HAL
open science

Adjustment Coefficient for Risk Processes in Some Dependent Contexts

Hélène Cossette, Etienne Marceau, Véronique Maume-Deschamps

To cite this version:

Hélène Cossette, Etienne Marceau, Véronique Maume-Deschamps. Adjustment Coefficient for Risk Processes in Some Dependent Contexts. Methodology and Computing in Applied Probability, 2010, pp.1. 10.1007/s11009-010-9182-y . hal-00504014

HAL Id: hal-00504014

https://hal.science/hal-00504014v1

Submitted on 19 Jul 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ADJUSTMENT COEFFICIENT FOR RISK PROCESSES IN SOME DEPENDENT CONTEXTS

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

Abstract. Following Müller and Pflug [18] and Nyrhinen [19, 20], we study the adjustment coefficient of ruin theory in a context of temporal dependency. We provide a consistent estimator for this coefficient, and perform some simulations.

1. Introduction

We consider $(Y_n)_{n\in\mathbb{N}}$ a sequence of random variables and R_u the event $\{Y_n > u \text{ for some } n \geq 1\}$. Y_n is interpreted as the value of the claim surplus process of a company at the end of year n. R_u is understood as the ruin event for an initial reserve $u > 0$. We write

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

where X_i is the loss/gain of the company during year i, such that $\mathbb{E}(X_i) < 0$ to satisfy the solvency condition. If U_n denotes the total reserve of the company at the end of year n, then $U_n = u - Y_n$, and the ruin event corresponds to the event that for some n the reserve U_n is negative.
In the classical risk model, the random variables (r.v.'s) X_i $(i = 1, \ldots, n)$ are assumed independent and identically distributed (i.i.d.). The adjustment coefficient $w > 0$ is defined as the unique positive solution to $\lambda(w) = 0$ with

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

assumed to be well defined. Mammitzsch ([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 ([11]). Let T_u be the ruin time for an initial reserve $u > 0$, $T_u = \inf\{k \in \mathbb{N} \ / \ Y_k > u\}$ and $T_u = \infty$ if ruin does not occur. Then, the expression for the ruin probability is given by

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

2000 *Mathematics Subject Classification.* 37A50, 60E15, 37D20.

Key words and phrases. Adjustement coefficient, risk process, ruin theory, non parametric estimation, weak dependence.

The research was financially supported by the Natural Sciences and Engineering Research Council of Canada and the Chaire d'actuariat de l'Université Laval. This work has been partially supported by the French Research National Agency (ANR) under the reference ANR-08-BLAN-0314-01. The authors would like to thank two anonimous referees whose comments helped to improve the redaction of the paper.

with $U_{T_u} = u - Y_{T_u}$, 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 :

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

Based on a general dependent setting considered by Nyrhinen [19] and Müller and Pflug [18] in which such an asymptotic result holds, we provide a consistent estimator for the adjustment coefficient. Our approach does not require a precise knowledge of the dependence structure, nor of the law of $(X_i)_{i\in\mathbb{N}}$ but only an information on the speed of mixing (given by Equation (2.1)). When it is defined, we consider

$$c_n(t) := \frac{\log \mathbb{E}\left[\exp(tY_n)\right]}{n}. \tag{1.1}$$

Following Nyrhinen [19] and Müller and Pflug [18], we assume that there exists a $t_0 > 0$ such that for all $0 < t < t_0$,

$$c(t) := \lim_{n\to\infty} c_n(t) \text{ exists.} \tag{1.2}$$

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

We shall provide a sufficient condition that implies the existence and uniqueness of a positive solution to $c(t) = 0$, provided that (1.2) is satisfied. We shall denote by 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 $(X_i)_{i\in\mathbb{N}}$ is i.i.d., then $w^i = w^d$.

Several results for sums of i.i.d. claims exist: Gerber [11] gave an exact formula for finite time ruin probabilities involving the adjustment coefficient w, Pitts et al. [21] provide a consistent and asymptotically normal estimator of w, V. Mammitzsch [15] gave a necessary and sufficient condition for the existence of w. In dependent contexts, let us cite Gerber [12] for auto-regressive processes, Christ and Steinebach [3] for an extension to ARMA processes and Cossette et al. [4] for the study of the adjustment coefficient in Markovian environments. The main objective of the present paper is to provide a non parametric estimation of the adjustment coefficient introduced in Nyrhinen [19, 20] and Müller and Pflug [18] in dependent contexts. We give a general dependent context (weak temporal dependency in the sense of Dedecker et al. [7]) for which our estimator is consistent.
The paper is organized as follows :

- Section 2 contains the definitions and elementary properties of weak dependent processes as well as for the adjustment coefficient. To make short, w^i 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 3, we prove that w^d may be seen as a limit (for $r \to \infty$) of independent coefficients w_r^i. We also provide some general examples for which the adjustment coefficient w^d may be defined.
- Section 4 is devoted to the estimation of coefficients w^i and w^d and contains the main results (see Theorems 4.3, 4.5 and 4.10). Note that in Christ and Steinebach [3], 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 5 we provide some simulations.

