Mixing properties of integer-valued GARCH processes

Paul Doukhan*, Naushad Mamode Khan and Michael H. Neumann

Professor at CY University,
UMR 8088 Analyse, Géométrie et Modélisation,
2, avenue Adolphe Chauvin,
95302 Cergy-Pontoise Cedex, France.
E-mail address: doukhan@cyu.fr
URL: https://doukhan.u-cergy.fr/

Associate Professor at University of Mauritius,
Department of Economics and Statistics,
Faculty of Social Sciences and Humanities,
Reduit, Mauritius.
E-mail address: n.mamodekhan@uom.ac.mu
URL: http://www.researchgate.net/profile/Naushad_Mamode_Khan

Professor at Friedrich-Schiller-Universität,
Institut für Mathematik,
Ernst-Abbe-Platz 2,
D – 07743 Jena, Germany.
E-mail address: michael.neumann@uni-jena.de
URL: http://www.stochastik.uni-jena.de/Mitarbeiter/Prof_+Dr_+M_+Neumann.html

Abstract. We consider models for count variables with a GARCH-type structure. Such a process consists of an integer-valued component and a volatility process. Using arguments for contractive Markov chains we prove that this bivariate process has a unique stationary regime. Furthermore, we show absolute regularity (β-mixing) with geometrically decaying coefficients for the count process. These probabilistic results are complemented by a statistical analysis and a few simulations.

1. Introduction and notation

Models involving integer-valued random variables have attracted increasing attention in the recent years. In most cases, the random variables are assumed to

be non-negative, since they represent numbers of counts as e.g. the series of road traffic accidents, the monthly or annual numbers of deaths due to accidents or diseases, or more recently, the striking daily numbers of infected people and deaths due the novel corona virus. Sometimes, it is however necessary to allow for both non-negative and negative integer-valued random variables. A typical field of application is the description of score differences in sports, e.g. the number of goals of the home team minus that of the away team. Another common example is that of price changes in finance, such as the tick by tick data in a Trade and Quote database, that are often represented as a combination of positive and negative integer values. Furthermore, similar types of observations also appear in differentiated series that are initially non-stationary. While in many applications a serial dependence between the observations is not taken into account, we consider models which allow to describe and exploit dependencies between consecutive variables. In view of their popularity in financial time series analysis and because of their flexibility we focus here on models with a GARCH-type structure. Adapting the structure of the classical GARCH model by Bollerslev (1986), Fokianos and Tjøstheim (2011); Fokianos and Tjø stheim (2012) considered such models for non-negative count variables and investigated conditional maximum likelihood estimators of the corresponding parameters. These authors assumed that the count variable X_t at time t conditioned on the past has a Poisson distribution with an intensity λ_t which itself is random and depends on lagged values of the count and intensity processes. Since in case of a Poisson distribution the variance is equal to the mean, a GARCH-type structure is imposed by the equation

$$\lambda_t = \omega + \alpha_1 X_{t-1} + \cdots + \alpha_p X_{t-p} + \beta_1 \lambda_{t-1} + \cdots + \beta_q \lambda_{t-q} \qquad (1.1)$$

or by nonlinear variants, $\lambda_t = f_\theta\big(X_{t-1}, \ldots, X_{t-q}, \lambda_{t-1}, \ldots, \lambda_{t-q}\big)$, where θ is a suitable parameter.

In this paper, we consider once more processes with a GARCH-type structure. In contrast to the papers mentioned above, we allow for integer-valued variables which can attain both non-negative and negative values. The most prominent example is the distribution introduced by Skellam (1946), which is the distribution of the difference of two independent Poisson variates with respective parameters λ_1 and λ_2, In the special case of $\lambda_1 = \lambda_2$ considered by Irwin (1937), the corresponding distribution has zero mean. Therefore and in contrast to (1.1), the conditional mean is no longer suitable to generate a GARCH-type structure. We will focus on second moments and consider processes where the integer-valued variables X_t have a conditional distribution Q_{v_t} where $\int x^2 \, dQ_v(x) = v$ and

$$v_t = f\big(X_{t-1}^2, \ldots, X_{t-p}^2, v_{t-1}, \ldots, v_{t-q}\big). \qquad (1.2)$$

Alomani et al. (2018) considered such a Skellam-GARCH process of order $p = q = 1$ and derived the estimating equations for a conditional maximum likelihood estimator of the parameters. However, perhaps because of the absence of suitable probabilistic tools for such models, they did not provide a further analysis of the asymptotic properties of this estimator. These authors also provided an overview of related results and applied the model to differences of non-negative data of counts of monthly drug crimes.

