ADJUSTMENT COEFFICIENT FOR RISK PROCESSES IN SOME DEPENDENT CONTEXTS

H. COSSETTE, E. MARCEAU, AND V. MAUME-DESCHAMPS

ABSTRACT. Following [18], we study the adjustment coefficient of ruin theory in a context of temporal dependency. We provide a consistent estimator of this coefficient, and perform some simulations.

Adjustment coefficient w for risk processes may describe the behavior of ruin probability. Several results for sums of i.i.d. claims exist: in [12], H.U. Gerber gave an exact formula for finite time ruin probabilities involving the adjustment coefficient w, [19] provide a consistent estimator of w, V. Mammisch [15] gave a necessary and sufficient condition for the existence of w. In dependent contexts, let us cite H.U. Gerber [13] for auto-regressive processes, [2] for an extension to ARMA processes and [3, 4] for the study of the adjustment coefficients in Markovian environments . The main objective of the parper is to provide a non parametric estimation of the adjustement coefficient introduced in [18] in dependent contexts. We give a general dependent context (weak temporal dependency in the sense of [7]) for which our estimator is consistent.

The paper is organized as follows :

- Section 1 contains the definitions and elementary properties of weak-dependent processes as well as adjustment coefficient. To make short, w^i, the independent coefficient, will be the adjustment coefficient if the process is i.i.d. while w^d will be the adjustment coefficient of a dependent sequence.
- In Section 2, we prove that w^d may be seen as a limit (for $r \to \infty$) of independent coefficients w^i_r. We also provide some general examples for which the adjustment coefficient w^d may be defined.
- Section 3 is devoted to the estimation of coefficients w^i and w^d and contains the main results : we construct consistent estimators (see Theorems 3.3, 3.5 and 3.10). Note that in [2], an estimation of w^d is given for ARMA processes which is based on the estimation of the ARMA parameters. Our procedure is completely non parametric.
- In Section 4 we provide some simulations.

1. SETTING

We consider $(Y_n)_{n \in \mathbb{N}}$ a sequence of random variables and R_u the event $\{Y_n > u$ for some $n \geq 1\}$. Y_n is interpreted as the value of the claim surplus

1

process of a company at the end of the year n. R_u is understood as the ruin event for an intial reserve $u > 0$. We could write

$$Y_n = \sum_{i=0}^{n} X_i$$

where X_i is the gain/loss of the company during the year i.

1.1. **Weak dependent processes.**

This last decade, *Doukhan and al.* ([7, 6, 9, 8]) have developed new dependence definitions that both extend classical probabilistic definitions (as α or Φ mixing) and are satisfied for several useful models (like $ARMA$ or $ARCH$) that are neither α nor Φ mixing in the standard way. Roughly speaking, in the classical probabilistic definitions of mixing, the functions f and g in the definition below (1.1) belong to the whole class of square integrable functions.

Define (see [7]) for a real valued or vector valued process $(X_t)_{t\in\mathbb{N}}$,

$$(1.1) \qquad \varepsilon(k) = \sup \frac{|\mathrm{Cov}(f(X_{i_1},\ldots,X_{i_u}), g(X_{j_1},\ldots,X_{j_v}))|}{c(f,g)}$$

where the supremum is taken over multi-indices $\mathbf{i} = (i_1,\ldots,i_u), \mathbf{j} = (j_1,\ldots,j_v)$ such that:

$$i_1 < \cdots < i_u \le i_u + k \le j_1 < \cdots < j_v$$

and all functions $f : \mathbb{R}^u \longrightarrow \mathbb{R}, g : \mathbb{R}^v \longrightarrow \mathbb{R}$ are bounded and Lipschitz functions, with respect to the distance :

$$d(x,y) = \sum_{i=1}^{p} |x_i - y_i|, \ x = (x_1,\ldots,x_p), \ y = (y_1,\ldots,y_p).$$

Remark. We could replace the space of Lipschitz functions by other spaces of regular functions (differentiable functions, functions of bounded variation ...), see [17] for a general condition of convenient functional spaces.

We define a notion of weak dependence according to the function $c(f,g)$.

Definition 1. Consider the following functions $c(f,g)$, defined for $f : \mathbb{R}^u \longrightarrow \mathbb{R}$ and $g : \mathbb{R}^v \longrightarrow \mathbb{R}$ bounded and Lipschitz functions, $\mathrm{lip}(f)$ is the Lipschitz coefficient of the function f.

  (1) $c(f,g) = v\|f\|_\infty \mathrm{lip}(g)$, we say that the sequence $(X_i)_{i\in\mathbb{N}}$ is θ-weakly dependent if the corresponding mixing coefficients sequence $(\varepsilon(k))_{k\in\mathbb{N}}$ is summable.
  (2) $c(f,g) = u\mathrm{lip}(f)\|g\|_\infty + v\|f\|_\infty \mathrm{lip}(g)$, we say that the sequence $(X_i)_{i\in\mathbb{N}}$ is η-weakly dependent if the corresponding mixing coefficients sequence $(\varepsilon(k))_{k\in\mathbb{N}}$ is summable.

This class of dependent processes is very rich and enjoy lots of nice probabilistic properties. Let us remark that weak dependent processes need not to be stationary.

For completeness, we recall the definitions of α, Φ and Ψ mixing. Ψ mixing may be defined in a formalism close to that of Definition 1 while for α and Φ mixing, it is not clear that the same formalism is meaningfull.

Definition 2. Consider mixing coefficients $\varepsilon(k)$ defined by Equation (1.1) where the supremum is taken over functions $f : \mathbb{R}^u \longrightarrow \mathbb{R}$ and $g : \mathbb{R}^v \longrightarrow \mathbb{R}$ in L^2 and $c(f,g) = \|f\|_1 \|g\|_1$. We say that the sequence $(X_i)_{i\in\mathbb{N}}$ is Ψ mixing if the corresponding mixing coefficients $\varepsilon(k)$ are bounded.

α and Φ mixing coefficients are defined as :

$$\alpha(\mathcal{U},\mathcal{V}) = \sup_{U\in\mathcal{U},\ V\in\mathcal{V}} |\mathbb{P}(U\cap V) - \mathbb{P}(U)\mathbb{P}(V)|,$$

$$\Phi(\mathcal{U},\mathcal{V}) = \sup_{U\in\mathcal{U},\ V\in\mathcal{V}} \left| \frac{\mathbb{P}(U\cap V)}{\mathbb{P}(U)} - \mathbb{P}(V) \right|.$$

A process $(X_t)_{\in\mathbb{Z}}$ is α (resp. Φ) mixing is the coeficients

$$\alpha_X(r) = \sup_{i\in\mathbb{Z}} \alpha(\sigma(X_t,\ t\leq i),\ \sigma(X_t,\ t\geq i+r)),$$

$$\text{resp. } \Phi_X(r) = \sup_{i\in\mathbb{Z}} \Phi(\sigma(X_t,\ t\leq i),\ \sigma(X_t,\ t\geq i+r))$$

go to 0.

1.2. **Adjustment coefficient.** In the classical i.i.d. (i.e. the X_i are i.i.d. random variables) model of ruin theory, the adjustment coefficient $w > 0$ is defined as the unique positive solution of $\lambda(w) = 0$ with

$$\lambda(t) = \log \mathbb{E}\left[\exp(tX_1)\right]$$

assumed to be well defined. Mammischt ([15]) gave a necessary and sufficient existence condition for w. The importance of the adjustment coefficient is revealed by the exact formula due to Gerber ([12]) : let T be the ruin time $(T = \inf\{k \in \mathbb{N} \ / \ Y_k > u\})$,

$$\mathbb{P}(R_u) = \mathbb{P}(T < \infty) = \frac{e^{-wu}}{\mathbb{E}[e^{-wY_T}|T < \infty]},$$

and the famous de Finetti bound follows :

$$\mathbb{P}(R_u) \leq e^{-wu}.$$

We shall focus on the following asymptotic result also due to Lundberg :

$$\lim_{u\to\infty} \frac{\log \mathbb{P}(R_u)}{u} = -w.$$

As already mentioned above, several attempts to extend these results to dependent and/or non stationary settings have been proposed. We wish to give a general dependent setting in which such an asymptotic result holds, as well as provide a consistent estimator to the adjustment coefficient in this dependent context. Our approach does not require a precise knowledge of the dependence structure, nor on the law of $(X_i)_{i\in\mathbb{N}}$ but only an information on the speed of mixing (given by Equation (1.1)).

Following [18] we assume that : there exists $t_0 > 0$ such that for all $0 < t < t_0$,

(1.2) $$c(t) = \lim_{n\to\infty} \frac{\log \mathbb{E}\left[\exp(tY_n)\right]}{n} \text{ exists.}$$

Also, there exists $0 < t < t_0$ such that $c(t) = 0$.

We shall provide a sufficient condition that implies existence and uniqueness of a positive solution to $c(t) = 0$, provided that (1.2) is satisfied. We shall denote w^d this unique solution. We shall also denote by w^i the unique positive solution to $\lambda(t) = 0$ with

$$\lambda(t) = \log \mathbb{E}\left[\exp(tX_1)\right].$$

Of course, if the sequence X_i is i.i.d. then $w^i = w^d$.

1.3. **Existence condition.** In Mammisch ([15]) it is proven that w^i exists (and is unique as a positive solution to $\lambda(t) = 0$) if and only if the following three conditions are satisfied. We shall denote by (E) these three conditions.

(1) $\mathbb{E}(X_1) < 0$,
(2) $\mathbb{P}(X_1 > 0) > 0$ and
(3) either $a < \infty$ and $\mathbb{E}(e^{aX_1}) \geq 1$ or $a = \infty$ where

(1.3) $$a = \sup\{t \geq 0, \ \mathbb{E}(e^{tX_1}) < \infty\}.$$

This condition, together with the weak dependence assumption is sufficient to get consistency and asymptotic normality of the estimator of w^i. In order to get existence and uniqueness (as a positive solution to $c(u) = 0$) of w^d, we shall need some additional conditions.

