial z_1} = \int_y f^{\star n}_{(u,v)}(z_1 - yh, y) P(y, [v_{\min}, z_2]) dxdy,$$

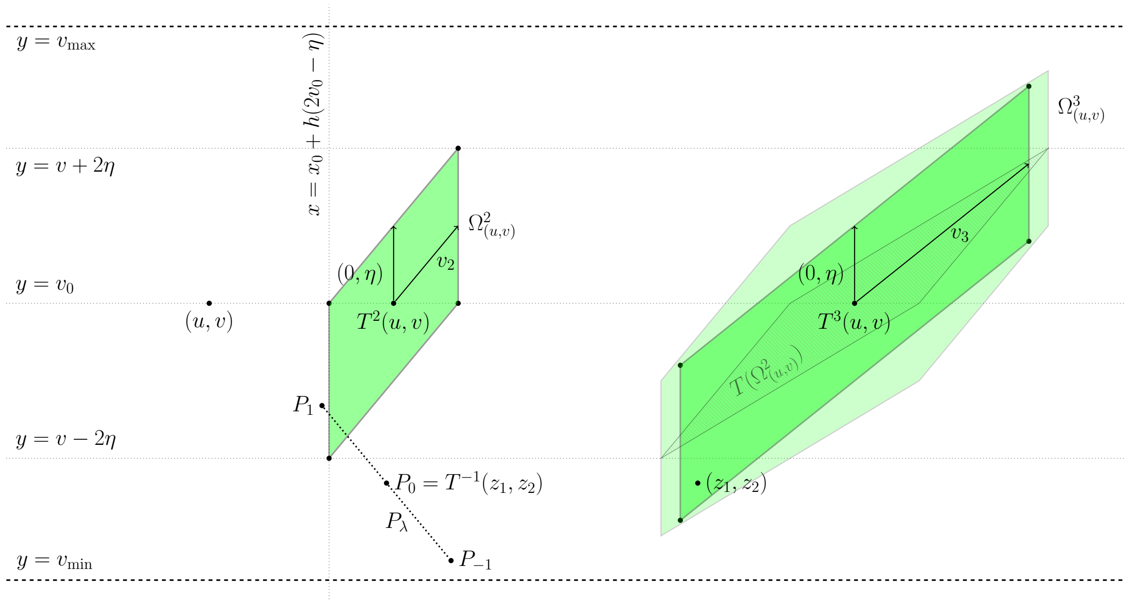

Figure 2: Illustration of $\Omega^n_{(u,v)}$ for $n = 2, 3$ and of the segment P_λ. Our argument consists in showing that for any $(z_1, z_2) \in \Omega^{n+1}_{(u,v)}$, the length of the intersection of P_λ with $\Omega^n_{(u,v)}$ is at least $\eta\epsilon/2$. While the dark green region is Ω^3, the lighter colour shows a larger region where the lower–bound is valid.

where f_v is defined in (19)*. For $z_2 \geq v_{\min}$, we get*

$$f^{\star n+1}_{(u,v)}(z_1, z_2) = \frac{\partial^2 \tilde{P}^{n+1}((u, v),]-\infty, z_1] \times [v_{\min}, z_2])}{\partial z_1 \partial z_2} = \int_I f^{\star n}_{(u,v)}(z_1 - yh, y) f_y(z_2) dy. \tag{F.14}$$

The expression in (F.14) *is similar to that from 1, except that it is integrated over I. We can lower–bound the integrand: $f^{\star n}_{(u,v)}$ is lower–bounded by $(\eta\epsilon/2)^{n-2}\,\mu_0^n$ on $\Omega^n_{(u,v)}$ for $n \geq 2$ and f_y by μ_0 on $[y - \eta, y + \eta] \cap I$. To lower-bound $f^{\star n+1}_{(u,v)}$, it remains to lower–bound the length of the integration domain. For the calculations, we take the following parametrisation of $[z_2 - \eta, z_2 + \eta] \cap I$,*

$$\left\{\lambda \in [-1, 1] \mid P_\lambda \coloneqq P_0 + \lambda\eta(-h, 1) \in \Omega^n_{(u,v)}\right\}, \qquad \textit{where } P_0 = T^{-1}(z_1, z_2). \tag{F.15}$$

Lemma F.3. *The length of the segment* (F.15) *is at least $\epsilon/2$.*

For the sake of readablity, we differ the proof of Lemma F.3 to Section F.2.1. Finally, going back to (F.14)*, we have the desired lower–bound*

$$f^{\star(n+1)}_{(u,v)}(z_1, z_2) \geq \left(\tfrac{\eta\epsilon}{2}\right)^{n-2} \mu_0^n \times \mu_0 \times \left(\tfrac{\eta\epsilon}{2}\right) = \left(\tfrac{\eta\epsilon}{2}\right)^{(n+1)-2} \mu_0^{n+1}, \qquad \textit{for } (z_1,\ z_2) \in \Omega^{n+1}_{u,v}.$$

In addition, for $\epsilon < 1\,/(1 + 3(v_{\max} - v_{\min})/2\eta)$,

$$(0, 1) \cdot (v_{n+1} - v_n) = (1 - \epsilon)\eta - \epsilon(0, 1) \cdot v_n > (1 - \epsilon)\eta - \epsilon(v_{\max} - v_{\min}) > \epsilon\tfrac{v_{\max} - v_{\min}}{2}.$$

The height of $\Omega^n_{u,v}$ grows with n, by at least a constant, positive term. Hence, it eventually reaches $v_{\max} - v_{\min}$, in which case $\Omega^n \cap R^c \neq \emptyset$.