2. Weak dependence and adjustment coefficient

2.1. **Weak dependent processes.** This last decade, Dedecker et 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 that are neither α nor Φ-mixing (for example ARMA processes with discrete innovation) in the standard way. Roughly speaking, in the classical probabilistic definitions of mixing, the functions f and g in the definition below (2.1) belong to the whole class of square integrable functions.
Given a non negative functional $c(f,g)$, we define, for a real valued or vector valued process $(X_t)_{t\in\mathbb{N}}$, the auto-correlation function as

$$\varepsilon(k) := \sup \frac{|\mathrm{Cov}(f(X_{i_1},\dots,X_{i_\ell}), g(X_{j_1},\dots,X_{j_m}))|}{c(f,g)} \tag{2.1}$$

(see Dedecker et al. [7]). In this definition, the supremum is taken over multi-indices $\mathbf{i} = (i_1,\dots,i_\ell)$, $\mathbf{j} = (j_1,\dots,j_m)$ such that

$$i_1 < \cdots < i_\ell \le i_\ell + k \le j_1 < \cdots < j_m$$

and that all functions $f : \mathbb{R}^\ell \longrightarrow \mathbb{R}$, $g : \mathbb{R}^m \longrightarrow \mathbb{R}$ are bounded and Lipschitz functions, with respect to the distance on $\mathbb{R}^p$ (with $p = \ell$ or $p = m$):

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

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

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

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

(1) $c(f,g) = m\|f\|_\infty \mathrm{lip}(g)$, 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) $c(f,g) = \ell \mathrm{lip}(f)\|g\|_\infty + m\|f\|_\infty \mathrm{lip}(g)$, 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.

Remark. It follows from the definitions above that θ-weakly dependent processes are η-weakly dependent processes as well.

This class of dependent processes is very rich and enjoys lots of nice probabilistic properties. Let us remark that weak dependent processes need not be stationary.
For completeness, we recall the definitions of α, Φ and Ψ-mixing coefficients (see Bradley [1] for a review on these strong mixing coefficients). Ψ-mixing coefficient 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 meaningful.

Definition 2. Consider mixing coefficients $\varepsilon(k)$ defined by Equation (2.1) where the supremum is taken over functions

$$f \ : \ \mathbb{R}^u \ \longrightarrow \ \mathbb{R} \text{ and } g \ : \ \mathbb{R}^v \ \longrightarrow \ \mathbb{R}$$

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 Ψ-mixing coefficients $\varepsilon(k)$ go to zero, taking f and g indicator functions lead to Bradley's ([1]) definition of Ψ-mixing.
We define α and Φ-mixing coefficients by:

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

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

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

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

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

go to 0.

Let us note that the above definition of Ψ-mixing is equivalent to the classical one. This follows from the fact that the bounded measurable functions are limits of simple functions.

2.2. **Adjustment coefficient: existence condition.** In Mammitzsch ([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.

(E1) $\mathbb{E}(X_1) < 0$,
(E2) $\mathbb{P}(X_1 > 0) > 0$, and
(E3) either $a < \infty$ and $\mathbb{E}(e^{aX_1}) \ge 1$ or $a = \infty$ where

$$a := \sup\{t \ge 0,\ \mathbb{E}(e^{tX_1}) < \infty\}. \tag{2.2}$$

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 $c(t) = 0$) of w^d, we shall need some additional conditions.

Property 2.1. Assume that the limit $c(t)$ is well defined on $[0, t_0[$, in particular, for all $t < t_0$, for n large enough, $\mathbb{E}(e^{tY_n}) < \infty$. Assume that (E) is satisfied for $0 < a \le t_0$, 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, $\lim_{t \to a-} \mathbb{E}(e^{tY_n}) \geq 1$, and
(4) $c'(0+) < 0$,

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

Proof. We adapt Mammitzsch'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 $[0, t_0[$ and converging to f on $[0, t_0[$, then f is a convex function and the convergence is uniform on any compact subset of $[0, t_0[$.
The function c is the limit in $[0, t_0[$ of convex functions c_n defined by (1.1). Thus c is a convex function with $c(0) = 0$ and we assume that $c'(0+) < 0$. This last condition is equivalent to $\exists\ t > 0$ such that $c(t) < 0$ by convexity. We consider two cases: $a < \infty$ and $a = \infty$.
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 a $w > 0$ 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, $\lim_{t \to \infty} \mathbb{E}(e^{tY_n}) = +\infty$. As a consequence, we have that $\lim_{t \to \infty} c(t) \geq 0$. Since we assume that $c(t) < 0$ for some t, the convexity of c then implies that there exists a 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 a $w > 0$ such that $c(w) = 0$. This solution is unique because of the convexity of c. □

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

3. Limit result and examples

3.1. **Asymptotic behavior for ruin probability.** In Müller and Pflug [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:

$$\lim_{u \to \infty} \frac{\log \mathbb{P}(R_u)}{u} = -w^d. \tag{3.1}$$

As a consequence to (3.1), we obtain that if it exists, w^d is the limit of the adjustment coefficient of the sequence $(Y_n)_{n \in \mathbb{N}}$.

Corollary 3.1. *Assume that the hypotheses of Property 2.1 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 Mammitzsch [15]: Y_n satisfies the existence hypotheses of Mammitzsch for n large enough. Applying Markov's inequality, we get for all $n > 0$,

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

Then, (3.1) 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, otherwise we would have $c(t) \geq 0$ for all $t \geq 0$ which contradicts the hypotheses of Property 2.1. We conclude that $w_n \overset{n\to\infty}{\longrightarrow} w^d$. □

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 [13].