Property 1.1. Assume that the limit $c(u)$ is well defined on $[0, u_0[$, in particular, for all $t < u_0$, for n large enough, $\mathbb{E}(e^{tY_n}) < \infty$. Assume that (E) is satisfied for $0 < a \leq u_0$, and

(1) $c_n(t)$ exists for all n and $0 \leq t < a$,
(2) for large enough n, $\mathbb{P}(Y_n > 0) > 0$
(3) if $a < \infty$ then for n large enough, $\lim\limits_{t \to a-} \mathbb{E}(e^{tY_n}) \geq 1$,
(4) $c'(0+) < 0$ or equivalently, $\exists t > 0$ such that $c(t) < 0$

then there exists a unique positive solution to $c(u) = 0$. This solution is denoted by w^d.

Proof. We adapt Mammish's arguments. Recall that any convex function is continuous and admits left and right derivatives on all points where it is defined. Moreover, if f_n is a sequence of convex functions defined on $[0, u_0[$ and converging to f on $[0, u_0[$ then f is a convex function and the convergence is uniform on any compact subset of $[0, u_0[$.
The function c is the limit on $[0, u_0[$ of convex functions

$$c_n(t) = \frac{1}{n} \log \mathbb{E}(e^{tY_n}).$$

Thus c is a convex function with $c(0) = 0$ and we assume that $c'(0+) < 0$ (which is equivalent to $\exists\, t > 0$ such that $c(t) < 0$ by convexity).

- If $a < \infty$ then, we assume that for large enough n, $\infty \geq \mathbb{E}(e^{aY_n}) \geq 1$ thus $\infty \geq c(a) \geq 0$. Since c is continuous (because it is convex), we deduce that there exists $w > 0$ such that $c(w) = 0$. This solution is unique because of the convexity of c.
- If $a = +\infty$, then $c(t)$ is well defined for any $t \in \mathbb{R}^+$. Because $\mathbb{P}(Y_n > 0) > 0$ for n large enough, we have that for large enough n, $\lim\limits_{t \to \infty} \mathbb{E}(e^{tY_n}) = +\infty$. As a consequence, we have that $\lim\limits_{t \to \infty} c(t) \geq 0$.

Since we assume that $c(t) < 0$ for some t, the convexity of c then implies that there exists t_0 such that $c'(t\pm) > 0$ for any $t > t_0$. We deduce that $c(t) > 0$ for t large enough. Since c is continuous (because it is convex), we deduce that there exists $w > 0$ such that $c(w) = 0$. This solution is unique because of the convexity of c.

$\square$

Remark. The condition $\mathbb{P}(Y_n > 0) > 0$ is necessary because if it exists $w > 0$ such that $c(w) = 0$ then by convexity, either there exists $t > 0$ such that $c(t) > 0$ or $c(t) = +\infty$ for all $t \geq w$ which implies that $\mathbb{E}(e^{tY_n}) > 1$ for large enough n which implies $\mathbb{P}(Y_n > 0) > 0$ for large enough n.

2. Limit result and examples

2.1. **Asymptotic behavior for ruin probability.** In [18], it is proven that if the adjustment coefficient w^d exists then it describes the asymptotic behavior of the ruin probability in the following sense:

$$(2.1) \qquad \lim_{u \to \infty} \frac{\log \mathbb{P}(R_u)}{u} = -w^d.$$

As a consequence to (2.1), we obtain that if it exists, w^d is the limit of the adjustment coefficients of the sequence $(Y_n)_{n \in \mathbb{N}}$.

Corollary 2.1. *Assume that the hypotheses of Property 1.1 are satisfied. For large enough n, there exists a unique $w_n > 0$ such that*

$$\mathbb{E}(e^{w_n Y_n}) = 1$$

and $w^d = \lim_{n \to \infty} w_n$.

Proof. The existence of w_n follows from [15]: Y_n satisfies the existence hypotheses of Mammish for n large enough. Applying Markov's inequality, we get for all $K > 0$,

$$w_n \leq \frac{-\log \mathbb{P}(Y_n > u)}{u}.$$

Then, (2.1) implies that any limit point w of the sequence w_n verifies : $w \leq w^d$. The convergence of functions c_n to c is uniform on $[0, w^d]$, thus we have that $c(w) = 0$, so that either $w = 0$ or $w = w^d$. Now, 0 cannot be a limit point of the sequence w_n because otherwise we would have that $c(t) \geq 0$ for all $t \geq 0$ which contradicts the hypotheses of Property 1.1. We conclude that $w_n \overset{n \to \infty}{\longrightarrow} w^d$. $\square$

Let us give some examples for which the function $c(t)$ is well defined. We recall the following result on approximate sub additive sequences due to Hammersley [14].

Lemma 2.2. *Assume $h : \mathbb{N} \longrightarrow \mathbb{R}$ be such that for all $n, m \geq 1$,*

$$h(n + m) \leq h(n) + h(m) + \Delta(m + n),$$

with Δ a non decreasing sequence satisfying :

$$(2.2) \qquad \sum_{r=1}^{\infty} \frac{\Delta(r)}{r(r+1)} < \infty.$$

Then, $\lambda = \lim_{n\to\infty} \dfrac{h(n)}{n}$ exists and is finite. Moreover, for all $m \geq 1$,

$$\lambda \leq \frac{h(m)}{m} - \frac{\Delta(m)}{m} + 4 \sum_{r=2m}^{\infty} \frac{\Delta(r)}{r(r+1)}.$$

Of course, for $\Delta(r) = O(1)$, then (2.2) is satisfied. Lemma 2.2 asserts that $\Delta(r)$ may go to infinity but not too fast.

2.2. **Ψ-mixing processes.** According to Definition 2, we consider the following classical Ψ-mixing condition :

(2.3) $$\Psi(k) = \sup \frac{\mathrm{Cov}(f(X_{i_1}, \ldots, X_{i_u}), g(X_{j_1}, \ldots, X_{j_v}))}{\|f\|_1 \|g\|_1} < \infty,$$

where the supremum is taken over functions f, $g \in L^2$ and over multi-indices $\mathbf{i} = (i_1, \ldots, i_u)$ and $\mathbf{j} = (j_1, \ldots, j_v)$ with $i_1 < \cdots < i_u < i_u + k \leq j_1 < \cdots < j_v$. The sequence if Ψ-mixing is bounded.

Proposition 2.3. *Assume that for $t \in [0, a[$, for all $n \in \mathbb{N}$, $\mathbb{E}(e^{2tS_n}) < \infty$ and $(X_n)_{n\in\mathbb{N}}$ is a Ψ-mixing process then*

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}(e^{tS_n})$$

exists for any $t \in [0, a[$.

Proof. Using (2.3), we have :

$$\left| \mathbb{E}(e^{tS_{n+m}}) - \mathbb{E}(e^{tS_n})\mathbb{E}(e^{tS_m}) \right| \leq \Psi(1)\mathbb{E}(e^{tS_n})\mathbb{E}(e^{tS_m}).$$

We conclude the proof by using Lemma 2.2.   □

Examples of Ψ-mixing processes are finite state Markov chains of any order but also Variable Length Markov Chains (VLMC) on a finite state (see Lemma 3.1 in [11]).
Even if Ψ (Φ, α)-mixing processes are often used in probability theory, lots of useful processes (like ARMA processes) are not Ψ (Φ, α)-mixing. In the following two subsections, we provide a class of η mixing processes for which the function $c(t)$ is well defined provided the mixing is sufficiently fast and the variables X_i are almost surely bounded. This condition is close to the one used in [1] by Bric and Dembo for one other class of mixing processes (namely α mixing processes). Then we prove that if the sequence has some structure (here, the sequence is a Bernoulli shift) then the condition on the speed of mixing may be weakened. We refer to [7] for other examples of θ and η weakly mixing processes (including ARMA and ARCH processes).

2.3. **Super mixing processes.** We prove that if the process $(X_i)_{i\in\mathbb{N}}$ is η weakly dependent (recall Definition 1) with dependence coefficient $\varepsilon(n) = O\left(\theta^{n(\ln n)^{\beta}}\right)$ with $0 < \theta < 1$ and $\beta > 1$ then the function $c(t)$ is well defined provided $|X_i| \leq M$ a.e. A η weakly dependent process with dependence coefficient $\varepsilon(n) = O\left(\theta^{n(\ln n)^{\beta}}\right)$ will be called a super mixing process.

Proposition 2.4. *Assume $(X_n)_{n\in\mathbb{N}}$ is a η weakly dependent process with mixing coefficients $\varepsilon(n) = O(e^{-cn(\ln n)^\beta})$ with $c > 0$, $\beta > 1$. Moreover, assume that there exists $M > 0$ such that $|X_i| \leq M$ a.e. Then the sequence $c(t)$ is well defined on $\mathbb{R}$.*

Proof. For any $0 < j < k$, let $S_j^k = \sum_{\ell=j}^{k} X_\ell$, because $|X_i| \leq M$ a.e.,

$$e^{-t(k-j)M} \leq \mathbb{E}(e^{tS_j^k}) \leq e^{t(k-j)M}.$$

Also, for any $j \leq \ell \leq j$,

$$e^{-t(\ell-j)M}\mathbb{E}(e^{tS_\ell^k}) \leq \mathbb{E}(e^{tS_j^k}) \leq e^{t(\ell-j)M}\mathbb{E}(e^{tS_\ell^k}).$$

Remark that the function : $x \mapsto e^{tx}$ is bounded above by e^{tM} and has Lipschitz constant te^{tM} for $x \in [-M, M]$.
Fix an integer $0 < r < \max(n, m)$. Firstly, assume that $n \leq m$ and, using the definition of η weak dependence, we get:

$$\mathbb{E}(e^{tS_{n+m}}) = \mathbb{E}(e^{tS_1^n}e^{tS_{n+1}^{n+m}})$$
$$= \mathbb{E}\left(e^{tS_1^n} \cdot e^{tS_{n+r+1}^{n+m}} \cdot e^{tS_{n+1}^{n+r}}\right)$$
$$\leq e^{trM}\left(\mathbb{E}(e^{tS_1^n}) \cdot \mathbb{E}(e^{tS_{n+r+1}^{n+m}}) + \varepsilon(r)(n+m)te^{tnM}e^{t(m-r)M}\right)$$
$$\leq e^{2trM}\mathbb{E}(e^{tS_n})\mathbb{E}(e^{tS_m}) + (n+m)\varepsilon(r)te^{2t(n+m)M}\mathbb{E}(e^{tS_n})\mathbb{E}(e^{tS_m})$$
$$\leq \mathbb{E}(e^{tS_n})\mathbb{E}(e^{tS_m})(e^{2trM} + (n+m)\varepsilon(r)te^{2t(n+m)M}).$$

We conclude the proof by choosing $r = O(\frac{n+m}{\ln(n+m)^\kappa})$, $1 < \kappa < \beta$ and applying Lemma 2.2 with a function:

$$\Delta(r) = O\left(\frac{r}{(\ln r)^\kappa}\right).$$

If $n > m$, the proof is similar but uses the decomposition :
$S_{n+m} = S_1^m + S_{m+1}^{m+r} + S_{m+r+1}^{n+m}$. $\qquad\qquad\square$

2.4. **Bernoulli shifts.** Causal Bernoulli shifts are processes defined as:

$$X_n = H(\xi_{n-j}, \ j \in \mathbb{N})$$

with H a measurable function and $(\xi_n)_{n\in\mathbb{Z}}$ an i.i.d. process. We shall assume the following regularity condition on H : define the continuity coefficients

$$d_n = \|\sup_{u=(u_0,u_{-1},\ldots)} |H(\xi_{n-i}, i \in \mathbb{N}) - H(\xi_n, \ldots, \xi_1, u_0, u_{-1}, \ldots)|\|_\infty$$

and assume that the sequence d_n is summable. By adapting the arguments of [7] we may prove that such a process is θ-dependent with mixing coefficient $\theta(n) = d_n$.