Step 3 (First non–empty intersection with the boundary)**.** *Let $n_0 \coloneqq \min\{n \in \mathbb{N} \mid \mu(\Omega^n \cap R^c) > 0\}$. Without loss of generality, Ω^{n_0} extends beyond $v_{\min}$. We will now construct a region $\tilde{\Omega}^{n_0+1} \subset R$ such that $f^{\star(n_0+1)}_{(u,v)} \geq (\eta\epsilon/2)^{n_0-2}\,\mu_0^{n_0}$ on $\tilde{\Omega}^{n_0+1}$ and for which $\tilde{\Omega}^{n_0+1} \cap (\mathbb{R} \times \{v_{\min}\})$ is lower–bounded. Since we can choose η arbitrarily small, we can treat the lower and upper boundaries independently, so we focus on the construction of $\tilde{\Omega}^{n_0+1}_{u,v}$ on the boundary $\mathbb{R} \times \{v_{\min}\}$ first.*

For $P \in \mathbb{R} \times \{v_{\min}\}$, we consider P_λ as in (F.15)*, under the constraint that the integration segment lies within R, that is, $\{\lambda \in [0, 1] \mid P_\lambda \in \Omega^{n_0}\}$. We denote the length of this segment by $L(P)$*

$$L(P) \coloneqq |\{\lambda \in [-1, 1] \mid P + \eta\lambda\,(-h, 1) \in \Omega^{n_0} \cap R\}|,$$

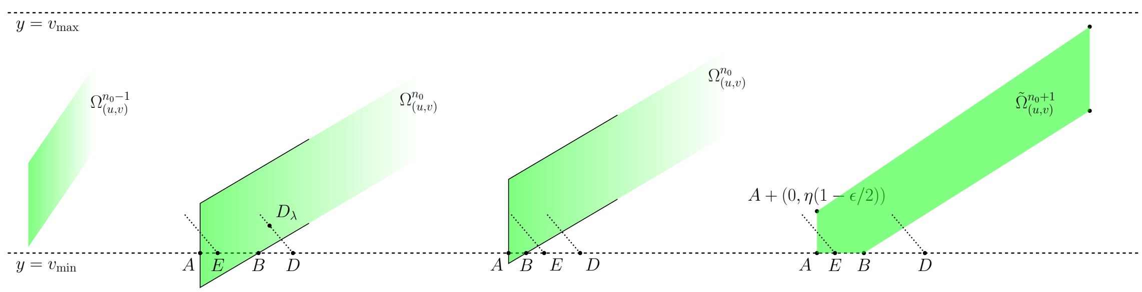

Figure 3: Illustration of 3 and the proof of (F.16). The leftmost polygon represents Ω^{n_0-1}, at the iteration before the first non-trivial intersection occurs. The two middle parallelograms illustrate the two cases, $x_E \leq x_B$ and $x_B \leq x_E$ respectively, from the proof of (F.16). On the right, the bottom part of the polygon $\tilde{\Omega}^{n_0+1}$ as constructed in 3. The dashed lines represent the integration segments, whose length is measured by L.

and we let A, B be the endpoints of $\Omega^{n_0} \cap (\mathbb{R} \times \{v_{\min}\})$. We rely on the following claim, whose proof is in Section F.2.2. Figure 3 illustrates the situation.

The set $\{P \in \Omega^{n_0} \cap \mathbb{R} \times \{v_{\min}\} \mid L(P) \geq \epsilon/2\}$ is not empty. If we denote by $D = A + (x_D, 0)$ and $E = A + (x_E, 0)$ its right- and left-endpoints, then for some $c_0 > 0$ independent of (u,v), we have (F.16)

$$x_B + c_0 \leq x_D - x_E.$$

In particular, $L(P) \geq \epsilon/2$ implies that $f^{\star n_0+1}_{(u,v)}(P) \geq (\eta\epsilon/2)^{n_0-1}\,\mu_0^{n_0+1}$. By convexity of $\Omega^{n_0} \cap R$, we have $L(P) \geq \epsilon/2$ for P on the segment $T(E)T(D)$, so we can include that segment in $\tilde{\Omega}^{n_0+1}$. As $L(E) = L(P)$, where $P = E + (1-\epsilon/2)k\eta\,(-h, 1)$, for $k \in [0,1]$, we have the same lower–bound on the density holds on $T(P)$. So, we can include a segment of height $\eta(1-\epsilon/2)$ above $T(E)$ in $\tilde{\Omega}^{n_0+1}$. Therefore, we can define $\tilde{\Omega}^{n_0+1}$ as the polygon with vertices $T(E) + (0, \eta(1-\epsilon/2))$, $T(E)$, $T(D)$, $T^{n_0+1}(u,v) - (0,\eta) + v_{n_0+1}$ and $T^{n_0+1}(u,v) + (0,\eta) + v_{n_0+1}$.

We have obtained a convex pentagon $\tilde{\Omega}^{n_0+1}$ on which $\tilde{P}^{n_0+1}((u,v),\cdot)$ is lower–bounded by a measure with density lower–bounded by $(\eta\epsilon/2)^{n_0-1}\,\mu_0^{n_0+1}$. Because T preserves lengths on horizontal cross-sections, (F.16) *implies that the length of $T(D)T(E)$ is equal to that of ED, which is longer by $c_0 = \eta h/4$ than the intersection at n_0.*

Step 4 (Induction for $n > n_0+1$)**.** *Assume that $\tilde{\Omega}^{n_0+1} \cap (\mathbb{R} \times \{v_{\max}\}) = \emptyset$. As a consequence of calculations for 2, $\tilde{\Omega}^n_{u,v}$ is growing upwards. Indeed, the calculations rely on Assumption* (19) *and the fact that v_n has a horizontal component whose length we control. Therefore, they adapt to $\tilde{\Omega}^n_{u,v}$, with v_n being the vector from $T(D)$ to $T^{n_0+1}(u,v) - (0,\eta) + v_{n_0+1}$.*

In addition, (F.16) *still holds. Indeed, redefine A, B, D and E, except with n_0 replaced by n_0+1 in the expression of L. We notice that A, B coincide with $T(E)$ and $T(D)$ from the previous iteration. Because AB is now of length at least $c_0 = h\eta/4$, the proof is easier as we fall in the first case. We define $\tilde{\Omega}^{n_0+2}$ as in 3.*

We can now iterate this procedure, obtaining a lower–bound of $f^{\star n}_{u,v}$ by a uniform constant, on a convex and polygonal domain $\tilde{\Omega}^n$. Crucially, both the height of $\tilde{\Omega}^n$ and the length of its intersection with $\mathbb{R} \times \{v_{\min}\}$ grow, by uniformly lower–bounded amounts.