Lemma 3.2. *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 :

$$\sum_{r=1}^{\infty} \frac{\Delta(r)}{r(r+1)} < \infty. \tag{3.2}$$

Then, $\lambda = \lim_{n\to\infty} \frac{h(n)}{n}$ 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 (3.2) is satisfied. Lemma 3.2 asserts that $\Delta(r)$ may go to infinity but not too fast.

3.2. **Bounded Ψ-mixing coefficients.** According to Definition 2, we consider the following classical Ψ-mixing coefficients :

$$\Psi(k) := \sup \frac{|\mathrm{Cov}(f(X_{i_1}, \ldots, X_{i_\ell}), g(X_{j_1}, \ldots, X_{j_m}))|}{\|f\|_1 \|g\|_1}, \tag{3.3}$$

where the supremum is taken over functions $f,\ g \in L^2$ and over multi-indices $\mathbf{i} = (i_1, \ldots, i_\ell)$ and $\mathbf{j} = (j_1, \ldots, j_m)$ with $i_1 < \cdots < i_\ell \leq i_\ell + k \leq j_1 < \cdots < j_m$.

Proposition 3.3. *Let*

$$S_n = \sum_{i=1}^{n} X_i.$$

Assume that for $t \in [0, a[$, for all $n \in \mathbb{N}$, $\mathbb{E}(e^{2tS_n}) < \infty$ and $(X_n)_{n\in\mathbb{N}}$ a stationary process with bounded Ψ-mixing coefficients, then

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

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

Proof. Using (3.3), 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 3.2. □

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 [10]).
Even if Ψ (Φ, α)-mixing processes are often used in probability theory, lots of useful processes (like ARMA processes with discrete innovations) 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 [2] 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 Dedecker et al. [7] for other examples of θ and η-weakly mixing processes (including ARMA and ARCH processes).

3.3. **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 3.4. *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 an $M > 0$ 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 $S_j^k = \sum_{\ell=j}^{k} X_\ell$ and, 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 $j \leq \ell \leq k$,

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

Remark that for $t \geq 0$, 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:

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

We conclude the proof by choosing $r = O(\frac{n+m}{\ln(n+m)^\kappa})$, $1 < \kappa < \beta$ and applying Lemma 3.2 with a function:

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

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

3.4. **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 Dedecker et al. [7] we may prove that such a process is θ-dependent with mixing coefficient $\theta(n) = d_n$.

Proposition 3.5. *Assume that $(X_n)_{n\in\mathbb{N}}$ is a Bernoulli shift satisfying the summability condition for the continuity coefficients $(d_n)_{n\in\mathbb{N}}$. Then, the sequence $\ln \mathbb{E}(e^{tS_n})$ satisfies the hypothesis of Lemma 3.2 and thus $c(t)$ is well defined.*

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

$$\begin{aligned}
S_i^\ell &= \sum_{j=i}^{\ell} X_j \\
&= \sum_{j=i}^{\ell} H(\xi_j, \ldots, \xi_i, \ldots) \\
&= \underbrace{\sum_{j=i}^{\ell} H(\xi_j, \ldots, \xi_i, u_0, u_{-1}, \ldots)}_{:=U_i^\ell} + \\
&\qquad \underbrace{\sum_{j=i}^{\ell} H(\xi_j, \ldots, \xi_i, \ldots) - H(\xi_j, \ldots, \xi_i, u_0, u_{-1}, \ldots)}_{:=d_i^\ell(u,\xi)}. \qquad (3.4)
\end{aligned}$$

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 (3.4), 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 3.2. □

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

In Dedecker et al. [7], it is mentioned that stationary ARMA processes are examples of Bernoulli shifts. Non linear autoregressive processes may also be examples of Bernoulli shifts. For example, ARMA processes with bounded innovation satisfy our hypotheses. Note that bounded variables might be meaningful for insurers in a "stop-loss" context where the insurer does not cover the claims above a certain fixed amount, these large claims being covered by a reinsurer. Propositions 3.4 and 3.5 give the existence of the function $c(t)$ under few structure hypotheses but are mainly applicable for bounded variables. For processes more specifically defined, the existence of $c(t)$ may be obtained for unbounded variables. We cite below two examples.

3.5. **Some Markov processes.** Following Nyhrinen [19] and Iscoe et al. [14] we recall that for $(X_n, n = 0, 1, \ldots)$ an irreductible, aperiodic Markov chain on a general state space E with σ-field $\mathcal{E}$ and with the transition probability function $\pi(x, A)$, $(x \in E, A \in \mathcal{E})$. The initial state of the chain is assumed to be fixed. Let $g : (E, \mathcal{E}) \to (\mathbb{R}, \mathcal{B})$ be a measurable map and $\xi_n = g(X_n)$. Assume there exists a probability measure ν on $(E, \mathcal{E})$, constants $0 < a \leq b < \infty$ and an integer m_0 such that for all $x \in E$ and $A \in \mathcal{E}$,

$$a\nu(A) \leq \pi^{m_0}(x, A) \leq b\nu(A).$$

Then the limit $c(t)$ exists and is finite for $t \in \mathcal{D}(h)$ with

$$\mathcal{D}(g) = \left\{ t \in \mathbb{R} \ / \ \int_E e^{tg(x)} d\nu(x) < \infty \right\}.$$