Proposition 2.5. *Assume that $(X_n)_{n\in\mathbb{N}}$ is a Bernoulli shift satisfying the summability condition for the continuity coefficients d_n. Then the sequence $\ln \mathbb{E}(e^{tS_n})$ satisfies the hypotheses of Lemma 2.2 and thus $c(t)$ is well defined.*

Proof. We fix a sequence of real numbers $u = (u_0, u_{-1}, \dots)$ and we write

$$
\begin{aligned}
S_i^\ell &= \sum_{j=i}^\ell X_j \\
&= \sum_{j=i}^\ell H(\xi_j, \dots, \xi_i, \dots) \\
&= \underbrace{\sum_{j=i}^\ell H(\xi_j, \dots, \xi_i, u_0, u_{-1}, \dots)}_{:=U_i^\ell} + \\
\end{aligned}
$$

(2.4)
$$
\underbrace{\sum_{j=i}^\ell H(\xi_j, \dots, \xi_i, \dots) - H(\xi_j, \dots, \xi_i, u_0, u_{-1}, \dots)}_{:=d_i^\ell(u,\xi)}.
$$

Using the stationarity of $(\xi_n)_{n \in \mathbb{N}}$, we have that

$$
d_i^\ell \leq \sum_{j=i}^\ell d_{j-i} = \sum_{j=0}^{\ell-i} d_j.
$$

Now, U_1^n and U_{n+1}^m are independent random variables and thus,

$$
\mathbb{E}(e^{tS_{n+m}}) \leq e^{t(\sum_{i=1}^n d_i + \sum_{i=1}^m d_i)} \mathbb{E}(e^{tU_1^n}) \mathbb{E}(e^{tU_{n+1}^{n+m}}).
$$

Applying once more (2.4), we get:

$$
\mathbb{E}(e^{tS_{n+m}}) \leq \exp\left[4t \sum_{i=1}^{\max(n,m)} d_i \right] \mathbb{E}(e^{tS_n}) \mathbb{E}(e^{tS_m}).
$$

If we denote $D = \sum_{j \in \mathbb{N}} d_j$, we get

$$
\mathbb{E}(e^{tS_{n+m}}) \leq e^{4tD} \mathbb{E}(e^{tS_n}) \mathbb{E}(e^{tS_m})
$$

and we conclude by applying Lemma 2.2. □

Remark. In the above proposition, we could replace the hypotheses of summability of the sequence $(d_j)_{j \in \mathbb{N}}$ by the summability of

$$
\frac{\Delta(r)}{r(r+1)} \text{ with } \Delta(r) = \sum_{j=0}^r d_j,
$$

Lemma 2.2 would apply as well and the limit $c(t)$ would be well defined.

In [7], it is mentionned that stationary ARMA processes are examples of Bernouilli shifts. Non linear autoregressive processes may also be examples of Bernouilli shifts. Nevertheless, in order to satisfy that the d_j are well defined, it requires that the innovation is bounded. We claim that Proposition 2.5 remains true for some unbounded Bernouilli shifts but we where unable to prove it.

3. ESTIMATION

3.1. **Definition of estimators.** In this section, we assume that the sequence $(X_i)_{i \geq 1}$ is stationary and that the hypotheses of Proposition 1.1 are satisfied. For $r \in \mathbb{N}$, the functions $\mathbb{E}(e^{tX_1})$ and $\mathbb{E}(e^{tY_r})$ may be estimated by their empirical moment versions: for $k \in \mathbb{N}$,

$$\widehat{m}_k(t) = \frac{1}{k} \sum_{i=1}^{k} e^{tX_i},$$

$$\widehat{M_k^r}(t) = \frac{1}{k} \sum_{i=0}^{k-1} e^{tZ_i^r},$$

where $Z_i^r = \sum_{j=1}^{r} X_{j+ir}$. Then we define $\widehat{w}^i$ as the positive solution to $\log \widehat{m}_k(t) = 0$ and $\widehat{w}_r$ as the positive solution to $\frac{1}{r} \log \widehat{M_k^r}(t) = 0$. We shall prove that $\widehat{w}^i$ is a consistent estimator of w^i and there exists an $r = r(k)$ such that $\widehat{w}_r = \widehat{w}^d$ is a consistent estimator of w^d. We shall also prove that they satisfy a central limit theorem.

Before stating and proving our main results on the asymptotic properties of the estimators $\widehat{w}^i$ and $\widehat{w}_r$, we prove that $\widehat{w}^i$ exists almost surely (the proof for $\widehat{w}_r$ will be done later because it requires some weak dependence property for $(Z_i^r)_{i \in \mathbb{N}}$ which is proven in Lemma 3.4).

Proposition 3.1. *Assume that the sequence $(X_i)_{i \in \mathbb{N}}$ is η or θ weakly dependent and satisfies Condition (E). Then, $\widehat{w}^i$ exists eventually almost surely as $k \longrightarrow \infty$.*

Proof. We begin by noting that the θ or η weak dependence implies ergodicity, see for example [6]. Following Section 1.3, we have that $\widehat{w}^i$ exists and is unique if and only if,

(1) $\frac{1}{k} \sum_{i=1}^{k} X_i < 0$,
(2) $\{i = 1, \ldots, k \ / \ X_i > 0\}$ is not empty.

Mammisch's third condition is satisfied with $a = \infty$ because here, expectations are finite sums. The two above conditions are eventually almost everywhere satisfied because of the ergodic theorem. $\square$

3.2. **Asymptotic properties of $\widehat{w}^i$.** Asymptotic properties of the estimators $\widehat{w}^i$ and $\widehat{w_r}$ are done by using the same approach as the one used to prove results on asymptotic properties for M-estimators in a parametric context (see [20] section 5). It is known (see [9, 7]) that the process $(X_i)_{i \in \mathbb{N}}$ satisfies a central limit theorem with asymptotic variance

$$\Gamma^2 = \sum_{i=0}^{\infty} \operatorname{Cov}(X_0, X_i),$$

provided that the sequence $\operatorname{Cov}(X_0, X_i)$ is summable.

We obtain a central limit theorem for $\widehat{m}_k(t)$ by proving that the sequence $(e^{tX_n})_{n \in \mathbb{N}}$ is also weakly dependent. Then we prove a central limit theorem for the M-estimator $\widehat{w}^i$.

Let us recall the following results from [7] (Theorem 7.1 and Section 7.5.4).

Theorem 3.2. *Let $(Z_n)_{n\in\mathbb{N}}$ be an η-weakly dependent sequence with $\varepsilon(n) = O(n^{-2-\kappa})$ for $\kappa > 0$. Then,*

$$\Gamma^2 = \sum_{n\geq 0} \mathrm{Cov}(Z_0, Z_n),$$

is well defined and

$$\sqrt{n}\left(\frac{1}{n}\sum_{i=0}^{n-1} Z_i - \mathbb{E}(Z_0)\right) \overset{n\to\infty}{\longrightarrow} \mathcal{N}(0, \Gamma^2).$$

As a consequence, we get the following consistency result for $\widehat{w}^i$ as well as asymptotic normality. In order to use Theorem 3.2, we first need to prove that the sequence $(e^{tX_n})_{n\in\mathbb{N}}$ is also η-weakly dependent with $\varepsilon(n)$ decreasing to zero sufficiently rapidly.

Theorem 3.3. *Assume $(X_n)_{n\in\mathbb{N}}$ is η-weakly dependent with $\varepsilon(n) = O(\theta^n)$, $0 < \theta < 1$. We have for any $t \in [0, u_0[$ that*

$$\Gamma^2(t) = \sum_{n\geq 0} \mathrm{Cov}(e^{tX_0}, e^{tX_n})$$

is well defined on $[0, u_0[$ and

$$\sqrt{n}\left(\widehat{m}_n(t) - \mathbb{E}(e^{tX_0})\right) \overset{k\to\infty}{\longrightarrow} \mathcal{N}(0, \Gamma^2(t)).$$

$\widehat{w}^i$ converges in probability to w^i and

$$\sqrt{n}\left(\widehat{w}^i - w^i\right) \overset{n\to\infty}{\longrightarrow} \mathcal{N}(0, \Gamma_i^2)$$

with $\Gamma_i^2 = \dfrac{\Gamma^2(w^i)}{\mathbb{E}(X_1 e^{w^i X_1})^2}$.

As already mentioned, we begin by proving that the sequence of random variables $(e^{tX_n})_{n\in\mathbb{N}}$ is η-weakly dependent.