In this contribution, we primarily focus on stochastic properties such as existence and uniqueness of a stationary distribution and absolute regularity of integer-valued GARCH processes. In the related case of Poisson-GARCH processes with linear or

nonlinear specifications $\lambda_t = f\big(X_{t-1}, \ldots, X_{t-p}, \lambda_{t-1}, \ldots, \lambda_{t-q}\big)$, there are several forerunners of the present work and it turns out that we can build on the methods derived there. Mixing properties of such processes have been derived for a first time in Neumann (2011), for INGARCH(1,1) processes under a contractive condition. This has been generalized in Doukhan and Neumann (2019) for the INGARCH(p,q) case and under a weaker semi-contractive condition which resulted in a somewhat unusual subexponential decay of the mixing coefficients. Doukhan et al. (2020) proved absolute regularity of the count process again in the INGARCH(1,1) case but allowing a possibly non-stationary (explosive) behavior of the process. Finally, Neumann (2020) proved absolute regularity with an exponential decay of the mixing coefficients for INGARCH(p,q) processes under a fully contractive condition,

$$\big|f(x_1, \ldots, x_p, \lambda_1, \ldots, \lambda_q) - f(x'_1, \ldots, x'_p, \lambda'_1, \ldots, \lambda'_q)\big|$$
$$\leq \sum_{i=1}^{p} c_i |x_i - x'_i| + \sum_{j=1}^{q} d_j |\lambda_j - \lambda'_j|, \tag{1.3}$$

where $c_1, \ldots, c_p, d_1, \ldots, d_q$ are non-negative constants are such that $\sum_{i=1}^{p} c_i + \sum_{j=1}^{q} d_j < 1$. We will also impose the contractive condition (1.3) on the volatility function f, however, in contrast to the papers mentioned above, the arguments of this function reflect second-order properties of the process $(X_t)_t$. Note that $(Y_t)_t$ and $(Z_t)_t$ with $Y_t = \big(X_t, \ldots, X_{t-p+1}, v_t, \ldots, v_{t-q+1}\big)$ and $Z_t = \big(X_t^2, \ldots, X_{t-p+1}^2, v_t, \ldots, v_{t-q+1}\big)$ are both (first-order) Markov chains. We show in Section 2 that the contractive condition on f yields a contraction property for $(Z_t)_t$ in terms of a suitable Wasserstein metric. This implies by the Banach fixed point theorem that $(Z_t)_t$ possesses a unique stationary distribution, and a simple extra argument shows that the same property holds true for the process $(Y_t)_t$ which is of actual interest here. Furthermore, we use the contraction property once more to prove almost effortlessly absolute regularity (β-mixing) with exponentially decaying coefficients of the count process $(X_t)_t$.

We are convinced that these results can serve as a basis for further work with such models without any hassle. As an example, population dynamics can be considered after differentiation, in order to evaluate the speed (differentiated series) or the acceleration (second order differentiated series) of the evolution of species under consideration; indeed both characteristics may be either positive or negative. As an illustration of their usefulness, we apply in Section 3 our results to prove asymptotic normality of a least squares estimator of the parameters of a Skellam-ARCH model. All proofs and a few auxiliary results are collected in a final Section 4.

2. Main results

2.1. *Assumptions and a preview of the results.* We consider a class of integer-valued processes $(X_t)_{t \in \mathbb{Z}}$ defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where, for all $t \in \mathbb{Z}$,

$$X_t \mid \mathcal{F}_{t-1} \quad \sim \quad Q_{v_t}, \tag{2.1a}$$
$$v_t \quad = \quad f(X_{t-1}^2, \ldots, X_{t-p}^2, v_{t-1}, \ldots, v_{t-q}), \tag{2.1b}$$

and $\mathcal{F}_s = \sigma(X_s, v_s, X_{s-1}, v_{s-1}, \ldots)$ denotes the σ-field generated by the random variables up to time s. Assuming that f takes values in some set $V \subseteq [0, \infty)$, this

function has to be defined on $\mathbb{N}_0^p \times V^q$. The parameter v_t stands for the conditional second moment of X_t, i.e. the family of distributions $(Q_v)_{v \in V}$ on $(\mathbb{Z}, \mathcal{P}(\mathbb{Z}))$ is parametrized such that

$$\int x^2 \, dQ_v(x) \; = \; v \qquad \forall v \in V. \tag{2.1c}$$

A frequently considered special case is that of a linear model of order p, q, where

$$f(x_1, \ldots, x_p, v_1, \ldots, v_q) \; = \; \omega + \sum_{i=1}^p \alpha_i x_i + \sum_{j=1}^q \beta_j v_j. \tag{2.2}$$