Step 5 (Intersection with both boundaries)**.** *For some $n_1 \in \mathbb{N}$, the intersection $\tilde{\Omega}^{n_1}_{u,v} \cap (\mathbb{R} \cap \{v_{\max}\})$ is not trivial. By a procedure analogue to that in 3, we can define $\tilde{\Omega}^{n_1+1}$, which non–trivially intersects both boundaries. Using the procedure from 4, it is clear that the intersection will not only remain non–trivial with n, but also increase.*

Step 6 (Cross-sections with length at least 1)**.** *By definition, $\tilde{\Omega}^n$ is delimited by a convex, polygonal domain. The length of any horizontal cross-section of $\tilde{\Omega}^n$ is lower–bounded by the minimum of the lengths of the* *intersections with the lower and upper boundaries*[1]*. Recall that by 4, these two are increasing, and this by at least* $h\eta/4$ *at each iteration. Hence, for some* $n = n_{u,v}$*, all horizontal sections of* $\tilde{\Omega}_{u,v}^{n_{u,v}}$ *are of length at least 1.*

By construction of $\tilde{\Omega}_{(u,v)}^{n}$*, we have obtained a region such that for any* $n \geq n_{u,v}$*,*

1. $\tilde{P}^n((u,v),\cdot)$ *is lower–bounded by* $(\eta\epsilon/2)^{n-2}\,\mu_0^n\mu$ *on* $\tilde{\Omega}_{u,v}^n$*, (μ being the Lebesgue measure)*
2. $\{\tilde{\Omega}_{(u,v)}^n + (k,0)\}_{k\in\mathbb{Z}}$ *is a cover of* R*.*

Step 7 (Uniform lower–bound)**.** *We now show that we can choose a uniform* $N \in \mathbb{N}$*, such that* $n_{u,v} \leq N$ *for all* $(u,v) \in R$*. Fix* $(u,v) \in R$ *and let* $\bar{\Omega}_{u,v}^2$ *be defined as in* (F.9)*, except with* $\eta/2$ *instead of* η*. We can then perform 1 to 6, so we obtain a domain* $\bar{\Omega}_{u,v}^{\bar{n}_{u,v}}$ *with cross-sections of length at least 1, for some* $\bar{n}_{u,v} \geq n_{u,v}$*.*

Notice that the shrinked parallelogram at $n = 2$ *is contained in parallelograms for different initial conditions. Specifically, we have* $\bar{\Omega}_{u,v}^2 \subset \Omega_{x,y}^2$ *for* $(x,y) \in C_{u,v}$*, where* $C_{u,v} = T^{-2}(\bar{\Omega}_{u,v}^2)$*. In particular,* $n_{x,y} \leq \bar{n}_{u,v}$*, for all* $(x,y) \in C_{u,v}$*. Since* $(\overset{\circ}{C}_{u,v})_{(u,v)\in[0,1]\times I}$ *is an open cover of* $[0,1]\times I$*, by compactness, we can find a finite cover* $\{C_{u_k,v_k}\}_{k=1}^K$*. Clearly,* $N = \max_{1\leq k\leq K} \bar{n}_{u_k,v_k} < \infty$ *gives a uniform bound on* $(n_{u,v})_{(u,v)\in[0,1]\times I}$*. The bound is also valid on* $R \times I$*, because the whole construction is invariant with respect to horizontal translations.*

Finally, for $(u,v) \in R$*, we have that* $\tilde{P}^N((u,v),\cdot)$ *is lower–bounded by* $(\eta\epsilon/2)^{N-2}\,\mu_0^N\mu$*, on* $\Omega_{u,v}$ *and* $\{\Omega_{(u,v)} + (k,0)\}_{k\in\mathbb{Z}}$ *is a cover of* R*, where* $\Omega_{u,v} \coloneqq \tilde{\Omega}_{u,v}^N$*.*

Step 8 (Conclusion)**.** *We can now go back to* $(\mathrm{frac}\gamma_n, V_n)$*. By lower–bounding* $\tilde{P}^N$ *with a uniform measure, we can use the same arguments as in Section F.1 to conclude. For* $A \in \mathcal{B}([0,1]\times I)$*, we have*

$$\begin{aligned}
\mathrm{frac}_\star\tilde{P}^N((u,v),A) &= \tilde{P}^N((u,v),\mathrm{frac}^{-1}(A)) \\
&\geq \tilde{P}^N((u,v),\mathrm{frac}^{-1}(A)\cap\Omega_{(u,v)}) \\
&\geq C\mu(\mathrm{frac}^{-1}(A)\cap\Omega_{(u,v)}) && (\textit{minorating on } \Omega_{(u,v)}) \\
&= C\mu(\cup_{k\in Z} A+(k,0)\cap\Omega_{(u,v)}) && (\{A+(k,0)\}_k \textit{ disjoint}) \\
&= C\sum_{k\in Z}\mu(A+(k,0)\cap\Omega_{(u,v)}) \\
&= C\sum_{k\in Z}\mu(A\cap(\Omega_{(u,v)}-(k,0))) && (\mu \textit{ translation–invariant}) \\
&\geq C\mu(\cup_{k\in Z} A\cap(\Omega_{(u,v)}-(k,0))) \\
&= C\mu(A) && (\{\Omega_{(u,v)}+(k,0)\}_{k\in\mathbb{Z}} \textit{ is a cover of } R),
\end{aligned}$$

where $C = C_{\eta,\epsilon,\mu_0,N} \coloneqq (\eta\epsilon/2)^{N-2}\,\mu_0^N$*. The lower–bound is uniform in* (u,v) *and also shows that the measure is non-trivial. We conclude the proof of Proposition 3.10 by applying Theorem F.1.*

F.2.1 Proof of Lemma F.3

We recall that for some $l, k \in [-1,1]$,

$$(z_1,z_2) = T^{n+1}(u,v) + l\,(0,\eta) + kv_n,$$

where v_n is given in (F.12), so

$$P_\lambda \coloneqq T^{-1}(z_1, z_2+\lambda\eta) = T^n(u,v) + \eta(l+\lambda)(-h,1) + k(1-\epsilon)\left((0,\eta)+v_n\right). \tag{F.17}$$

For a parallelogram Ω generated by vectors x, y and centered around the origin, we have

$$P\in\Omega \iff \left\{\begin{array}{ccccc} x^Ty^\perp & \leq & P^Ty^\perp & \leq & -x^Ty^\perp \\ -y^Tx^\perp & \leq & P^Tx^\perp & \leq & y^Tx^\perp, \end{array}\right. \tag{F.18}$$

[1] To see this, consider the parallelogram on the 4 vertices of $\tilde{\Omega}^n$ which belong to the boundary. That parallelogram is included in $\tilde{\Omega}^n$ by convexity, so the lengths of the horizontal sections between the length of both bases.