Lemma 3.4. *Assume $(X_n)_{n\in\mathbb{N}}$ is η-weakly dependent with mixing coefficient $\varepsilon(r)$. Then, for any $t \in [0, u_0[$, the sequence of random variables $(e^{tX_n})_{n\in\mathbb{N}}$ is η-weakly dependent with mixing coefficient $\varepsilon_t(r) \leq 2\mathbb{E}(e^{(t+\kappa)X_1})\varepsilon(r)^{\frac{\kappa}{t+\kappa}}$ with $\kappa > 0$ such that $t + \kappa \in [0, u_0[$.*

Proof. We follow the proof of Proposition 2.1 in [7]. Let f and g be two Lipschitz functions and for $M > 0$ fixed, $x \in \mathbb{R}$, denote $x^{(M)} = \min(x, M)$. Assume $(\mathbf{i}, \mathbf{j})$ are multi-indices such that

$$i_1 < \cdots < i_u \leq i_u + r \leq j_1 < \cdots < j_v,$$

and define:

$$F(X_{\mathbf{i}}) = f(e^{tX_{i_1}}, \ldots, e^{tX_{i_u}}) \qquad F^{(M)}(X_{\mathbf{i}}) = (e^{tX_{i_1}^{(M)}}, \ldots, e^{tX_{i_u}^{(M)}}),$$

$$G(X_{\mathbf{j}}) = g(e^{tX_{j_1}}, \ldots, e^{tX_{j_v}}) \qquad G^{(M)}(X_{\mathbf{j}}) = (e^{tX_{j_1}^{(M)}}, \ldots, e^{tX_{j_v}^{(M)}}).$$

Then,

$$\begin{aligned}
|\mathrm{Cov}(F(X_{\mathbf{i}}), G(X_{\mathbf{j}}))| \leq{} & 2\|f\|_\infty \mathbb{E}(|G(X_{\mathbf{j}}) - G^{(M)}(X_{\mathbf{j}})|) \\
& + 2\|g\|_\infty \mathbb{E}(|F(X_{\mathbf{i}}) - F^{(M)}(X_{\mathbf{i}})|) \\
& + |\mathrm{Cov}(F^{(M)}(X_{\mathbf{i}}), G^{(M)}(X_{\mathbf{j}}))|.
\end{aligned}$$

Recall that $\mathbb{E}(e^{sX_i}) < \infty$ for any $s \in [0, u_0[$ and let $\kappa > 0$ be such that
$$t + \kappa \in [0, u_0[,$$
then $\mathbb{E}(|e^{(t+\kappa)X_i}|) < \infty$, using the Markov inequality, we get:

$$\mathbb{E}(|G(X_{\mathbf{j}}) - G^{(M)}(X_{\mathbf{j}})|) \leq \operatorname{lip}(g) \sum_{k=1}^{v} \mathbb{E}(|e^{tX_{j_k}} - e^{tX_{j_k}^{(M)}}|)$$

$$\leq 2v\operatorname{lip}(g)e^{-\kappa M}\mathbb{E}(e^{(t+\kappa)X_1}).$$

Also, since $(X_n)_{n \in \mathbb{N}}$ is η-weakly dependent,

$$|\operatorname{Cov}(F^{(M)}(X_{\mathbf{i}}), G^{(M)}(X_{\mathbf{j}}))| \leq (u\operatorname{lip}(f)\,\|g\|_\infty + v\operatorname{lip}(g)\|f\|_\infty)te^{tM}\varepsilon(r).$$

Finally, we obtain for $M \geq 1$,

$$|\operatorname{Cov}(F(X_{\mathbf{i}}), G(X_{\mathbf{j}}))| \leq (u\operatorname{lip}(f)\,\|g\|_\infty + v\operatorname{lip}(g)\|f\|_\infty)$$
$$\times (te^{tM}\varepsilon(r) + 2\mathbb{E}(e^{(t+\kappa)X_1})e^{-\kappa M}).$$

To conclude, we choose $M = -\frac{1}{t+\kappa}\ln(\varepsilon(r))$. $\square$

Proof of Theorem 3.3. Lemma 3.4 together with Theorem 3.2 imply that for any $t \in [0, u_0[$,

$$\sqrt{n}\left(\widehat{m}_n(t) - \mathbb{E}(e^{tX_0})\right) \overset{n\to\infty}{\longrightarrow} \mathcal{N}(0, \Gamma^2(t))$$

provided that the sequence $\varepsilon_t(r) = \varepsilon(r)^{\frac{\kappa}{t+\kappa}} = O(r^{-2-\alpha})$ for some $\alpha > 0$. Since we assume that $\varepsilon(r) = O(\theta^r)$, $0 < \theta < 1$, this condition is satisfied. This convergence in law also leads to

$$\widehat{m}_n(t) \overset{n\to\infty}{\longrightarrow} \mathbb{E}(e^{tX_0}) \text{ in probability,}$$

moreover we have that this convergence takes place almost everywhere because of the ergodic theorem.

Now, let us consider the estimator $\widehat{w}^i$ of w^i. Following the proof of Lemma 5.10 in [20], we have that $\widehat{w}^i$ converges to w^i in probability (this uses the convergence in probability of $\widehat{m}_k(t)$, the continuity of the map $t \mapsto \widehat{m}_k(t)$ and the uniqueness of $\widehat{w}^i$ as a positive solution to $\widehat{m}_k(t) = 1$). The central limit theorem follows now from the Δ method :

$$\widehat{m}_k(\widehat{w}^i) - \widehat{m}_k(w^i) = (\widehat{w}^i - w^i)\frac{\partial \widehat{m}_k(w^i)}{\partial t} + \frac{1}{2}(\widehat{w}^i - w^i)^2\frac{\partial^2 \widehat{m}_k(\widetilde{w})}{\partial t^2},$$

with $\widetilde{w} \in [\min(w^i, \widehat{w}^i), \max(w^i, \widehat{w}^i)]$. Thus,

$$\sqrt{k}(\widehat{w}^i - w^i) = \frac{\sqrt{k}(\widehat{m}_k(\widehat{w}^i) - \widehat{m}_k(w^i))}{\dfrac{\partial \widehat{m}_k(w^i)}{\partial t} + \dfrac{1}{2}(\widehat{w}^i - w^i)\dfrac{\partial^2 \widehat{m}_k(\widetilde{w})}{\partial t^2}}.$$

We have that

$$\sqrt{k}(\widehat{m}_k(\widehat{w}^i) - \widehat{m}_k(w^i)) = \sqrt{k}(1 - \widehat{m}_k(w^i))$$
$$= \sqrt{k}(\mathbb{E}(e^{w^i X_1}) - \widehat{m}_k(w^i))$$

and therefore it is asymptotically normal with zero mean and variance $\Gamma^2(w^i)$. Moreover,

$$\frac{\partial \widehat{m}_k(w^i)}{\partial t} = \frac{1}{k}\sum_{j=1}^{k} X_j e^{w^i X_j}.$$

This quantity converges in probability to $\mathbb{E}(X_1 e^{w^i X_1})$ and we have that
$(\widehat{w}^i - w^i)\dfrac{\partial^2 \widehat{m}_k(\widetilde{w})}{\partial t^2}$ goes to zero in probability. Finally, we have proven that

$$\sqrt{n}(\widehat{w}^i - w^i) \overset{k\to\infty}{\longrightarrow} \mathcal{N}\left(0, \frac{\Gamma^2(w^i)}{\mathbb{E}(X_1 e^{w^i X_1})^2}\right).$$

$\square$

Remark. We could relax the hypotheses that $\varepsilon(r)$ decreases to 0 exponentially fast. It is sufficient that $\varepsilon(r)^{\frac{\kappa}{t+\kappa}} = O(r^{-2-\alpha})$ for some $\alpha > 0$. For example, some intermediate speed of mixing like $\varepsilon(r) = O\left(e^{-K(\ln r)^\beta}\right)$, $K > 0$, $\beta > 1$ or $\varepsilon(r) = O\left(\theta^{n^\alpha}\right)$, $0 < \theta < 1$, $0 < \alpha < 1$ is convenient.

3.3. Asymptotic properties of $\widehat{w}^d$.
Now, we are interested in the consistency of $\widehat{w}^d$.

Theorem 3.5. *Assume that hypotheses of Theorem 3.3 are satisfied. Then there exists a sequence $r = r(k) \overset{k\to\infty}{\longrightarrow} \infty$ such that $\widehat{w}^d$ converges in probability to w^d.*

Theorem 3.5 will be proven by rewriting 3.3 for $\widehat{w}_r$ instead of $\widehat{w}^i$ and then by using Corollary 2.1. We only need to prove that the sequence $(Z_i^r)_{i\in\mathbb{N}}$ satisfies some weak dependence property.

Lemma 3.6. *Assume that $(X_n)_{n\in\mathbb{N}}$ is η-weakly dependent with mixing coefficient $\varepsilon(k)$. Then, the sequence $(Z_i^r)_{i\in\mathbb{N}}$ is η-weakly dependent with mixing coefficient $\varepsilon_Z(k) = r\varepsilon(r(k-1))$.*

Proof. Let f and g be two Lipschitz functions. Assume $(\mathbf{i}, \mathbf{j})$ are multi-indices such that

$$i_1 < \cdots < i_u \le i_u + k \le j_1 < \cdots < j_v.$$

$$Then, \mathrm{Cov}(f(Z_{i_1}^r, \ldots, Z_{i_u}^r), g(Z_{j_1}^r, \ldots, Z_{j_v}^r))$$

$$= \mathrm{Cov}\left(\widetilde{f}(X_{i_1 r+1}, \ldots, X_{i_1(r+1)}, X_{i_2 r+1}, \ldots, X_{(i_u+1)r}),\right.$$

$$\left. \widetilde{g}(X_{j_1 r+1}, \ldots, X_{j_1(r+1)}, X_{j_2 r+1}, \ldots, X_{(j_v+1)r})\right)$$

$$\le r\varepsilon((k-1)r)\left(u\mathrm{lip}f\|g\|_\infty + v\|f\|_\infty\mathrm{lip}f\right),$$

where $\widetilde{\varphi}(x_1, \ldots, x_r, \ldots, x_{rk}) = \varphi\left(\displaystyle\sum_{i=1}^r x_i, \ldots, \sum_{i=1}^r x_{r(k-1)+i}\right)$. $\square$

As a corollary to Lemma 3.6, we deduce that for any r, $\widehat{w}_r$ exists eventually almost surely.