Nyhrinen [19] also presents a condition of existence of $c(t)$ for partial sums of moving averages. Indeed, its result can be extended to $MA(\infty)$ processes as follows. Let $(\varepsilon_n)_{n\in\mathbb{Z}}$ be i.i.d. random variables. Let $(a_i)_{i\in\mathbb{N}}$ be a sequence of *re-summable* real numbers, that is : $\sum_{i\in\mathbb{N}} |a_i| < \infty$ and $\sum_{j\in\mathbb{N}} |b_j| < \infty$ where $b_j = \sum_{i\geq j} a_i$. Denote $a = \sum_{i\in\mathbb{N}} a_i$, $X_n = \sum_{i\in\mathbb{N}} a_i\varepsilon_{n-i}$, $c_\varepsilon(t) = \log \mathbb{E}(e^{t\varepsilon_1})$ and $S_n = \sum_{i=1^n} X_i$. Then,

$$c_n(t) = \frac{1}{n}\log(\mathbb{E}(e^{tS_n}) = \frac{1}{n}\left[\sum_{k=0}^{n-1} c_\varepsilon\left(t\sum_{\ell=0}^{n-k-1} a_\ell\right)\right] + \frac{1}{n}\sum_{j\geq 1} c_\varepsilon\left(t\sum_{k=j}^{n+j-1} a_k\right).$$

Because $(a_i)_{i\in\mathbb{N}}$ is re-summable, $\left(c_\varepsilon\left(t\sum_{k=j}^{n+j-1} a_k\right)\right)_{j\geq 1}$ is summable and then $c_n(t) \to c_\varepsilon(at)$.

4. Estimation

4.1. **Definition of estimators.** In this section, we assume that the sequence $(X_i)_{i\geq 1}$ is stationary and that the hypotheses of Proposition 2.1 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 that 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 properties for $(Z_i^r)_{i\in\mathbb{N}}$ which is proven in Lemma 4.4).

Proposition 4.1. *Assume that the sequence $(X_i)_{i\in\mathbb{N}}$ is η or θ-weakly dependent and satisfies (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 Dedecker and Doukhan [6]. Following Section 2.2, 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.

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

4.2. **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 of M-estimators in a parametric context (see van der Vaart [22] section 5). It is known (see Dedecker et al. [7], Douckhan and Louhichi [9]) that provided the sequence $\mathrm{Cov}(X_0, X_i)$, $i \in \mathbb{N}$, is summable, the process $(X_i)_{i\in\mathbb{N}}$ satisfies a central limit theorem with asymptotic variance

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

in other words

$$\sqrt{n}\left(\frac{1}{n}\sum_{i=0}^{n-1} X_i - \mathbb{E}(X_0)\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \Gamma^2) \quad \text{as } n \to \infty.$$

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 4.2. *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

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

As a consequence, we get the following consistency result for $\widehat{w}^i$ as well as asymptotic normality. In order to use Theorem 4.2, 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 4.3. *Assume $(X_n)_{n\in\mathbb{N}}$ is η-weakly dependent with $\varepsilon(n) = O(\theta^n)$, $0 < \theta < 1$. We have for any $t \in [0, t_0[$ that*

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

is well defined on $[0, t_0[$ and

$$\sqrt{n}\left(\widehat{m}_n(t) - \mathbb{E}(e^{tX_0})\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \Gamma^2(t)) \quad \text{as } n \to \infty.$$

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

$$\sqrt{n}\left(\widehat{w}^i - w^i\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \Gamma_i^2) \quad \text{as } n \to \infty,$$

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 4.4. *Assume $(X_n)_{n\in\mathbb{N}}$ is η-weakly dependent with mixing coefficient $\varepsilon(r)$. Then, for any $t \in [0, t_0[$, the sequence of random variables $(e^{tX_n})_{n\in\mathbb{N}}$ is η-weakly dependent with mixing coefficient $\varepsilon_t(r) \leq 2[t + \mathbb{E}(e^{(t+\kappa)X_1})]\varepsilon(r)^{\frac{\kappa}{t+\kappa}}$ with $\kappa > 0$ such that $t + \kappa \in [0, t_0[$.*

Proof. We follow the proof of Proposition 2.1 in Dedecker et al. [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, we have

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

Recall that $\mathbb{E}(e^{sX_i}) < \infty$ for any $s \in [0, t_0[$ and let $\kappa > 0$ be such that $t + \kappa \in [0, t_0[$. Then $\mathbb{E}(|e^{(t+\kappa)X_i}|) < \infty$ and, using the Markov inequality, we get

$$\begin{aligned} \mathbb{E}\left(e^{tX_j} - e^{tX_j^{(M)}}\right) &\leq 2\mathbb{E}\left(e^{tX_j}\mathbf{1}_{\{X_j > M\}}\right) \\ &\leq 2e^{-\kappa M}\mathbb{E}\left(e^{(t+\kappa)X_j}\right). \end{aligned}$$

This leads to:

$$\begin{aligned} \mathbb{E}(|G(X_{\mathbf{j}}) - G^{(M)}(X_{\mathbf{j}})|) &\leq \mathrm{lip}(g)\sum_{k=1}^{v}\mathbb{E}(|e^{tX_{j_k}} - e^{tX_{j_k}^{(M)}}|) \\ &\leq 2v\mathrm{lip}(g)e^{-\kappa M}\mathbb{E}(e^{(t+\kappa)X_1}). \end{aligned}$$

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

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

Finally, we obtain for $M \geq 1$

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

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

Proof of Theorem 4.3. Lemma 4.4 together with Theorem 4.2 imply that for any $t \in [0, t_0[$,

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

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 van der Vaart [22], 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, it follows that

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

We have that

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

and therefore it is asymptotically normal with mean zero 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

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

□

Remark. We could relax the hypothesis that $\varepsilon(r)$ decreases to 0 exponentially fast. It is sufficient to have $\varepsilon(r)^{\frac{\kappa}{t+\kappa}} = O(r^{-2-\alpha})$ 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.