It is clear that the processes $\mathbf{Y} = (Y_t)_{t \in \mathbb{Z}}$ and $\mathbf{Z} = (Z_t)_{t \in \mathbb{Z}}$ with
$Y_t = (X_t, \ldots, X_{t-p+1}, v_t, \ldots, v_{t-q+1})$ and $Z_t = (X_t^2, \ldots, X_{t-p+1}^2, v_t, \ldots, v_{t-q+1})$
are time-homogeneous Markov chains with state spaces $S = \mathbb{Z}^p \times V^q$ and $S^{\geq} = \mathbb{N}_0^p \times V^q$, respectively. The following conditions ensure existence and uniqueness of a stationary distribution of $(Z_t)_{t \in \mathbb{Z}}$, and eventually of $(Y_t)_{t \in \mathbb{Z}}$ as well. Furthermore, they also yield absolute regularity of the count process $(X_t)_{t \in \mathbb{Z}}$.

(A1) There exist non-negative constants $c_1, \ldots, c_p, d_1, \ldots, d_q$ such that

$$\sum_{i=1}^p c_i + \sum_{j=1}^q d_j < 1$$

and

$$\left| f(x_1, \ldots, x_p, v_1, \ldots, v_q) - f(x_1', \ldots, x_p', v_1', \ldots, v_q') \right|$$
$$\leq \sum_{i=1}^p c_i \, |x_i - x_i'| + \sum_{j=1}^q d_j \, |v_j - v_j'|$$

holds for all $(x_1, \ldots, x_p, v_1, \ldots, v_q), (x_1', \ldots, x_p', v_1', \ldots, v_q') \in S^{\geq}$.

(A2) The family of distributions $(Q_v)_{v \in V}$ is increasing in the following sense: If $v < v'$, $X \sim Q_v$, and $X' \sim Q_{v'}$, then $|X|$ is stochastically not greater than $|X'|$, i.e.

$$P\big(|X| \leq k\big) \; \geq \; P\big(|X'| \leq k\big), \qquad \forall k \in \mathbb{N}.$$

Remark. If (2.1b) is replaced by $v_t = \tilde{f}(|X_{t-1}|, \ldots, |X_{t-p}|, v_{t-1}, \ldots, v_{t-q})$, then remark that the function $x \mapsto \sqrt{x}$ admits the derivative $1/(2\sqrt{x})$ bounded above by $1/2$ if $x \in \mathbb{N}_0$, thus if

$$\left| \tilde{f}(x_1, \ldots, x_p, v_1, \ldots, v_q) - \tilde{f}(x_1', \ldots, x_p', v_1', \ldots, v_q') \right|$$
$$\leq \sum_{i=1}^p \tilde{c}_i \, |x_i - x_i'| + \sum_{j=1}^q d_j \, |v_j - v_j'|$$

this model again writes as before in (2.1b) with an additional factor $1/2$ in (A1), $c_j = \tilde{c}_j/2$. The mixture of both situations may also be considered and the condition and the current conditions appear in fact as more restrictive with squares that without them. An interesting open question is to remove absolute values, it cannot be considered as before.

Examples

1) **Symmetric Skellam distributions**
   Let, for $v \in V$, $Q_v = \mathrm{Skellam}(v/2, v/2)$ be a Skellam distribution with parameters $v/2$, $v/2$, i.e. Q_v is the distribution of two independent Poisson variates with parameter $v/2$ each. Suppose that $X_1, X_2 \sim \mathrm{Pois}(v/2)$ are independent. Then

   $$\int x^2 \, dQ_v(x) = \mathrm{var}(X_1 - X_2) = \mathrm{var}(X_1) + \mathrm{var}(X_2) = v,$$