Corollary 3.7. *Assume that the hypotheses of Property 1.1 are satisfied and that the sequence $(X_i)_{i\in\mathbb{N}}$ is η-weakly dependent. Then, for any r, $\widehat{w}_r$ exists eventually almost surely.*

Proof. This is a direct consequence of Proposition 3.1 and Lemma 3.6. $\square$

Proof of Theorem 3.5. From Lemma 3.6 and Theorem 3.3, we get

$$\sqrt{k}\left(\widehat{M}_k^r(t) - \mathbb{E}(e^{tZ_1^r})\right) \overset{\mathcal{L}}{\longrightarrow} \mathcal{N}(0, \Gamma_r^2(t))$$

with

$$\Gamma_r^2(t) = \sum_{n \geq 0} \mathrm{Cov}(e^{tZ_0^r}, e^{tZ_n^r})$$

and if we denote by w_r the i-adjustment coefficient of the sequence Z_0^r and $\widehat{w_r}$ the positive solution to $\widehat{M_k^r}(t) = 1$,

$$\sqrt{k}(\widehat{w_r} - w_r) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \Gamma_r^2)$$

with

$$\Gamma_r^2 = \frac{\Gamma_r^2(w_r)}{\mathbb{E}(Z_1^r e^{w_r Z_1^r})^2}.$$

This implies that $\widehat{w_r}$ goes to w_r in probability, as k goes to infinity. We conclude the proof of Theorem 3.5 by using Corollary 2.1 : there exists a sequence $r(k) \xrightarrow{k \to \infty} \infty$ such that $\widehat{w_r}$ converges to w^d in probability. $\qquad \square$

Theorem 3.5 is interesting from a theoretical point of view but it is not so useful from a practical point of view. Indeed, it proves that $\widehat{w_r}$ converges to w^d for a sequence $r = r(k)$ but we have no information on how to choose r with respect to k. Moreover, provided that $\frac{\Gamma^2(w_r)}{\mathbb{E}(Z_1^r e^{w_r Z_1^r})^2}$ is converging, we could obtain a central limit theorem for $\sqrt{k}(\widehat{w_r} - w^d)$ with limit variance $\Gamma_d^2 = \lim_{r \to \infty} \frac{\Gamma^2(w_r)}{\mathbb{E}(Z_1^r e^{w_r Z_1^r})^2}$. This expression of the asymptotic variance in the central limit theorem is not useful from a practical point of view. We might nevertheless use moment and Bienaimé-Tchebitchev inequalities in order to get a useful relationship between r and k.
We apply the following inequality on the order 2 moment.

Proposition 3.8. *Let* $(W_i)_{i \in \mathbb{N}}$ *be a sequence of centered random variables and*

$$C_{j,2} = \sup_{\substack{t_1, \, t_2 \\ t_2 - t_1 = j}} \mathrm{Cov}(W_{t_1}, W_{t_2}).$$

$$\mathbb{E}(S_n^2) \leq 2n \sum_{j=0}^{n-1} C_{j,2},$$

where $S_n = \sum_{i=1}^{n} W_i$.
As a consequence, if W_i *is* η-*dependent and stationary, with mixing coefficient* $\varepsilon(k) \leq C\theta^k$, *we have that*

$$\mathbb{E}(S_n^2) \leq 16n M_m^{\frac{2}{m}} C^{\frac{m-2}{m}} \frac{1}{1 - \theta^{\frac{m-2}{m}}}$$

where $M_m = \mathbb{E}(|W_i|^m)$.

Proof. The first part of the proof of Proposition 3.8 may be found in [7] (see Lemma 4.6 p.79). For the second part, we proceed as in the proof of

Lemma 3.4 (see also [10]) : let $W_t^{(M)} = \max(\min(W_t, M), -M)$, (so that $W_t^{(M)} = W_t$ provided that $|W_t| \leq M$). Then,

$$
\begin{aligned}
\mathrm{Cov}(W_t, W_{t+j}) &\leq \mathrm{Cov}(W_t^{(M)}, W_{t+j}^{(M)}) + \mathrm{Cov}(W_t^{(M)}, (W_{t+j} - W_{t+j}^{(M)})) + \\
&\quad \mathrm{Cov}(W_{t+j}^{(M)}, W_j - W_j^{(M)}) + \mathrm{Cov}((W_{t+j} - W_{t+j}^{(M)}), W_j - W_j^{(M)}) \\
&\leq 2M^2\varepsilon(j) + 2M\|W_t - W_t^{(M)}\|_1 + \|(W_t - W_t^{(M)})^2\|_1 \\
&\leq 2M^2\varepsilon(j) + 6M^{2-m}M_m,
\end{aligned}
$$

where the last line is obtained by noting that

$$
\begin{aligned}
\|W_t - W_t^{(M)}\|_1 &= \int \mathbf{1}_{|W_t|>M}|W_t - W_t^{(M)}|d\mathbb{P} \\
&\leq 2\int \mathbf{1}_{|W_t|>M}|W_t|d\mathbb{P} \\
&\leq 2M^{1-m}M_m.
\end{aligned}
$$

Similarly, we obtain

$$
\|(W_t - W_t^{(M)})^2\|_1 \leq 4M^{2-m}M_m.
$$

We conclude by choosing $M = \left(\dfrac{M_m}{\varepsilon(j)}\right)^{\frac{1}{m}}$.   □

Proposition 3.9. *Assume that $(X_i)_{i\in\mathbb{N}}$ is an η weakly-dependent process, with mixing coefficient $\varepsilon(k) \leq C\theta^k$, $C > 0$, $0 < \theta < 1$. Then, for any $t \in [0, u_0[$, such that $3t \leq u_0$, for any $v > 0$, we have:*

$$
\mathbb{P}\left(|\widehat{M_k^r}(t) - \mathbb{E}(e^{tY_r})| > v\right) \leq \frac{4(Cr+3)\mathbb{E}(e^{3tY_r})^{2/3}}{v^2k(1-\theta^{\frac{1}{6}})}.
$$

Proof. We apply the Bienaimé-Tchebitchev inequality and Proposition 3.8 with $m = 3$. Let

$$
W_\ell(t) = e^{t\sum_{i=1}^r X_{i+r\ell}}.
$$

Following the lines of the proof of Proposition 3.8, we have that the covariance coefficients associated to $(W_\ell(t))_{\ell\in\mathbb{N}}$ are:

$$
C_{2,j} \leq M_m^{\frac{2}{m}}\theta^{\frac{j}{2}\frac{m-2}{m}}(2Cr+6),
$$

with $M_m = \mathbb{E}(e^{tmY_r})$. We have used the fact that

$$
\mathrm{Cov}(W_\ell^{(M)}(t), W_{\ell+j}^{(M)}(t)) \leq 2M_m^2 r\varepsilon(r(j-1)+1) \leq 2M_m^2 rC\theta^{\frac{j}{2}}.
$$

We choose $m = 3$ and apply Proposition 3.8, so that

$$
\mathbb{E}([k\widehat{M_k^r}(t) - k\mathbb{E}(e^{tY_r})]^2) \leq 2k(2Cr+6)[\mathbb{E}(e^{3tY_r})]^{\frac{2}{3}}\frac{1}{1-\theta^{\frac{1}{6}}}.
$$

We conclude by using the Bienaimé-Tchebitchev inequality.   □

This proposition shows that in order to get the consistency for $\widehat{w}_r$ we should choose $r(k) = o(\ln k)$.

Theorem 3.10. *Assume that $(X_i)_{i\in\mathbb{N}}$ is an η weakly-dependent process, with mixing coefficient $\varepsilon(k) = O(\theta^k)$, $0 < \theta < 1$ and that $3w^d < u_0$. Then, for $r = r(k) = o(\ln k)$, $\widehat{w}_r$ goes to w^d in probability.*

Proof. We have that

$$\mathbb{E}(e^{w_r Z_1^r}) - \widehat{M}_k^r(w_r) = \widehat{M}_k^r(\widehat{w}_r) - \widehat{M}_k^r(w_r)$$

$$= (\widehat{w}_r - w_r)\frac{\partial \widehat{M}_k^r(w_r)}{\partial t} + \frac{\widehat{w}_r - w_r}{2}\int_{\widehat{I}_r} \frac{\partial^2 \widehat{M}_k^r(w)}{\partial t^2} dw,$$

with $\widehat{I}_r = [\min(\widehat{w}_r, w_r), \max(\widehat{w}_r, w_r)]$, so that

$$|\widehat{w}_r - w_r| = |\mathbb{E}(e^{w_r Z_1^r}) - \widehat{M}_k^r(w_r)| \left[\frac{\partial \widehat{M}_k^r(w_r)}{\partial t} + \frac{1}{2}\int_{\widehat{I}_r} \frac{\partial^2 \widehat{M}_k^r(w)}{\partial t^2} dw\right]^{-1}$$

$$\leq |\mathbb{E}(e^{w_r Z_1^r}) - \widehat{M}_k^r(w_r)| \left[\frac{\partial \widehat{M}_k^r(w_r)}{\partial t}\right]^{-1} \quad \text{remark that } \frac{\partial^2 \widehat{M}_k^r(w)}{\partial t^2} > 0.$$

For any $L > 0$,

$$\mathbb{P}(|\widehat{w}_r - w_r| > u) = \mathbb{P}\left(|\mathbb{E}(e^{w_r Z_1^r}) - \widehat{M}_k^r(w_r)| > u\frac{\partial \widehat{M}_k^r(w_r)}{\partial t}\right)$$

$$\leq \mathbb{P}(|\widehat{M}_k^r(w_r) - \mathbb{E}(e^{w_r Z_1^r})| > uL) + \mathbb{P}\left(\left|\frac{\partial \widehat{M}_k^r(w_r)}{\partial t}\right| \leq L\right).$$

Denote $\alpha_r(t) = \mathbb{E}\left(\frac{\partial \widehat{M}_k^r(t)}{\partial t}\right) = \mathbb{E}(Z_1^r e^{tZ_1^r}) = \frac{\partial}{\partial t}\mathbb{E}(e^{tY_r})$ and take $L = \frac{\alpha_r(w_r)}{2}$.