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

Theorem 4.5. *Under the assumptions of Theorem 4.3 there exists a sequence $r = r(k) \overset{k\to\infty}{\longrightarrow} \infty$ such that $\widehat{w}^d$ converges in probability to w^d, as k goes to infinity.*

Theorem 4.5 will be proven by rewriting 4.3 for $\widehat{w}_r$ instead of $\widehat{w}^i$ and then by using Corollary 3.1. We only need to prove that the sequence $(Z_i^r)_{i\in\mathbb{N}}$ satisfies some weak dependence property.

Lemma 4.6. *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

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

Then, we have

$$\begin{aligned}
&\mathrm{Cov}(f(Z_{i_1}^r,\ldots,Z_{i_u}^r), g(Z_{j_1}^r,\ldots,Z_{j_v}^r)) \\
&= \mathrm{Cov}\Big(\widetilde{f}(X_{i_1 r+1},\ldots,X_{i_1(r+1)}, X_{i_2 r+1},\ldots,X_{(i_u+1)r}), \\
&\quad \widetilde{g}(X_{j_1 r+1},\ldots,X_{j_1(r+1)}, X_{j_2 r+1},\ldots,X_{(j_v+1)r})\Big) \\
&\leq r\varepsilon((k-1)r)\left(u\mathrm{lip}f\|g\|_\infty + v\|f\|_\infty \mathrm{lip}f\right),
\end{aligned}$$

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

As a corollary to Lemma 4.6, we deduce that for any r, $\widehat{w}_r$ exists eventually almost surely.

Corollary 4.7. *Assume that the hypotheses of Property 2.1 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 4.1 and Lemma 4.6. □

Proof of Theorem 4.5. From Lemma 4.6 and Theorem 4.3, we get

$$\sqrt{k}\left(\widehat{M}_k^r(t) - \mathbb{E}(e^{tZ_1^r})\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0,\Gamma_r^2(t)) \quad \text{as } k\to\infty$$

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 independent adjustment coefficient of the sequence Z_0^r and $\widehat{w_r}$ the positive solution to $\widehat{M}_k^r(t) = 1$,

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

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 4.5 by using Corollary 3.1 : there exists a sequence $r(k) \overset{k\to\infty}{\longrightarrow} \infty$ such that $\widehat{w}_r$ converges to w^d in probability. □

Theorem 4.5 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.
The following lemma is required in the proof of Proposition 4.9.

Lemma 4.8. *Let $(W_i)_{i\in\mathbb{N}}$ be a sequence of centered random variables and*

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

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

where $S_n = \sum_{i=1}^{n} W_i$.
As a consequence, if W_i is η-dependent and stationary, with mixing coefficient $\varepsilon(k) \le C\theta^k$, we have that for any $m > 2$ such that $M_m = \mathbb{E}(|W_i|^m) < \infty$,

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

Proof. The first part of the proof of Lemma 4.8 may be found in Dedecker et al. [7] (see Lemma 4.6 p.79). For the second part, we proceed as in the proof of Lemma 4.4 : let $W_t^{(M)} = \max(\min(W_t, M), -M)$, (so that $W_t^{(M)} = W_t$ provided that $|W_t| \le M$). Then,

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

where the last line is obtained by noting that

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

Similarly, we obtain

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

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

Proposition 4.9. *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 $t \in [0, t_0[$, such that $3t \leq t_0$, for any $v > 0$, we have:*

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

Proof. We apply the Bienaimé-Tchebitchev inequality and Lemma 4.8 with $m = 3$. Let

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

Following the lines of the proof of Lemma 4.8, 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 4.8, 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 4.10. *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 $3w^d < t_0$. Then, for $r = r(k) = o(\ln k)$, $\widehat{w}_r$ goes to w^d in probability.*

Proof. We have that

$$\begin{aligned}
\mathbb{E}(e^{w_r Z_1^r}) - \widehat{M_k^r}(w_r) &= \widehat{M_k^r}(\widehat{w}_r) - \widehat{M_k^r}(w_r) \\
&= (\widehat{w}_r - w_r)\frac{\partial \widehat{M_k^r}(w_r)}{\partial t} + \frac{\widehat{w}_r - w_r}{2}\int_{\widehat{I}_r} \frac{\partial^2 \widehat{M_k^r}(w)}{\partial t^2} dw,
\end{aligned}$$

with $\widehat{I}_r := [\min(\widehat{w}_r, w_r), \max(\widehat{w}_r, w_r)]$. It follows that

$$\begin{aligned}
|\widehat{w}_r - w_r| &= |\mathbb{E}(e^{w_r Z_1^r}) - \widehat{M_k^r}(w_r)| \left[\frac{\partial \widehat{M_k^r}(w_r)}{\partial t} + \frac{1}{2}\int_{\widehat{I}_r} \frac{\partial^2 \widehat{M_k^r}(w)}{\partial t^2} dw\right]^{-1} \\
&\leq |\mathbb{E}(e^{w_r Z_1^r}) - \widehat{M_k^r}(w_r)| \left[\frac{\partial \widehat{M_k^r}(w_r)}{\partial t}\right]^{-1}.
\end{aligned}$$