   i.e. (2.1c) is satisfied. To prove that **(A2)** is fulfilled by the family $(Q_v)_{v \in V}$, suppose that $v < v'$. If $Y \sim \mathrm{Skellam}(v/2, v/2)$ and $Z \sim \mathrm{Skellam}((v' - v)/2, (v' - v)/2)$ are independent, it follows from the properties of Poisson distributions that $Y + Z \sim \mathrm{Skellam}(v'/2, v'/2)$. Since the probability mass function of Y is symmetric and unimodal (see e.g. Alzaid and Omair, 2010) we have that

   $$P(|Y| \le k) \ge P(|Y + l| \le k), \qquad \forall (l, k) \in \mathbb{Z} \times \mathbb{N}_0,$$

   which implies that

   $$P(|Y| \le k) \ge \sum_{l \in \mathbb{Z}} P(|Y + l| \le k) P(Z = l) = P(|Y + Z| \le k), \qquad \forall k \in \mathbb{N}_0.$$

   Hence, $|Y|$ is stochastically not greater than $|Y + Z|$.

2) **Mixtures of symmetric Skellam distributions**
   Let G be the distribution of a non-negative random variable with $\int_{[0,\infty)} x \, dG(x) = \mu \in (0, \infty)$.
      Then

   $$Q_v = \int_{[0,\infty)} \mathrm{Skellam}\left(\frac{vx}{2\mu}, \frac{vx}{2\mu}\right) dG(x)$$

   is a mixture of symmetric Skellam distributions. We have that

   $$\int x^2 \, dQ_v(x) = \int_{[0,\infty)} \int x^2 \, d\,\mathrm{Skellam}\left(\frac{vx}{2\mu}, \frac{vx}{2\mu}\right) dG(x)$$
   $$= \int_{[0,\infty)} \frac{vx}{\mu} \, dG(x) = v,$$

   and, for $v < v'$,

   $$Q_v(\{-k, \ldots, k\}) = \int_{[0,\infty)} \mathrm{Skellam}\left(\frac{vx}{2\mu}, \frac{vx}{2\mu}\right)(\{-k, \ldots, k\}) \, dG(x)$$
   $$\ge \int_{[0,\infty)} \mathrm{Skellam}\left(\frac{v'x}{2\mu}, \frac{v'x}{2\mu}\right)(\{-k, \ldots, k\}) \, dG(x)$$
   $$= Q_{v'}(\{-k, \ldots, k\}),$$

   i.e. (2.1c) and **(A2)** are satisfied.
      A notable special case is that of a zero-inflated Skellam distribution, where $B \sim G$ follows a Bernoulli distribution with parameter $p \in (0, 1)$. If B and $X \sim \mathrm{Skellam}(v/(2p), v/(2p))$ are independent, then BX has a zero-inflated Skellam distribution. Such a distribution was used by Karlis and Ntzoufras (2006) and Andersson and Karlis (2014) to account for an excess of zero counts in certain medical data.

3) **Poisson distributions**
   If $X \sim \text{Pois}(\lambda)$, then $EX^2 = \lambda^2 + \lambda$. This is equal to $v > 0$ if and only if $\lambda = \sqrt{v + 1/4} - 1/2$. In order to obey (2.1c) we choose $Q_v = \text{Pois}(\sqrt{v + 1/4} - 1/2)$. For $v < v'$ we have that $\sqrt{v + 1/4} - 1/2 < \sqrt{v' + 1/4} - 1/2$. Since $X \sim \text{Pois}(\sqrt{v + 1/4} - 1/2)$ is stochastically not greater than $X \sim \text{Pois}(\sqrt{v' + 1/4} - 1/2)$ we see that condition **(A2)** is satisfied.

4) **Mixtures of Poisson distributions**
   Let, as in Example 2, G be the distribution function of a non-negative random variable with $\int_0^\infty x \, dG(x) = \mu \in (0, \infty)$. Then

$$Q_v = \int_0^\infty \text{Pois}\left(\sqrt{vx/\mu + 1/4} - 1/2\right) dG(x)$$

is a mixture of Poisson distributions. Then condition **(A2)** is obviously fulfilled. Furthermore, since

$$\int x^2 \, dQ_v(x) = \int_{[0,\infty)} \int x^2 \, d\text{Pois}\left(\sqrt{vx/\mu + 1/4} - 1/2\right) dG(x)$$
$$= \int_{[0,\infty)} \frac{vx}{\mu} \, dG(x) = v$$

we see that (2.1c) is also satisfied. Poisson distributions can be used for modeling data from various fields, e.g. the number of financial transactions within a certain time period or the number of claims in an insurance context. When dealing with a collection of individual transactions corresponding to different trading strategies or with a collection of claim numbers from persons with different features (age, health state,...) an appropriate mixture of Poisson distributions seems to be more adequate. Notable special cases are that of a zero-inflated Poisson distribution which appears in case of G a Bernoulli distribution with parameter $p \in (0, 1)$ or a negative binomial distribution, if G has a Gamma distribution. A negative binomial distribution is often preferred to a Poisson distribution if data are overdispersed, i.e. if their variance is greater than their mean (as this is the case for all mixed Poison distributions).