Note that $\alpha_r(w_r) > 0$ because of the convexity of $\mathbb{E}(e^{tZ_1^r})$ and the fact that w_r is the unique positive solution to $\mathbb{E}(e^{tZ_1^r}) = 1$. Hence,

$$\mathbb{P}\left(\left|\frac{\partial \widehat{M}_k^r(w_r)}{\partial t}\right| \leq L\right) \leq \mathbb{P}\left(\left|\frac{\partial \widehat{M}_k^r(w_r)}{\partial t} - \alpha_r(w_r)\right| > \frac{\alpha_r(w_r)}{2}\right)$$

$$\leq \frac{16(Cr + 3)\mathbb{E}(Y_r^3 e^{3w_r Y_r})^{\frac{2}{3}}}{k\alpha_r(w_r)^2(1 - \theta^{\frac{1}{6}})}$$

(we proceed as in the proof of Proposition 3.9). Finally,

$$(3.1) \quad \mathbb{P}(|\widehat{w}_r - w_r| > v) \leq \frac{16(Cr + 3)}{\alpha_r(w_r)^2 k(1 - \theta^{\frac{1}{6}})}(\mathbb{E}(e^{3w_r Y_r})^{\frac{2}{3}}\frac{1}{v^2} + \mathbb{E}(Y_r^3 e^{3w_r Y_r})^{\frac{2}{3}}).$$

Remark that $\alpha_r(w_r) = rE(e^{w_r Y_r})c_r'(w_r)$. Consider an interval $[u_1, u_2] \subset [0, u_0[$ such that $w^d \in [u_1, u_2]$, $3u_2 < u_0$, c is non decreasing on $[u_1, u_2]$, $c(u_1) < 0$, $c(u_2) > 0$ (this exists because of the convexity of the function c). Since c_r converges uniformly to c and w_r converges to w^d, for r large enough, $w_r \in [u_1, u_2]$, and

$$E(e^{3w_r Y_r}) \leq e^{r(\varepsilon + c(3u_2))}.$$

We also have that for r large enough, $\alpha_r(w_r) \geq \dfrac{1 - e^{r(c(u_1) - \varepsilon)}}{u_2 - u_1}$. By taking $r(k) = o(\ln k)$, we have that $\widehat{w}_r$ goes to w^d in probability. $\square$

4. SIMULATIONS

We conclude by giving some simulation results. We present some models for which the adjustment coefficient is computable - namely MA and AR linear processes with an innovation following an exponential law. We provide also a non linear example. We refer to [5] for non linear and computable examples. These examples are also more realistic from an actuarial point of view.

Recall that if ξ_i follows an exponential law with parameter $\theta > 0$ then for $0 \leq t < \theta$,

$$\mathbb{E}(e^{t\xi_i}) = \frac{\theta}{\theta - t}.$$

Let $\varepsilon_i = \xi_i - c$, with $c\theta > 1$. Then, the independent adjustment coefficient w^i is the positive solution to :

$$e^{-tc}\frac{\theta}{\theta - t} = 1.$$

The simulation results are summarized below. The graphs represent the estimator $\log \widehat{m}_k(t)$ and $\frac{1}{r}\log \widehat{M}_k^r(t)$ of $\lambda(t)$ and $c_r(t)$ respectively in grey and in black.

4.1. **Independent case.** We have simulated an independent sample of $\varepsilon_i = \xi_i - c$ of length 10000 and $\theta = 1.2$, $c = 1$. We have computed $\widehat{w}^i$ and $\widehat{w}^d$:

$r = 6$	$w^i = w^d = 0.38$	$\widehat{w}^i = 0.36$	$\widehat{w}^d = 0.37$
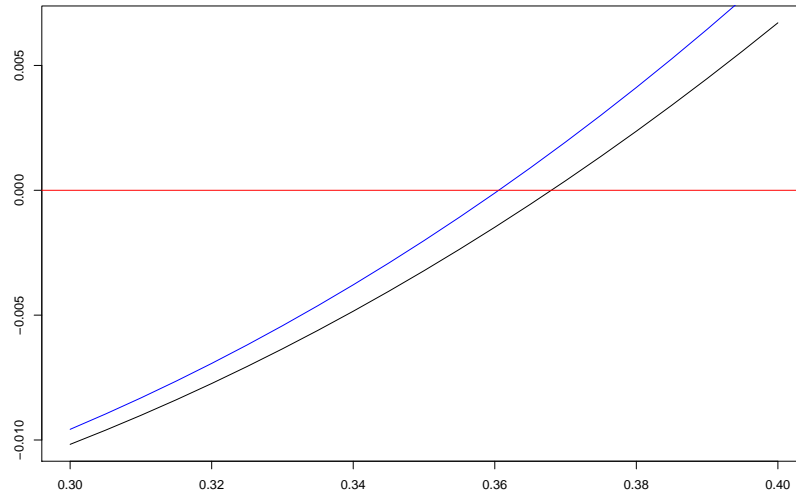
4.2. $AR(1)$ **model.** We consider the following $AR(1)$ model: $X_n = aX_{n-1} + \varepsilon_n$. Following [13], we have that $w^d = (1-a)w^i$. We have simulated a sample of length 10000 for $\theta = 1.2$, $c = 1$, $a = 0.3$. Then,

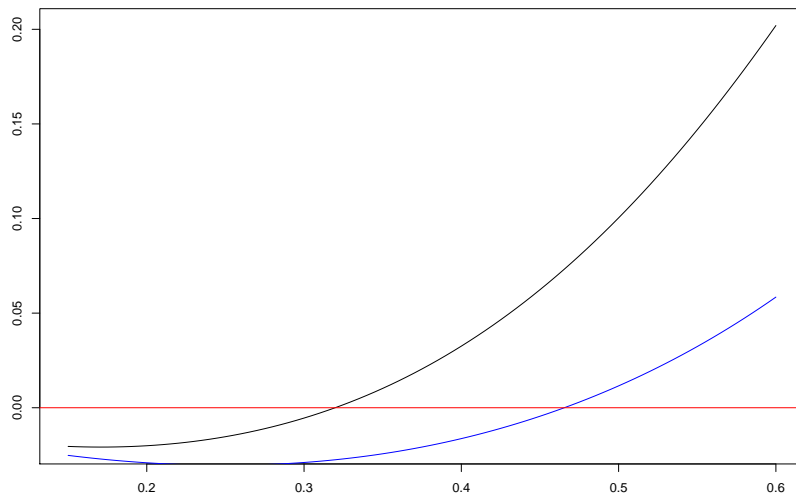| $r = 6$	$\widehat{w}^i = 0.45$	$w^d = 0.26$	$\widehat{w}^d = 0.27$
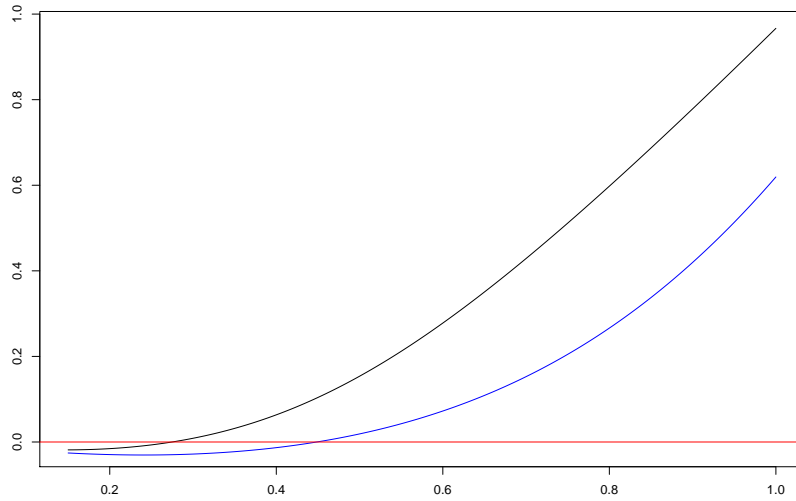
4.3. $MA(1)$ model.

We consider the following $MA(1)$ model: $X_n = \varepsilon_n + a\varepsilon_{n-1}$, with $\theta = 1.2$, $c = 1$, $a = 0.2$. Then, w^d is the positive solution to:

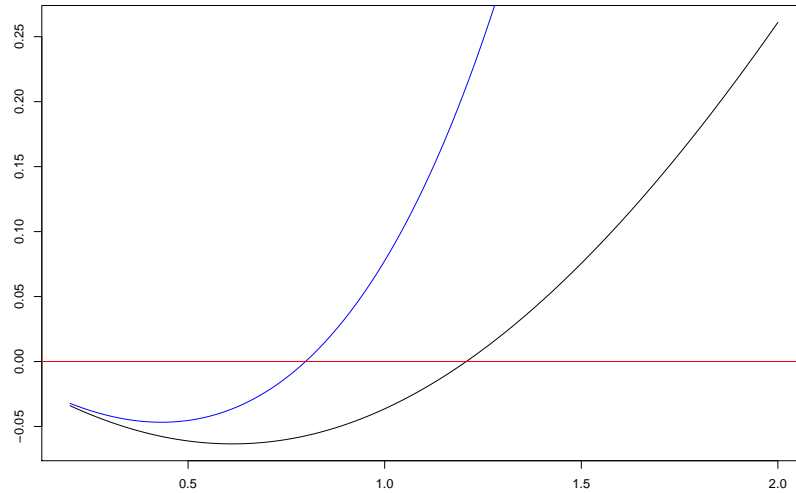$$-tc(1 + a) + \ln\theta - \ln(\theta - t(1 + a)) = 0.$$

We have simulated a sample of length 10000,

$r = 6$	$\widehat{w}^i = 0.47$	$w^d = 0.31$	$\widehat{w}^d = 0.32$

4.4. A non linear $AR(1)$ model.

We consider the following non linear $AR(1)$ model (which may be seen as a particular case of Bernoulli shifts, see [7]): $X_n = aX_{n-1}^2 + 0.7\varepsilon_n$. We have simulated a sample of length 10000, with $\theta = 1.2$, $c = 1$, $a = -0.2$.