where $(x_1, x_2)^\perp = (x_2, -x_1)$. Combining (F.17) with (F.18), we have that $P_\lambda \in \Omega^n_{(u,v)}$ if and only if

$$\left\{\begin{array}{c} \eta(0,1)\cdot v_n^\perp \leq \eta(l+\lambda)(-h,1)\cdot v_n^\perp + k(1-\epsilon)(\eta\,(0,1)+v_n)\cdot v_n^\perp \leq -\eta(0,1)\cdot v_n^\perp \\ -\eta v_n\cdot(0,1)^\perp \leq \eta^2(l+\lambda)(-h,1)\cdot(0,1)^\perp + k(1-\epsilon)(\eta^2\,(0,1)+\eta v_n)\cdot(0,1)^\perp \leq \eta v_n\cdot(0,1)^\perp \end{array}\right.$$
$$\iff \left\{\begin{array}{ccccc} (1,0)\cdot v_n(-1+k(1-\epsilon)) & \leq & -(l+\lambda)(1,h)\cdot v_n & \leq & (1,0)\cdot v_n(1+k(1-\epsilon)) \\ v_n\cdot(1,0)(-1-k(1-\epsilon)) & \leq & (-h\eta)(l+\lambda) & \leq & v_n\cdot(1,0)(1-k(1-\epsilon)). \end{array}\right.$$

As $(1,h)\cdot v_n > 0$ and denoting

$$a_n \coloneqq \frac{1}{h\eta}(1,0)\cdot v_n, \qquad b_n \coloneqq 1 - \frac{(0,h)\cdot v_n}{(1,h)\cdot v_n},$$

we have

$$P_\lambda \in \Omega^n \iff \left\{\begin{array}{ccccc} b_n(-1-k(1-\epsilon))-l & \leq & \lambda & \leq & b_n(1-k(1-\epsilon))-l \\ a_n(-1+k(1-\epsilon))-l & \leq & \lambda & \leq & a_n(1+k(1-\epsilon))-l. \end{array}\right.$$

Finally, taking into account that $\lambda \in [-1,1]$, we obtain that

$$\begin{aligned} \lambda \in [\max(-1, b_n(-1-k(1-\epsilon))-l, a_n(-1+k(1-\epsilon))-l),& \\ \min(1, b_n(1-k(1-\epsilon))-l, a_n(1+k(1-\epsilon))-l)],& \end{aligned}$$

which is of length

$$\begin{aligned} &\min(1, b_n(1-k(1-\epsilon))-l, a_n(1+k(1-\epsilon))-l)+ \\ &+\min(1, b_n(1+k(1-\epsilon))+l, a_n(1-k(1-\epsilon))+l) = \\ &= \min(2\min(1,a_n,b_n), (a_n+b_n)(1-k(1-\epsilon)), 1+b_n(1+k(1-\epsilon))+l, \\ &\qquad 1+a_n(1-k(1-\epsilon))+l, 1+b_n(1-k(1-\epsilon))-l, 1+a_n(1+k(1-\epsilon)-l) \end{aligned}$$

We claim that for $n \geq 2$,

$$a_n \geq 1, \qquad b_n \geq \frac{1}{2}. \tag{F.19}$$

Combining (F.19) with $l, k \in [-1,1]$, $0 < \epsilon \leq 1/2$, we conclude that the length of (F.15) is at least $\epsilon/2$.

It remains to show (F.19). For a_n, we proceed by induction. Using $v_2 = T\,(0,\eta)$ for $n=2$ and $v_3 = (1-\epsilon)\eta\,(3h,2)$, we verify that $a_2,\, a_3 \geq 1$. Notice that $(T\,(x,y))\cdot(1,0) = ((x,y)+(yh,0))\cdot(1,0) = (x,y)\cdot(1,h)$. Then,

$$v_{n+1}\cdot(1,0) = (1-\epsilon)(T\,(0,\eta)+T(v_n))\cdot(1,0) = (1-\epsilon)\left[h\eta + v_n\cdot(1,h)\right].$$

Using the induction hypothesis, $v_n\cdot(1,0) \geq h\eta$ combined with $v_n\cdot(0,h)\geq 0$,

$$v_{n+1}\cdot(1,0) \geq 2h\eta(1-\epsilon) \geq h\eta,$$

since $\epsilon \leq 1/2$.

For b_n, we can calculate directly $b_2 = (h\eta)(h\eta+h\eta) = 1/2$. For $n \geq 3$, we can express v_n using (F.12), so that

$$\begin{aligned} \frac{(1,0)\cdot v_n}{(1,h)\cdot v_n} =& \frac{(1,0)\cdot(T\,(0,\eta)+T(v_{n-1}))}{(1,h)\cdot(T\,(0,\eta)+T(v_{n-1}))} \\ =& \frac{h\eta+(1,0)\cdot T(v_{n-1})}{2h\eta+(1,h)\cdot T(v_{n-1})} \\ =& \frac{1}{2}+\frac{1}{2}\frac{((1,0)-(0,h))\cdot T(v_{n-1})}{2h\eta+(1,h)\cdot T(v_{n-1})} \\ \geq& \frac{1}{2}, \end{aligned}$$

where the last inequality follows from

$$((1,0)-(0,h))\cdot T(v_{n-1}) = (1,0)\cdot v_{n-1} + (0,h)\cdot v_{n-1} - (0,h)\cdot v_{n-1} \geq 0.$$

□

F.2.2 Proof of (F.16)

Notice first that $L(A) = 0$ and $L(P) = 0$ for any $P \in R \times \{v_{\min}\}$ to the left of A, so that $0 \leq x_B, x_D, x_E$. Second, consider $P = A + (x_B + (1-\epsilon/2)\eta h(1-b), 0)$. As $L(P) = \epsilon/2$, we have that $\{P \mid L(P) \geq \epsilon/2\} \neq \emptyset$, so D and E exist. In addition, we know that $x_D \geq x_B + (1-\epsilon/2)\eta h/(1-b)$. In particular, $b \leq 1$ implies that $x_B < x_D$.

Since $A \in \Omega^{n_0}_{u,v}$, we can write $A = T^{n_0}(u,v) + l_A \eta\,(0,\eta) - v_{n_0}$. By definition, n_0 is the first time such that $\Omega^n_0 \cap R^c$ has non-trivial measure, so, using the relation between Ω^{n_0-1} and Ω^{n_0}, we can conclude that $l_A \leq 0 \leq 1 - \epsilon/2$.