5) **Binomial distributions**
   A $\text{Bin}(n, p)$ distribution ($p \in (0, 1)$) can be used for modeling underdispersed data since its variance $np(1 - p)$ is less than its mean np. To satisfy (2.1c), we set

$$Q_v = \text{Bin}\left(n, g(v)\right),$$

where $g \colon (0, n^2) \to (0, 1)$ is a strictly monotonic function such that

$$\int x^2 \, d\text{Bin}(n, g(v)) = ng(v) + n(n - 1)\left(g(v)\right)^2 = v.$$

To see that **(A2)** is fulfilled, let $U_1, \ldots, U_n$ be independent and uniformly distributed on $[0, 1]$. Let $v, v' \in (0, n^2)$. Then $X := \sum_{i=1}^n \mathbb{1}_{\{U_i \leq g(v)\}} \sim Q_v$ and $X' := \sum_{i=1}^n \mathbb{1}_{\{U_i \leq g(v')\}} \sim Q_{v'}$. If $v < v'$, then it follows from the construction that $X \leq X'$ with probability one which implies that X is stochastically not greater than X'.

6) **Some asymmetric distributions over $\mathbb{Z}$**

Let Q_v ba any of the above distributions and let $Y \sim Q_v$ and R be independent, where $P(R \in \{-1, 1\}) = 1$. If $P(R = 1) \neq 1$, then the distributions of the random variable $X = RY$ is asymmetric over $\mathbb{Z}$ and obeys (2.1c). The corresponding family of distributions satisfies **A2**) as soon as $(Q_v)_{v \in V}$ does.

In the following we derive a contraction property of $(Z_t)_t$ in terms of a suitable Wasserstein metric. As shown in Eberle (2019, Chapter 3) and Douc et al. (2018, Theorem 20.3.4), this implies by the Banach fixed point theorem that $(Z_t)_t$ possesses a unique stationary distribution. A simple extra argument shows that the same property holds true for the process $(Y_t)_t$ which is of actual interest here. Furthermore, we use the contraction property once more to prove almost effortlessly absolute regularity (β-mixing) with exponentially decaying coefficients of the count process $(X_t)_t$.

2.2. *Contraction.* First of all, we transfer the contraction condition **(A1)** for the intensity process into a contraction property for the Z_t.

We consider the following metric on $S^{\geq}$:

$$\Delta_{\gamma,\delta}\big((x_1, \ldots, x_p, v_1, \ldots, v_q), (x'_1, \ldots, x'_p, v'_1, \ldots, v'_q)\big)$$
$$= \sum_{i=1}^{p} \gamma_i \, |x_i - x'_i| + \sum_{j=1}^{q} \delta_j \, |v_j - v'_j|,$$

where $\gamma_1, \ldots, \gamma_p, \delta_1, \ldots, \delta_q$ are strictly positive constants. Let $y = (x_1, \ldots, x_p, v_1, \ldots, v_q)$, $y' = (x'_1, \ldots, x'_p, v'_1, \ldots, v'_q) \in S$ be arbitrary and, accordingly $z = (x_1^2, \ldots, x_p^2, v_1, \ldots, v_q)$, $z' = (x_1'^2, \ldots, x_p'^2, v'_1, \ldots, v'_q) \in S^{\geq}$. With an appropriate choice of $\gamma_1, \ldots, \gamma_p, \delta_1, \ldots, \delta_q$, we can construct random vectors $Y = (X, x_1, \ldots, x_{p-1}, \lambda, \lambda_1, \ldots, \lambda_{q-1})$ and $Y' = (X', x'_1, \ldots, x'_{p-1}, \lambda', \lambda'_1, \ldots, \lambda'_{q-1})$ on a suitable probability space $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{P})$ such that