$r = 6$	$\widehat{w}^i = 0.8$	$\widehat{w}^d = 1.21$

4.5. **How to choose the r-parameter ?** When performing the estimation of the w^d coefficient, we are faced with the choice of the parameter r. Following Theorem 3.10, we should take $r = o(\ln(k))$ but the practical choice of r for n given is not clear. We have performed several simulations for the independent, $MA(1)$, $AR(1)$, non linear $AR(1)$ models, for several values of r. These experiments tend to show that when r increases, the estimator $\widehat{w}^d$ behaves monotonically in the beginning and then has a more chaotic behavior. We propose to choose r as the largest integer for which $\widehat{w}^d$ is monotonic on $[0, r]$. This is illustrated in the graphs below for several models.

4.5.1. *Independent case.* We have simulated an independent sample of $\varepsilon_i = \xi_i - c$ of length 10000 and $\theta = 1.2$, $c = 1$. Below is represented $\widehat{w}^d$ for $r = 1, \ldots, 35$, $w^i = w^d = 0.38$.
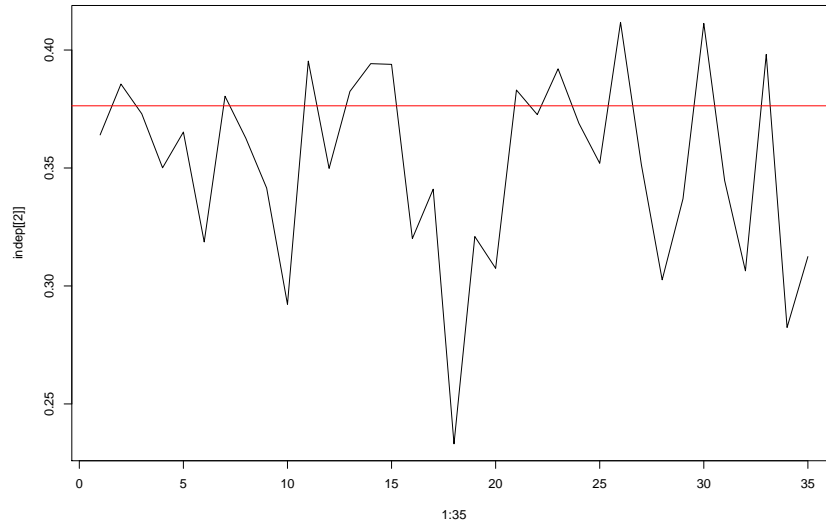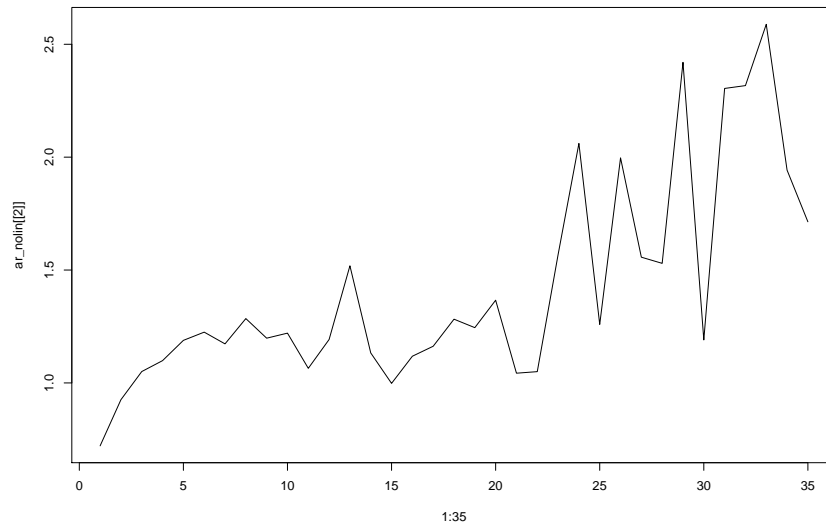
4.5.2. *Linear $MA(1)$*. We consider the following $MA(1)$ model: $X_n = \varepsilon_n + a\varepsilon_{n-1}$, with $\theta = 1.2$, $c = 1$, $a = 0.3$. We have simulated a sample of size 10000 and represented below $\widehat{w}^d$ for $r = 1, \ldots, 40$, $w^d = 0.26$.
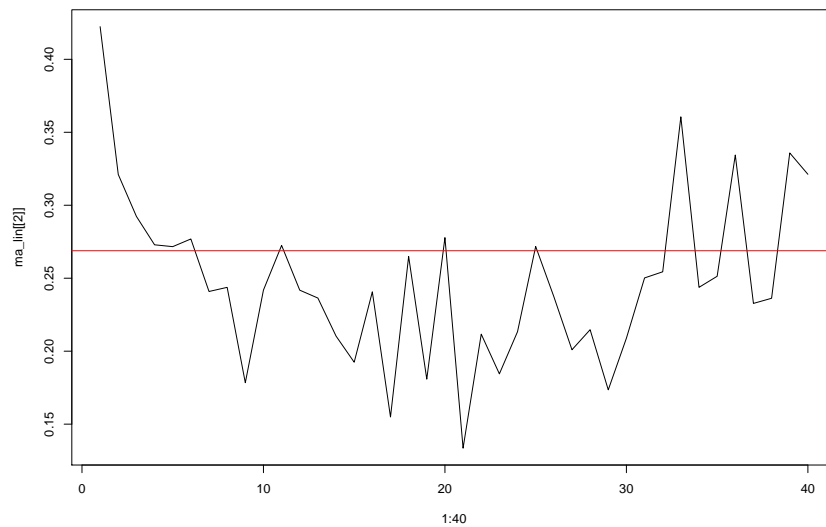
4.5.3. *Linear $AR(1)$*. We consider the following $AR(1)$ model: $X_n = aX_{n-1} + \varepsilon_n$. We have simulated a sample of length 10000 for $\theta = 1.2$, $c = 1$, $a = 0.4$, $w^d =$.
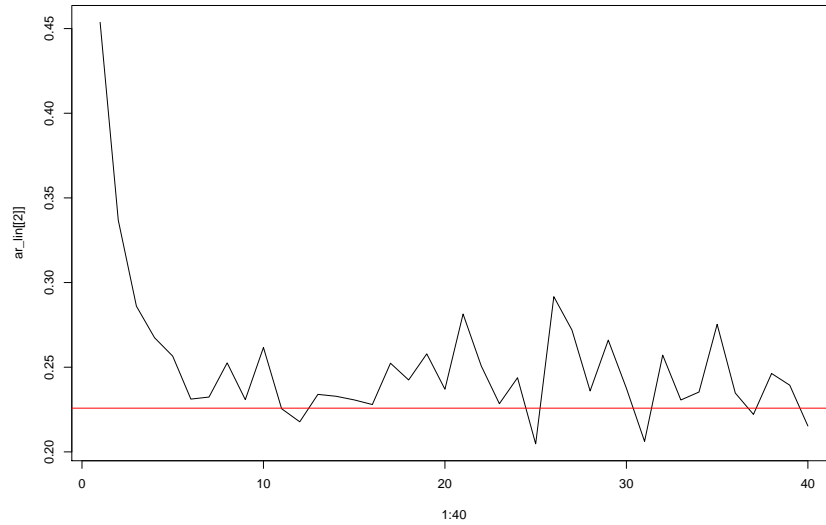
4.6. **Non linear** $AR(1)$. We have represented below $\widehat{w}^d$ for $r = 1, \ldots, 35$ for the non linear $AR(1)$ model of section 4.4.

4.7. **On the empirical distribution of** $\widehat{w}^d$. We conclude this simulation section with a short study of the empirical distribution of $\widehat{w}^d$. We have performed 100 simulations of a sample of size 10000 in the $MA(1)$ model (Section 4.3). The mean value of $\widehat{w}^d$ is 0.317, with standard deviation 0.04. The computed value of w^d is 0.314. The histogram and a Shapiro test indicate that the distribution of $\widehat{w}^d$ is probably asymptotically normal.

```
Shapiro-Wilk normality test
W = 0.9871, p-value = 0.4462
```

Histogram of comp_r[[2]]
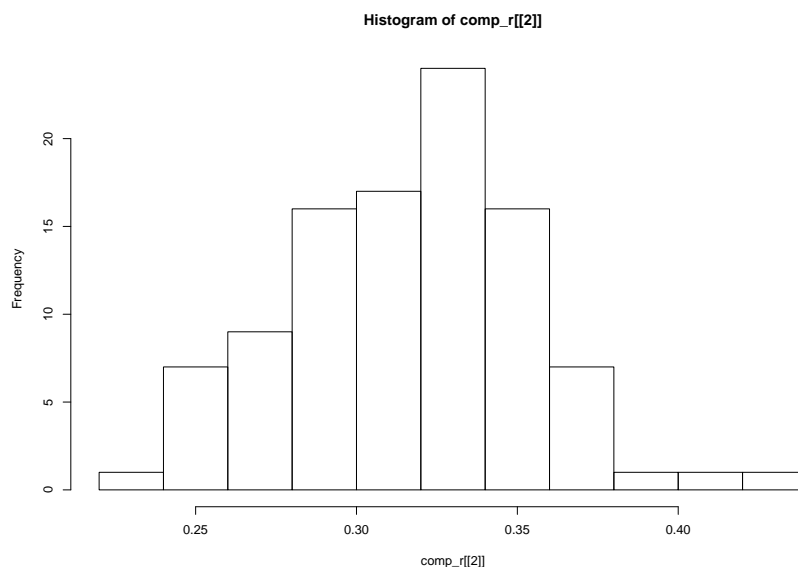
References

[1] W. Bric, Dembo, *Large deviations and strong mixing*, Ann. IHP, B, (1996), 32, 549-569.

[2] R. Christ, J. Steinebach, *Estimating the adjustment coefficient in an $ARMA(p,q)$ risk model.* Insurance: Mathematics and Economics, (1995), 17, 149-161.

[3] Cossette, H., Landriault, D. & E. Marceau (2003). *Ruin probabilities in the compound Markov binomial model*, Scandinavian Actuarial Journal, 301-323.

[4] H. Cossette, D. Landriault & E. Marceau *Compound binomial risk model in a markovian environment.* Insurance: Mathematics and Economics 35, (2004), 425-443.

[5] H. Cossette, E. Marceau, V. Maume-Deschamps

[6] J. Dedecker, P. Doukhan, *A new covariance inequality and applications.* Stochastic Process. Appl. 106 (2003), no. 1, 63–80.

[7] J. Dedecker, P. Doukhan, G. Lang, JR. León R., S. Louhichi, C. Prieur, *Weak Dependence: With Examples and Applications* Lect. Notes Stat. **190**. (2007).