We distinguish two cases, depending which one of $\eta\epsilon h/2$ or x_B is greater. First, if $\eta h\epsilon/2 \leq x_B$, then $x_E = \eta h \epsilon/2$. Indeed, the triangle formed by A, E and $A + (0, \eta\epsilon/2)$ is in Ω^{n_0}, so $L\left(A + (\eta h\epsilon/2, 0)\right) \geq \epsilon/2$. Therefore,

$$\begin{aligned} x_D - x_E &\geq x_B + (1-\tfrac{\epsilon}{2})\eta h \tfrac{1}{1-b} - \eta h \tfrac{\epsilon}{2} \\ &\geq x_B + \eta h(2(1-\tfrac{\epsilon}{2}) - \tfrac{\epsilon}{2}) \\ &\geq x_B + \tfrac{5}{4}\eta h, \end{aligned}$$

where in the last two inequalities, we have used that $\epsilon \leq 1/2 \leq b$.

Next, if $x_B < \eta h\epsilon/2$, then $\eta h\epsilon/2 \leq x_E$. So, for $x \leq \eta h$,

$$L\left(A + (x, 0)\right) = \frac{1}{\eta h}\left(x - \tfrac{(0,h)\cdot v_n}{(1,h)\cdot v_n}\,(x - x_B)_+\right) = \frac{1}{\eta h}(xb - x_B(1-b)). \tag{F.20}$$

Notice that $L\left(A + (\eta h, 0)\right) \geq \epsilon/2$ for $\epsilon \leq 2/5$ small enough,

$$\begin{aligned} L\left(A + (\eta h, 0)\right) - \frac{\epsilon}{2} &= \frac{1}{\eta h}(\eta h b - x_B(1-b)) - \frac{\epsilon}{2} \\ &= b - (1-b)\frac{x_B}{\eta h} - \frac{\epsilon}{2} \\ &\geq b - (1-b)\frac{\epsilon}{2} - \frac{\epsilon}{2} \\ &= b\left(1 + \frac{\epsilon}{2}\right) - \frac{3}{2}\epsilon \\ &\geq \frac{1}{2}\left(1 - \frac{5}{2}\epsilon\right), \end{aligned}$$

so $x_E \leq \eta h$. Using (F.20), we find that $x_E = (\eta h\epsilon/2 + x_B(1-b))/b$. Finally,

$$\begin{aligned} x_D - x_E - x_B &= x_B\left(1 - \tfrac{1-b}{b}\right) + \eta h \tfrac{1-\epsilon/2}{1-b} - \tfrac{1}{2b}\eta h \epsilon - \eta h \tfrac{\epsilon}{2} \\ &= x_B(2 - \tfrac{1}{b}) + \eta h\left(\tfrac{1}{1-b} + \tfrac{\epsilon}{2}(\tfrac{1}{b} - \tfrac{1}{1-b})\right) - \eta h \tfrac{\epsilon}{2} \\ &= x_B(2 - \tfrac{1}{b}) + \tfrac{\eta h}{1-b}\left(1 - \tfrac{\epsilon}{b}(b - \tfrac{1}{2})\right) - \eta h \tfrac{\epsilon}{2}. \end{aligned}$$

Since $1/2 \leq b \leq 1$, we have $(b - 1/2)/(b) \leq 1$ and $1/(1-b) \geq 2$, so that

$$x_D - x_E - x_B \geq 2(1-\epsilon)\eta h - \eta h\epsilon/2 \geq \eta h(1 - \tfrac{3}{2}\epsilon).$$

Combining the two cases with $\epsilon < 1/2$, we conclude that

$$x_D - x_E \geq x_B + \eta h \min\left(\left(1 - \tfrac{\epsilon}{2}\right), \tfrac{5}{4}\right) \geq x_B + \tfrac{1}{4}\eta h.$$

□

G Mixing-preserving operations: mixing coefficients of $(\mathbf{X}_n)_{n\in\mathbb{N}}$

Proposition G.1. *Let $\mathbf{X}_n$ be as in* (20). *For any $k \in \mathbb{N}$,*

$$\beta_{\mathbf{X}}(k + M - 1) \leq \beta_X(k) \leq \beta_{\mathrm{frac}(\gamma)}(k) + \beta_W(k).$$

The proposition is a consequence of Lemmata G.2 and G.3, combined with the fact that ϕ is continuous, so $\beta_{\phi(\gamma)}(k) \leq \beta_{\mathrm{frac}(\gamma)}(k)$. The proofs of the lemmata essentially consist in manipulating the definitions.

Lemma G.2. *For two random variables $U : (\Omega_U, \sigma^U) \to \mathbb{R}$, $V : (\Omega_V, \sigma^V) \to \mathbb{R}$, we have*

$$\beta_{U+V}(k) \leq \beta_U(k) + \beta_V(k),$$

if $(\beta_U(k))_{k\in\mathbb{N}}$, $(\beta_V(k))_{k\in\mathbb{N}} \in \ell^1$. Moreover, the same holds true if U and V are defined on the same probability space, but are independent.

Proof. Define $Z := U + V$. Then, Z is (Ω_Z, σ^Z)-measurable, where $\Omega_Z = \Omega_U \times \Omega_V$ and $\sigma^Z = \sigma^U \otimes \sigma^V$. As σ^Z is generated by products of elements from σ^U and σ^V, we only need to consider (countable) partitions $\mathcal{A}_U, \mathcal{B}_{\mathcal{U}}$ and $\mathcal{A}_V, \mathcal{B}_V$ of $\sigma^U_{-\infty,0}, \sigma^U_{k,\infty}$ and $\sigma^V_{-\infty,0}, \sigma^V_{k,\infty}$ respectively. For any $A_U \in \mathcal{A}_U, A_V \in \mathcal{A}_V$ and $B_U \in \mathcal{B}_U, B_V \in \mathcal{B}_V$, by definition of the product probability measure,

$$\begin{aligned} P((A_U \times A_V) \cap (B_U \times B_V)) - P(A_U \times A_V) P(B_U \times B_V) = \\ = (P_U(A_U \cap B_U) - P_U(A_U)P_U(B_U))P_V(A_V \cap B_V) \\ + P_U(A_U)P_U(B_U)(P_V(A_V \cap B_V) - P_V(A_V)P_V(B_V)). \end{aligned}$$