$$\widetilde{P}^Y = \mathbb{P}^{Y_t | Y_{t-1} = y} = \mathbb{P}^{Y_t | Z_{t-1} = z}, \qquad \widetilde{P}^{Y'} = \mathbb{P}^{Y_t | Y_{t-1} = y'} = \mathbb{P}^{Y_t | Z_{t-1} = z'}, \qquad (2.3)$$

and, for $Z = (X^2, x_1, \ldots, x_{p-1}, \lambda, \lambda_1, \ldots, \lambda_{q-1})$ and $Z' = (X'^2, x'_1, \ldots, x'_{p-1}, \lambda', \lambda'_1, \ldots, \lambda'_{q-1})$,

$$\widetilde{E}\Delta_{\gamma,\delta}(Z, Z') \leq \kappa \, \Delta_{\gamma,\delta}(z, z') \qquad (2.4)$$

holds for some $\kappa < 1$. Actually, according to the model equation (2.1b), we have to set $v = f(x_1^2, \ldots, x_p^2, v_1, \ldots, v_q)$ and $v' = f(x_1'^2, \ldots, x_p'^2, v'_1, \ldots, v'_q)$. Suppose that $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{P})$ admits the construction of independend random variables U and V, both following a uniform distribution on $[0, 1]$. Let G_v and $G_{v'}$ be the respective distribution functions of $\mathbb{P}^{X_t^2 | v_t = v}$ and $\mathbb{P}^{X_t^2 | v_t = v'}$. We define versions of X^2 and X'^2 by $W := G_v^{-1}(U)$ and $W' := G_{v'}^{-1}(U)$, where G^{-1} denotes the generalized inverse of a generic distribution function G, $G^{-1}(t) = \inf\{x \colon G(x) \geq t\}$. We still have to determine the signs of X and X', taking into account that the values of $X^2 = W$ and $X'^2 = W'$ are already determined. With a view to our proof of absolute regularity, and since the probability $P(|X| = |X'|)$ is under control, we will do this in such way that the probability $P(|X| = |X'|, X \neq X')$ is as small

as possible. Let $p_l = \mathbb{P}(X_t = l \mid v_t = v)$ and $p'_l = \mathbb{P}(X_t = l \mid v_t = v')$. If $(p_k \wedge p'_k) + (p_{-k} \wedge p'_{-k}) \leq P(W = W' = k^2)$, then we can couple the signs of X and X' such that

$$
\begin{aligned}
P(W = W' = k^2, X = X' = k) &= p_k \wedge p'_k, \qquad \text{and} \\
P(W = W' = k^2, X = X' = -k) &= p_{-k} \wedge p'_{-k}.
\end{aligned}
$$

In this case,

$$P(|X| = |X'| = |k|, X \neq X') = P(|X| = |X'| = |k|) - p_k \wedge p'_k - p_{-k} \wedge p'_{-k}. \tag{2.5a}$$

On the other hand, if $(p_k \wedge p'_k) + (p_{-k} \wedge p'_{-k}) > P(W = W' = k^2)$, then we couple X and X' such that

$$P(W = W' = k^2, X = X' = \pm k) = p_{\pm k} \wedge p'_{\pm k} \frac{P(W = W' = k^2)}{(p_k \wedge p'_k) + (p_{-k} \wedge p'_{-k})},$$

which leads to

$$P(|X| = |X'| = |k|, X \neq X') = 0. \tag{2.5b}$$

We denote the corresponding Markov kernels by $\pi^{Y,Y'}$ and $\pi^{Z,Z'}$, respectively. This construction produces a pair (X, X') such that

$$
\begin{aligned}
\widetilde{P}(X = k) &= \mathbb{P}(X_t = k \mid Y_{t-1} = y), \\
\text{and} \quad \widetilde{P}(X' = k) &= \mathbb{P}(X_t = k \mid Y_{t-1} = y'), \ \forall k \in \mathbb{Z},
\end{aligned}
$$

which means that (2.3) is satisfied. Furthermore, it follows from **(A2)** that X^2 is stochastically not greater than X'^2 if $v \leq v$ and vice versa. Since the coupling of these random variables is based on the quantile transform we obtain by (2.1c)

$$\widetilde{E}|X^2 - X'^2| = |\widetilde{E}[X^2 - X'^2]| = |v - v'|. \tag{2.6}$$