[8] J. Dedecker, C. Prieur, *New dependence coefficients. Examples and applications to statistics.* Prob. Th. Rel. Fields, **132** (2005), 203-236.

[9] P. Doukhan, S. Louhichi, *A new weak dependence condition and applications to moment inequalities.* Stochastic Process. Appl. 84 (1999), no. 2, 313–342.

[10] P. Doukhan, P. Neumann

[11] A. Galves, V. Maume-Deschamps, B.Schmitt *Exponential inequalities for VLMC empirical trees.* ESAIM Prob. Stat., (2008), 12, 119–229.

[12] H.U. Gerber, *An introduction to Mathematical Risk Theory* Huebner Foundation monograph n°8, Irwin Homewood IL (1979).

[13] H.U. Gerber, *Ruin theory in the linear model*, Insurance: Mathematics and Economics, (1982), 1, 177-184.

[14] J.M. Hammersley, *Generalization of the fundamental theorem on subadditive functions*, Math. Proc. Camb. Philos. Soc., (1962), 58, 235-238.

[15] V. Mammitzsch *A note on the adjustment coefficient in ruin theory*, Insurance: Mathematics and Economics, 5, (1986), 147-149.

[16] V. Maume-Deschamps *Exponential inequalities and estimation of conditional probabilities* in *Dependence in probability and statistics*, Lect. notes in Stat., Springer, Vol. 187 Bertail, Patrice; Doukhan, Paul; Soulier, Philippe (Eds.), (2006).

[17] V. Maume-Deschamps *Exponential inequalities and functional estimations for weak dependent data ; applications to dynamical systems.* Stochastics and Dynamics 6, no 4, (2006), 535-560.

[18] A. Müller, G. Pflug *Asymptotic ruin probabilities for risk processes with dependent increments.*, Insurance: Mathematics and Economics, 28, (2001), 381-392.

[19] S.M. Pitts, R. Grübel, P. Embrechts, *Confidence bounds for the adjustment coefficient*, Ad. in Applied probability, 28 n°3, (1996), 802-827.

[20] A.W. van der Vaart *Asymptotic Statistics*, Cambridge Series in Statistical and Probabilistic Mathematics, (1998).
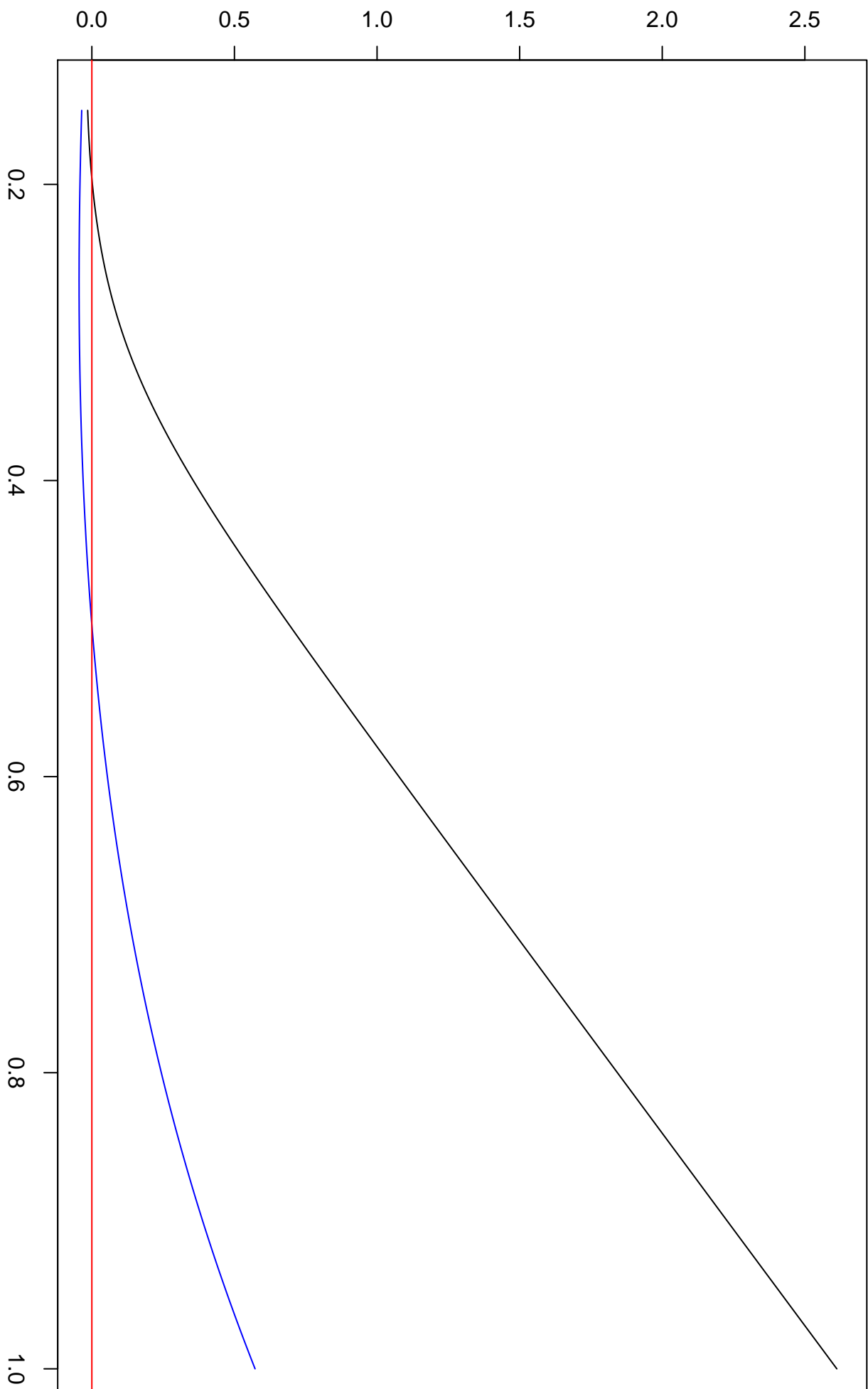
Université Laval
E-mail address: `helene.cossette@act.ulaval.ca`
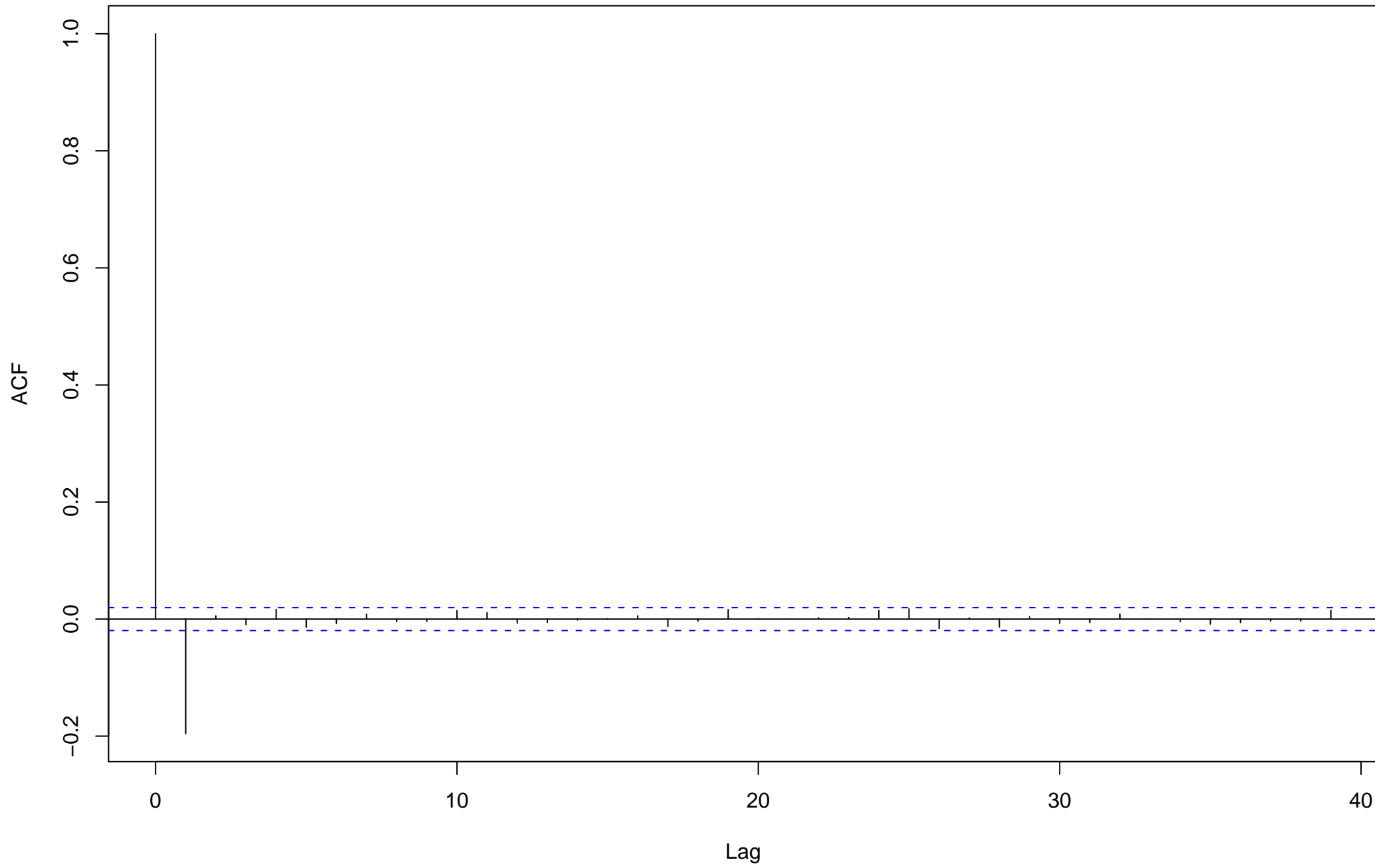*E-mail address*: `etienne.marceau@act.ulaval.ca`

Université de Lyon, Université Lyon 1, ISFA, laboratoire SAF
E-mail address: `veronique.maume@univ-lyon1.fr`

Series nonlin_ma

Series nonlin_ma