Remark that $\dfrac{\partial^2 \widehat{M}_k^r(w)}{\partial t^2} > 0$. For any $L > 0$, we have

$$\begin{aligned}\mathbb{P}(|\widehat{w}_r - w_r| > u) &= \mathbb{P}\left(|\mathbb{E}(e^{w_r Z_1^r}) - \widehat{M}_k^r(w_r)| > u\frac{\partial \widehat{M}_k^r(w_r)}{\partial t}\right)\\ &\leq \mathbb{P}(|\widehat{M}_k^r(w_r) - \mathbb{E}(e^{w_r Z_1^r})| > uL) + \mathbb{P}\left(\left|\frac{\partial \widehat{M}_k^r(w_r)}{\partial t}\right| \leq L\right).\end{aligned}$$

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, we obtain

$$\begin{aligned}\mathbb{P}\left(\left|\frac{\partial \widehat{M}_k^r(w_r)}{\partial t}\right| \leq L\right) &\leq \mathbb{P}\left(\left|\frac{\partial \widehat{M}_k^r(w_r)}{\partial t} - \alpha_r(w_r)\right| > \frac{\alpha_r(w_r)}{2}\right)\\ &\leq \frac{16(Cr+3)\mathbb{E}(Y_r^3 e^{3w_r Y_r})^{\frac{2}{3}}}{k\alpha_r(w_r)^2(1-\theta^{\frac{1}{6}})}\end{aligned}$$

(we proceed as in the proof of Proposition 4.9). Finally, we get

$$\mathbb{P}(|\widehat{w}_r - w_r| > v) \leq \frac{16(Cr+3)}{\alpha_r(w_r)^2 k(1-\theta^{\frac{1}{6}})}(\mathbb{E}(e^{3w_r Y_r})^{\frac{2}{3}}\frac{1}{v^2} + \mathbb{E}(Y_r^3 e^{3w_r Y_r})^{\frac{2}{3}}). \tag{4.1}$$

Remark that $\alpha_r(w_r) = rE(e^{w_r Y_r})c_r'(w_r)$. Consider an interval $[t_1, t_2] \subset [0, t_0[$ such that $w^d \in [t_1, t_2]$, $3t_2 < t_0$, c is non decreasing on $[t_1, t_2]$, $c(t_1) < 0$, $c(t_2) > 0$ (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, $w_r \in [t_1, t_2]$, and

$$E(e^{3w_r Y_r}) \leq e^{r(\varepsilon + c(3t_2))}.$$

We also have that for r large enough, $\alpha_r(w_r) \geq \dfrac{1 - e^{r(c(t_1)-\varepsilon)}}{t_2 - t_1}$. By taking $r(k) = o(\ln k)$, we have that $\widehat{w}_r$ goes to w^d in probability. □

5. Simulations

We conclude by providing 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 Cossette et al. [5] for non linear and computable examples. The examples of that later article are also more realistic from an actuarial point of view.

We consider an auxiliary random variable ξ_i with exponential law. 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 associated to $(\varepsilon_i)_{i\in\mathbb{N}}$ is the positive solution to :

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

We shall perform simulations using independent copies of ε_i as innovations of auto regressive and moving average models.
The simulation results are summarized below. The simulations have been done on a standard laptop running under `Linux`. The scripts have been written in `R`. The computations may be done within reasonable time. Indeed, in the case of the non linear $AR(1)$ model (see section 5.1.4), in order to simulate one sample of size 10000 and to perform the estimation of $w^d(r)$ for $r = 1, \ldots, 15$, it tooks $8s$.

5.1. **Large samples.** For each model (independent, linear $AR(1)$ and $MA(1)$) we produce large samples from which we compute the estimators $\widehat{w}^i$ and $\widehat{w}^d$. We provide 50 simulations of length 10000, we also consider the results for several values of r (for $r = 1$ to $r = 10$). This study gives indications on the most appropriate value for the parameter r in order to calculate the estimator $\widehat{w}^d$.

5.1.1. *Independent case.* We have simulated independent samples of $\varepsilon_i = \xi_i - c$ of length 10000 and $\theta = 1.2$, $c = 1$. In that case, we know that $w^i = w^d = 0.37$. The table below gives the mean of the estimated value $\widehat{w}^d(r)$, the standard deviation and square root of the mean squared error over the 50 samples, for r going from 1 to 15.

r	1	2	3	4	5	6	7	8	9
mean	0.363	0.367	0.366	0.367	0.367	0.371	0.374	0.378	0.38
deviation	0.036	0.029	0.036	0.04	0.036	0.039	0.043	0.043	0.04
$\sqrt{\text{mse}}$	0.037	0.031	0.038	0.042	0.037	0.04	0.043	0.043	0.039
r	10	11	12	13	14	15			
mean	0.384	0.383	0.388	0.382	0.389	0.394			
deviation	0.043	0.05	0.039	0.048	0.047	0.059			
$\sqrt{\text{mse}}$	0.043	0.046	0.04	0.048	0.048	0.061			

Theoretically, we should find values very close to 0.36 for all r (because of independence). We remark that the result may be considered as correct for $r = 1$ to $r = 6$. This is consistent with the theoretical condition that r has to be $o(\ln n)$.