Since β_U is is summable, $\sum_{A_U, B_U} |P_U(A_U \cap B_U) - P_U(A_U)P_U(B_U)| < \infty$ (idem for V), so we can regroup terms and

$$\begin{array}{rll} \sum_{\substack{A_U \in \mathcal{A}_U, A_V \in \mathcal{A}_V, \\ B_U \in \mathcal{B}_U, B_V \in \mathcal{B}_V}} & P((A_U \times A_V) \cap (B_U \times B_V)) - P(A_U \times A_V) P(B_U \times B_V) & \\ = & \sum_{A_U \in \mathcal{A}_U,\, B_U \in \mathcal{B}_U} (P_U(A_U \cap B_U) - P_U(A_U)P_U(B_U)) & \\ & \times \sum_{A_V \in \mathcal{A}_V,\, B_V \in \mathcal{B}_V} P_V(A_V \cap B_V) & (=1) \\ & + \sum_{A_U \in \mathcal{A}_U,\, B_U \in \mathcal{B}_U} P_U(A_U)P_U(B_U) & (=1) \\ & \times \sum_{A_V \in \mathcal{A}_V,\, B_V \in \mathcal{B}_V} (P_V(A_V \cap B_V) - P_V(A_V)P_V(B_V)) & \\ \leq & \beta_U(k) + \beta_V(k). & \end{array}$$

We conclude by taking the sup over partitions of Ω_Z. □

Lemma G.3. *Consider $(X_i)_{i\in\mathbb{N}}$ with coefficients $\beta_X(k)$ and define $\mathbf{X}_n = (X_n, \dots, X_{n+M-1})$. Then,*

$$\beta_{\mathbf{X}}(k + M - 1) \leq \beta_X(k).$$

Proof. First, note that the σ-algebra generated by a vector coincides with the σ-algebra generated by its components

$$\begin{aligned} \sigma(\mathbf{X}_{n_1}, \dots \mathbf{X}_{n_2}) &= \sigma((X_{n_1}, \dots, X_{n_1+M-1}), \dots, (X_{n_2}, \dots, X_{n_2+M-1})) \\ &= \sigma(X_{n_1}, \dots, X_{n_2+M-1}) \\ &= \sigma^X_{n_1, n_2+M-1}. \end{aligned}$$

Then, any partition $\mathcal{A} \subset \sigma^{\mathbf{X}}_{n_1,n_2}$ is also in $\sigma^X_{n_1,n_2+M-1}$. Since $\beta_{\mathbf{X}}$ is defined as a sup over such partitions, $\beta_{\mathbf{X}}(k + M - 1) \leq \beta_X(k)$. For $k \leq M$, we can take $\mathcal{A} = \mathcal{B} \subset \sigma(X_k)$. Since X_k is a continuous random variable, $\beta_{\mathbf{X}}(k) = 1$. □

H Gaussian approximation for dependent data

Theorem H.1 (Kosorok (2008, Theorem 11.24))**.** *Let $(X_n)_{n\in\mathbb{N}} \subset \mathbb{R}^d$ be a stationary sequence and consider a functional family $\mathcal{F} = (F_t)_{t\in\mathbb{U}}$ with finite bracketing entropy. Suppose there exists $r \in]2, \infty[$, such that*

$$\sum_{k=1}^{\infty} k^{\frac{2}{r-2}} \beta_X(k) < \infty, \tag{H.21}$$

Then, $\sqrt{N}(\hat{F}_t - F_t)$ converges to a tight, zero–mean Gaussian G process with covariance (23).

Theorem H.2 (Bühlmann (1995, Theorem1))**.** *Let $(X_n)_{n\in\mathbb{N}} \subset \mathbb{R}^d$ be a stationary sequence and consider a functional family $\mathcal{F} = (F_t)_{t\in\mathbb{U}}$ with finite bracketing entropy. Suppose that $\beta_X(k) \xrightarrow[k\to\infty]{} 0$ decrease exponentially and that $\mathcal{F}$ satisfies (4,6). Let the bootstrap sample be generated with the Moving Block Bootstrap, where the block size $L(n)$ satisfying $L(n) \to \infty$ and $L(n) = \mathcal{O}(n^{1/2-\epsilon})$ for some $0 < \epsilon < 1/2$. Then,*

$$\sqrt{N}(\hat{F}^*_N - \mathbb{E}^*[\hat{F}^*_N]) \to^* G \qquad \textit{in probability,}$$

where G is the zero-mean Gaussian Process with the covariance (23).

I Bracketing entropy for functionals

To show the convergence of functionals on random functions, we need to control the complexity of the family $\mathcal{F} := (\bar{\rho}_u)_{u\in K}$, where $\bar{\rho}_u : \mathbb{R}^M \to \mathbb{R}$. In particular, we need the bracketing entropy of $\mathcal{F}$ to be finite, which is a well-known result and a consequence of (4, 5).

Proposition I.1. *Let $N_{[]}(\epsilon, \mathcal{F}, \|\cdot\|)$ denote the bracketing number of $\mathcal{F}$, with brackets $[u, l]$ of size $\|u-l\| \le \epsilon$. Consider $\bar{\rho}$ as in* (3) *with k satisfying (4, 5). Then, for any probability measure P on $\mathbb{R}^M$ and $r \ge 1$,*

$$N_{[]}(\epsilon, \{\bar{\rho}_u\}_{u\in\mathbb{U}}, \|\cdot\|_{L_r(P)}) \le \frac{2^{D+1}L^D \operatorname{diam}(K)}{\epsilon^D},$$

where D is the dimension of $\mathbb{U}$. As a consequence, the bracketing entropy $J_{[]}(\infty, \mathcal{F}, \|\cdot\|_{L_r(P)})$ is finite

$$J_{[]}(\infty, \mathcal{F}, \|\cdot\|_{L_r(P)}) := \int_0^\infty \sqrt{\log N_{[]}(\epsilon, \mathcal{F}, \|\cdot\|_{L_r(P)})}d\epsilon < \infty.$$

Proof. First, since P is a probability measure, $\|\bar{\rho}_u\|_{L_r(P)} = \left(\int |\bar{\rho}_u|^r dP\right)^{1/r} \le \|\bar{\rho}_u\|_\infty \int dP = \|\bar{\rho}_u\|_\infty$, so $N_{[]}(\epsilon, \{\bar{\rho}_u\}_{u\in\mathbb{U}}, \|\cdot\|_{L_r(P)}) \le N_{[]}(\epsilon, \{\bar{\rho}_u\}_{u\in\mathbb{U}}, \|\cdot\|_\infty)$. Combining (5) with the fact that $\bar{\rho}(x)$ is a weighted average of k, for any $x \in \mathbb{R}^M$ and $s,\, t \in \mathbb{U}$, the normalized functional is L-Lipschitz in time

$$|\bar{\rho}_u(x) - \bar{\rho}_s(x)| \le Ld(t, s).$$

Let K be given by (4). Then, Kosorok (2008, Theorem 9.22) states that

$$N_{[]}(2\epsilon L, \{\bar{\rho}_u\}_{u\in K}, \|\cdot\|_\infty) \le N(\epsilon, K, d),$$

where $N(\epsilon, K, d)$ is the covering ϵ-number of (K, d). By assumption, $\mathbb{U}$ is of finite dimension that we will denote by D. By compacity of K, it has a finite diameter, say U. Therefore, $N(\epsilon, K, d) \le \max(1, U/\epsilon^D)$.