5.1.2. $AR(1)$ *model.* We consider the following $AR(1)$ model: $X_n = aX_{n-1} + \varepsilon_n$. Following Gerber [12], we have that $w^d = (1-a)w_0$, where w_0 is the (independent) adjustment coefficient of ε_1. We have simulated 50 samples of length 10000 for $\theta = 1.2$, $c = 1$, $a = 0.3$. The theoretical value of the (dependent) adjustment coefficient is $w^d = 0.263$. The table below gives the mean of the estimated value $\widehat{w}^d(r)$, the standard deviation and squared root of the mean squared error over the 50 samples, for r going from 1 to 15.

r	1	2	3	4	5	6	7	8	9
mean	0.45	0.36	0.32	0.31	0.29	0.27	0.28	0.28	0.28
deviation	0.046	0.036	0.029	0.028	0.028	0.026	0.033	0.028	0.035
$\sqrt{\text{mse}}$	0.183	0.103	0.063	0.052	0.043	0.038	0.038	0.035	0.042
r	10	11	12	13	14	15			
mean	0.28	0.28	0.28	0.28	0.28	0.29			
deviation	0.034	0.035	0.032	0.038	0.033	0.037			
$\sqrt{\text{mse}}$	0.038	0.039	0.038	0.042	0.04	0.039			

We remark that the choice of r is not easy to determine, the best choice in this example seems to be $r = 6$ or $r = 7$. We remark also that our estimator seems to overestimate a bit the adjustment coefficient.

5.1.3. *$MA(1)$ 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 find $w^d = 0.31$. We have simulated 50 samples of length 10000. The table below gives the mean of the estimated value $\widehat{w}^d(r)$, the standard deviation and squared root of the mean squared error over the 50 samples, for r going from 1 to 15.

r	1	2	3	4	5	6	7	8	9
mean	0.41	0.35	0.33	0.32	0.33	0.32	0.32	0.32	0.32
deviation	0.034	0.031	0.034	0.033	0.027	0.032	0.032	0.038	0.034
$\sqrt{\text{mse}}$	0.1	0.051	0.039	0.034	0.03	0.033	0.032	0.038	0.033
r	10	11	12	13	14	15			
mean	0.32	0.32	0.32	0.32	0.33	0.32			
deviation	0.031	0.035	0.04	0.037	0.047	0.038			
$\sqrt{\text{mse}}$	0.032	0.035	0.041	0.038	0.047	0.039			

For this example, a good choice for r seems to be $r = 5$ or $r = 6$ or even $r = 7$. We observe the same stabilization phenomenon of the values of the estimator but when r increases, the standard deviation increases also.

5.1.4. *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 Dedecker et al. [7]): $X_n = aX_{n-1}^2 + 0.7\varepsilon_n$. We have simulated 50 samples of length 10000, with $\theta = 1.2$, $c = 1$, $a = -0.2$. For that model we don't know the exact value of w^d. The table below gives the mean of the estimated value $\widehat{w}^d(r)$ and the standard deviation over the 50 samples, for r going from 1 to 15.

r	1	2	3	4	5	6	7	8	9
mean	0.88	1.0	1.06	1.14	1.13	1.22	1.21	1.26	1.28
deviation	0.22	0.17	0.23	0.19	0.23	0.24	0.21	0.27	0.26
r	10	11	12	13	14	15			
mean	1.28	1.33	1.26	1.44	1.41	1.42			
deviation	0.26	0.27	0.30	0.32	0.3	0.41			

Because for the previous example $r = 6$ or $r = 7$ was a good choice for r, also because for larger values of r, the standard deviation is higher, we propose to consider $\widehat{w^d} = 1.22$ as a good approximation of w^d.

5.1.5. *An example with a Gaussian innovation.* In order to give an idea of the effect of the distribution of the innovation, we have done an $AR(1)$ simulation with a Gaussian innovation. In that case, we consider $\xi_i \rightsquigarrow \mathcal{N}(0, \sigma^2)$ a i.i.d sequence and an innovation $\varepsilon_i = \xi_i - c$. In that case,

$$\mathbb{E}(e^{t\xi_i}) = e^{t^2/2},$$

and the adjustment coefficient w associated to the i.i.d. sequence $(\varepsilon_i)_{i\in\mathbb{Z}}$ is the solution to

$$\frac{t^2\sigma^2}{2} - tc = 0,$$

that is $w = \frac{2c}{\sigma^2}$.
We consider the $AR(1)$ model $X_n - aX_{n-1} = \varepsilon_n$. We have simulated 50 samples of length 10000, with $\sigma = 2$, $c = 1.2$, $a = 0.4$. For that model, we have $w_0 = 0.5$ and $w^d = (1-a)w_0 = 0.3$. The table below gives the mean of the estimated value $\widehat{w}^d(r)$, the standard deviation and squared root of the mean squared error over the 50 samples, for r going from 1 to 15.