Let $u_0 \notin K,\, u_1 \in K$. We have that $\bar{\rho}_{u_1}$ is uniformly bounded,

$$|\bar{\rho}_{u_1}(x)| \le |\bar{\rho}_{u_0}(x)| + Ld(u_0, u_1) = Ld(u_0, u_1),$$

so that $\bar{\rho}_{u_0} = 0 \in [\bar{\rho}_{u_1} - \epsilon L, \bar{\rho}_{u_1} - \epsilon L]$, for $\epsilon > d(u_0, u_1)$. The brackets in the proof of Kosorok (2008, Theorem 9.22) are of the form $[\bar{\rho}_u - \epsilon L, \bar{\rho}_u + \epsilon L]$, so that $N_{[]}(2\epsilon L, \{\bar{\rho}_u\}_{u\in\mathbb{U}}, \|\cdot\|_\infty) \le N(\epsilon, K, d)$ for $\epsilon > d(u_0, u_1)$. In particular, one bracket is enough for $\epsilon > \max(U^{1/D}, d(u_0, u_1))$, while, for $\epsilon \le \max(U^{1/D}, d(u_0, u_1))$, we have $N_{[]}(2\epsilon L, \{\bar{\rho}_u\}_{u\in\mathbb{U}}, \|\cdot\|_\infty) \le 1 + N_{[]}(2\epsilon L, \{\bar{\rho}_u\}_{u\in K}, \|\cdot\|_\infty) \le 1 + N(\epsilon, K, d) \le 2N(\epsilon, K, d)$.

Finally, since $L_r(P)$ is dominated by $\|\cdot\|_\infty$ for any probability measure P,

$$\begin{aligned}
J_{[]}(\delta, \{\bar{\rho}_u\}_{u\in\mathbb{U}}, L_r(P)) &= \int_0^\delta \sqrt{\log(N_{[]}(\epsilon, \{\bar{\rho}_u\}_{u\in\mathbb{U}}, L_r(P)))}d\epsilon \\
&\le \int_0^\delta \sqrt{\log(N_{[]}(\epsilon, \{\bar{\rho}_u\}_{u\in\mathbb{U}}, \|\cdot\|_\infty))}d\epsilon \\
&\le \int_0^{\min(\delta, 2L\max(U^{1/D}, d(u_0,u_1)))} \sqrt{\log(N(\tfrac{\epsilon}{2L}, K, d))}d\epsilon \\
&\le \int_0^{\min(\delta, 2L\max(U^{1/D}, d(u_0,u_1)))} \sqrt{\log\left(2^{D+1}UL^D\right) - \frac{1}{D}\log(\epsilon)}d\epsilon.
\end{aligned}$$

As $\lim_{\delta\to 0} \int_\delta^1 \sqrt{-\log(\epsilon)}d\epsilon < \infty$, we conclude that $J_{[]}(\delta, \{\bar{\rho}_u\}_{u\in\mathbb{U}}, L_r(P)) < \infty$.

□

J Moments of the Hölder constant of a stochastic process

Let $(W_t)_{t\in[0,T]}$ be a stochastic process. A path $t \mapsto W_t(\omega)$ is said to be α-Hölder if $|W_t(\omega) - W_s(\omega)| \le \Lambda_{W(\omega)}|s-t|^\alpha$, for any $s, t \in [0, T]$. Many processes, for example Gaussian processes, do not admit a uniform constant. Based on Azaïs and Wschebor (2009); Hu and Le (2013); Shevchenko (2017), we will now give a condition under which $\Lambda_{W,\omega}$ is a random variable and we will calculate its moments.

Proposition J.1 (Azaïs and Wschebor (2009, Proposition 1.11)). *Suppose W satisfies* (16) *with K_{r_2,r_1} and let $\alpha \in]0, r_1/r_2[$. Then, there exists a version $(W'_t)_{t\in[0,1]}$ of W and a random variable $\Lambda_{W',\alpha} > 0$, such that, for all $s, t \in [0,1]$,*

$$P(|W'_t - W'_s| \leq \Lambda_{W',\alpha}|t-s|^{\alpha}) = 1 \quad \textit{and} \quad P(W(t) = W'(t)) = 1.$$

Theorem J.2 (Shevchenko (2017)). *Let $r_2 \in \mathbb{N}$ be such that $K_{r_2,\alpha r_2} < \infty$ and $1-\alpha > 1/r_2$, $r_2 \geq 2$,*

$$\mathbb{E}[\Lambda_W] \leq 16\ \tfrac{\alpha+1}{\alpha} T K^{1/r_2}_{r_2,r_2\alpha+1}.$$

In addition,

$$\mathbb{E}[\Lambda^k_W] \leq \begin{cases} \left(2^{3+2/r_2}\ \frac{\alpha+2/r_2}{\alpha}\right)^k K^{k/r_2}_{r_2,r_2\alpha+1}, & \textit{for } 0 < k \leq r_2, \\ \left(2^{3+2/r_2}\ \frac{\alpha+2/r_2}{\alpha}\right)^k K_{k,k(\alpha+2/r_2)-1}, & \textit{for } k > r_2. \end{cases}$$

Lemma J.3 (Garsia–Rodemich–Rumsey Inequality (Hu and Le, 2013, Lemma 1.1)). *Let $G : \mathbb{R}_+ \to \mathbb{R}_+$ be a non–decreasing function with $\lim_{x\to\infty} G(x) = \infty$ and $\delta : [0,T] \to [0,T]$ continuous and non–decreasing with $\delta(0) = 0$. Let G^{-1} and δ^{-1} be lower–inverses. Let $f : [0,T] \to \mathbb{R}$ be a continuous functions such that*

$$\int_0^T \int_0^T G\left(\frac{|f(x)-f(y)|}{\delta(x-y)}\right) dxdy \leq B < \infty.$$