r	1	2	3	4	5	6	7	8	9
mean	0.7	0.5	0.43	0.39	0.38	0.36	0.35	0.37	0.37
deviation	0.022	0.025	0.028	0.03	0.034	0.04	0.045	0.063	0.065
$\sqrt{\text{mse}}$	1.34	0.67	0.45	0.32	0.29	0.25	0.24	0.31	0.32
r	10	11	12	13	14	15			
mean	0.38	0.39	0.40	0.44	0.46	0.44			
deviation	0.06	0.079	0.1	0.14	0.15	0.16			
$\sqrt{\text{mse}}$	0.1	0.11	0.14	0.19	0.22	0.21			

For that example also, we remark that the estimation of w^d is decreasing and then increases or becomes non monotonic. The best estimation is given for r corresponding to this change of behavior. Here $r = 6$ or $r = 7$. Also, the estimation is much worst than in the case of an innovation with exponential law. This is probably due to the variance of the $(\varepsilon_n)_{n\in\mathbb{Z}}$ process which is higher in this example.

5.2. **How to choose the r-parameter ?** When performing the estimation of the w^d coefficient, we have to choose the value of the parameter r. Following Theorem 4.10, we should take $r = o(\ln(k))$ but the practical choice of r for n given is not clear. The study above for several samples is instructive but does not give the answer of the choice of r for a given sample. 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.

5.2.1. *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 $r = 1, \ldots, 35$, $w^i = w^d = 0.38$. The horizontal line corresponds to the exact value of w^d.

5.2.2. *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 $r = 1, \ldots, 40$, $w^d = 0.26$.

5.2.3. *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 $\theta = 1.2$, $c = 1$, $a = 0.4$, $w^d = 0.23$.

5.2.4. *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 5.1.4.

The choice of the r parameter is the main issue of our estimation method but it is also a general problem when dealing with non linear dependent processes for which we often have to make blocks of size r.

5.3. **Smaller samples.** We conclude this section with simulations on smaller samples in the linear and non linear $AR(1)$ cases. The tables below give the mean of $\widehat{w}^d$ and the squared root of the *mse* (mean squared error) for several values of r. For the linear $AR(1)$ case, based on 200 samples of size n ($n = 500,\ 1000$), our simulations indicate that our estimation procedure could be applied for $n \geq 1000$. Christ and Steinebach [3] had the same conclusion. Remark that our results are not directly comparable to theirs because for each size, we have produced 200 independent samples while they considered several sub-samples extracted from the same large sample of size 10000. Moreover, we point out that our estimation procedure is non parametric and is available for non linear processes for which no estimation procedure exists.

5.3.1. *Linear $MA(1)$ model.* The model is the same as in section 5.1.3. Recall that $w^d = 0.31$.

	r	$\widehat{w}^d$	$\sqrt{mse}$
$n = 500$	1	0.43	0.17
$n = 500$	3	0.36	0.134
$n = 1000$	1	0.43	0.16
$n = 1000$	5	0.34	0.096

The r parameter has been chosen to be considered as $o(\ln n)$, also we have computed the mean of $\widehat{w}_r^d$ for several values of r and have chosen a stabilization point : $r \mapsto \text{mean}(\widehat{w}_r^d)$ is monotonic before this point and non monotonic after. It seems to be a convenient choice.

5.3.2. *Non linear $AR(1)$ model.* The model is the same as in section 5.1.4. Recall that for a large sample ($n = 10000$), we got $\widehat{w}^d = 1.21$.

	r	$\widehat{w}^d$
$n = 500$	1	0.79
$n = 500$	3	1.24
$n = 1000$	1	0.79
$n = 1000$	5	1.22

We choose r as in the $MA(1)$ case. Our simulations indicate that our estimation procedure should be useful, especially for non linear processes.

5.4. **Conclusion.** This simulation section emphasizes that two important questions have to be considered further. We observe that the choice of r may be guided by the behavior of $\widehat{w}^d$ computed for several values of r but should be confirmed theoretically (maybe using a cross-validation procedure). Also, it would be useful to construct confidence intervals, which would require further work.

References

[1] R.C. Bradley, *Basic properties of strong mixing conditions. A survey and some open questions. Update of, and a supplement to, the 1986 original.* Probab. Surv. 2 (2005), 107–144.

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

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

[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, *Discrete-time risk models based on times series for count random variables.* ASTIN Bulletin, (2010), to appear.

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

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

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

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

[14] I. Iscoë, P. Ney, Nummelin, *Large deviations of uniformly recurrent Markov additive processes.* Adv. Appl. Math. 6 , (1985), 373-412.

[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] H. Nyrhinen, *Rough descriptions of ruin for a general class of surplus processes.* Advances in Applied Probability 30, (1998), 1008–1026.

[20] H. Nyrhinen, *Large deviations for the time of ruin.* Journal of Applied Probability 36, (1999), 733–746.

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

[22] A.W. van der Vaart, *Asymptotic Statistics.* Cambridge Series in Statistical and Probabilistic Mathematics, (1998).

E-mail address: `helene.cossette@act.ulaval.ca`
E-mail address: `etienne.marceau@act.ulaval.ca`
E-mail address: `veronique.maume@univ-lyon1.fr`

Université Laval

Université de Lyon, Université Lyon 1, ISFA, laboratoire SAF