Then, for any $s, t \in [0,T]$,

$$|f(s) - f(t)| \leq 8 \int_0^{|s-t|} G^{-1}(4B/u^2) d\delta(u).$$

Proof of Theorem J.2. Consider a path $W(\omega)$ of the stochastic process and set $B(\omega) := \int_0^T \int_0^T G\left(|W_t(\omega)W_s(\omega)|/\delta(t-s)\right) dtds$, where $G(u) = u^{r_2}$ and $\delta(u) = u^{\alpha+2/r_2}$. Then, $G^{-1}(u) = u^{1/r_2}$ and $d\delta/du = (\alpha+2/r_2)u^{\alpha+2/r_2-1}$. Applying Lemma J.3,

$$\begin{aligned} |W_t(\omega) - W_s(\omega)| &\leq 8 \int_0^{|s-t|} G^{-1}(4B(\omega)/u^2) d\delta(u) \\ &\leq 8 \int_0^{|t-s|} \left(\frac{4B(\omega)}{u^2}\right)^{1/r_2} (\alpha+2/p) u^{\alpha+2/r_2-1} du \\ &\leq 8(4B(\omega))^{1/r_2}(\alpha+2/r_2) \int_0^{|t-s|} u^{\alpha-1} du \\ &= 8(4B(\omega))^{1/r_2}\ \tfrac{\alpha+2/r_2}{\alpha}|t-s|^{\alpha}. \end{aligned}$$

As this is valid for any $s,\ t \in [0,T]$, $\Lambda_W(\omega) \leq 8(4B(\omega))^{1/r_2}\ (\alpha+2/r_2)/\alpha$. By Jensens' inequality,

$$\mathbb{E}[\Lambda_W] \leq 2^{3+2/r_2}\ \tfrac{\alpha+2/r_2}{\alpha} \mathbb{E}[B(\omega)^{1/r_2}] \leq 2^{3+2/r_2}\ \tfrac{\alpha+2/r_2}{\alpha} \mathbb{E}[B(\omega)]^{1/r_2}. \tag{J.22}$$

By linearity of expectation,

$$\begin{aligned} \mathbb{E}\left[\int_0^T \int_0^T G\left(\frac{|W_t(\omega)W_s(\omega)|}{\delta(t-s)}\right) dtds\right] &= \int_0^T \int_0^T \frac{\mathbb{E}[|W_t(\omega)W_s(\omega)|^{r_2}]}{\delta(t-s)^{r_2}} dtds \\ &= \int_0^T \int_0^T \frac{\mathbb{E}[|W_t(\omega)W_s(\omega)|^{r_2}]}{|t-s|^{p\alpha+2}} dtds \\ &\leq \int_0^T \int_0^T K_{p,p\alpha+1} dtds \\ &= T^2 K_{r_2,r_2\alpha+1}. \end{aligned}$$

Finally, $\mathbb{E}[\Lambda_W] \leq 2^{3+2/r_2}\ T^{2/r_2} K^{1/r_2}_{r_2,r_2\alpha+1}\ (\alpha+2/r_2)/\alpha$, as long as $r_2\alpha + 1 \leq r_2$ and we can simplify the constants if $r_2 > 2$. Consider now the higher moments. If $k \leq r_2$, we can still apply Jensens' inequality in (J.22):

$$\begin{aligned}
\mathbb{E}[\Lambda_W^k] &\leq \left(2^{3+2/r_2}\ \tfrac{\alpha+2/r_2}{\alpha}\right)^k \mathbb{E}[B(\omega)^{k/r_2}] \\
&\leq \left(2^{3+2/r_2}\ \tfrac{\alpha+2/r_2}{\alpha}\right)^k \mathbb{E}[B(\omega)]^{k/r_2} \\
&\leq \left(2^{3+2/r_2}\ \tfrac{\alpha+2/r_2}{\alpha}\right)^k K_{r_2,r_2\alpha+1}^{k/r_2}.
\end{aligned}$$

However, if $k \geq r_2$,

$$\begin{aligned}
\mathbb{E}\left[\left(\int_0^T \int_0^T G\left(\frac{|W_t(\omega)W_s(\omega)|}{\delta(t-s)}\right) dtds\right)^{k/r_2}\right] &= \int_0^T \int_0^T \frac{\mathbb{E}[|W_t(\omega)W_s(\omega)|^k]}{\delta(t-s)^k} dtds \\
&= \int_0^T \int_0^T \frac{\mathbb{E}[|W_t(\omega)W_s(\omega)|^k]}{|t-s|^{k\alpha+2k/r_2}} dtds \\
&\leq \int_0^T \int_0^T K_{k,k(\alpha+2/r_2)-1} dtds \\
&= T^2 K_{k,k(\alpha+2/r_2)-1}.
\end{aligned}$$

□

References

Azaïs, J.-M. and Wschebor, M. (2009). *Level Sets and Extrema of Random Processes and Fields*. John Wiley & Sons, Inc., Hoboken, NJ, Hoboken, NJ, USA.

Bühlmann, P. (1995). The blockwise bootstrap for general empirical processes of stationary sequences. *Stochastic Process. Appl.*, 58(2):247–265.

Chazal, F., de Silva, V., Glisse, M., and Oudot, S. (2016). *The Structure and Stability of Persistence Modules*. Number 2191-8198 in SpringerBriefs in Mathematics. Springer, Cham.

Cohen-Steiner, D., Edelsbrunner, H., Harer, J., and Mileyko, Y. (2010). Lipschitz Functions Have L p -Stable Persistence. *Found. Comput. Math.*, 10(2):127–139.

Doukhan, P. (1995). *Mixing: Properties and Examples*, volume 85 of *Lecture Notes in Statistics*. Springer New York, NY.

Hu, Y. and Le, K. (2013). A multiparameter Garsia-Rodemich-Rumsey inequality and some applications. *Stochastic Process. Appl.*, 123(9):3359–3377.

Kosorok, M. R. (2008). *Introduction to Empirical Processes and Semiparametric Inference*. Springer Series in Statistics. Springer New York, New York, NY.

Meyn, S. P. and Tweedie, R. L. (2012). *Markov chains and stochastic stability*. Springer Science & Business Media.

Perez, D. (2022). On C0-persistent homology and trees. page 41.

Radulović, D. (1996). The bootstrap for empirical processes based on stationary observations. *Stochastic Process. Appl.*, 65(2):259–279.

Shevchenko, G. (2017). Kolmogorov continuity theorem and holder norm. MathOverflow. https://mathoverflow.net/q/279085 (version: 2017-08-19).

Steinwart, I. (2022). Mathematics of Gaussian Processes for Machine Learning. Course at the 3-Day meeting of statisticians in Paris.