Since $P(|X| = |X'| = |k|) \leq (p_k + p_{-k}) \wedge (p'_k \vee p'_{-k}) \leq (p_k + p'_k + p_{-k} + p'_{-k})/2$ we obtain from (2.5a) and (2.5b) that

$$
\begin{aligned}
P(|X| = |X'|, X \neq X') &= \sum_{k=1}^{\infty} P(|X| = |X'| = k, X \neq X') \\
&\leq \sum_{k=1}^{\infty} (p_k + p'_k + p_{-k} + p'_{-k})/2 - p_k \wedge p'_k - p_{-k} \wedge p'_{-k} \\
&\leq \sum_{k=-\infty}^{\infty} \frac{p_k + p'_k}{2} - p_k \wedge p'_k = \frac{1}{2} \sum_{k=-\infty}^{\infty} |p_k - p'_k|. \tag{2.7}
\end{aligned}
$$

It follows from (2.6) and by **(A1)** that

$$
\begin{aligned}
\widetilde{E} &\Delta_{\gamma,\delta}(Z, Z') \\
&= (\gamma_1 + \delta_1) \left| f(x_1^2, \ldots, x_p^2, v_1, \ldots, v_q) - f(x_1'^2, \ldots, x_p'^2, v_1', \ldots, v_q') \right| \\
&\quad + \sum_{i=2}^{p} \gamma_i |x_{i-1}^2 - x_{i-1}'^2| + \sum_{j=2}^{q} \delta_j |v_{j-1} - v_{j-1}'| \tag{2.8} \\
&\leq (\gamma_1 + \delta_1) \left(\sum_{i=1}^{p} c_i |x_i^2 - x_i'^2| + \sum_{j=1}^{q} d_j |v_j - v_j'| \right)
\end{aligned}
$$

$$+ \sum_{i=2}^{p} \gamma_i \, |x_{i-1}^2 - x_{i-1}'^{\,2}| + \sum_{j=2}^{q} \delta_j \, |v_{j-1} - v_{j-1}'|.$$

The desired relation of $\widetilde{E}\Delta(Z, Z') \le \kappa\Delta(z, z')$ would be guaranteed to hold if we find strictly positive $\gamma_1, \ldots, \gamma_p, \delta_1, \ldots, \delta_p$ such that the right-hand side of (2.8) is less than or equal to $\kappa\Delta(z, z') = \kappa\big(\sum_{i=1}^{p} \gamma_i |x_i^2 - x_i'^{\,2}| + \sum_{j=1}^{q} \delta_j|v_j - v_j'|\big)$, for all $(x_1^2, \ldots, x_p^2, v_1, \ldots, v_q)$, $(x_1'^{\,2}, \ldots, x_p'^{\,2}, v_1', \ldots, v_q') \in S^{\ge}$. The following lemma provides a bridge from the contraction property **(A1)** for the volatility function to a contraction property for Z_t.

Lemma 2.1. *Let $c_1, \ldots, c_p, d_1, \ldots, d_q$ be non-negative constants with*

$$\sum_{i=1}^{p} c_i + \sum_{j=1}^{q} d_j < 1.$$

Then there exist strictly positive constants $\gamma_1, \ldots, \gamma_p, \lambda_1, \ldots, \lambda_q$ and some $\kappa < 1$ such that

$$
(\gamma_1 + \delta_1) \Big(\sum_{i=1}^{p} c_i \, y_i + \sum_{j=1}^{q} d_j \, z_j \Big) + \sum_{i=2}^{p} \gamma_i \, y_{i-1} + \sum_{j=2}^{q} \delta_j \, z_{j-1}
$$
$$
\le \kappa \Big(\sum_{i=1}^{p} \gamma_i \, y_i + \sum_{j=1}^{q} \delta_j \, z_j \Big)
$$

(2.9)

holds for all $y_1, \ldots, y_p, z_1, \ldots, z_q \ge 0$.

Let $\pi^{Y,Y'}$ and $\pi^{Z,Z'}$ be the Markov kernels which provide the above coupling, that is, for the above pairs of random variables (Y, Y') and (Z, Z') we have that $(Y, Y') \sim \pi^{Y,Y'}((y, y'), \cdot)$ and $(Z, Z') \sim \pi^{Z,Z'}((z, z'), \cdot)$, respectively.

The following proposition provides the contraction property which will be instrumental for the proof of the existence and uniqueness of a stationary distribution as well as for the derivation of absolute regularity of the count process.

Proposition 2.2. *Suppose that conditions **(A1)** and **(A2)** are fulfilled. Let $\gamma_1, \ldots, \gamma_p, \delta_1, \ldots, \delta_p$ and $\kappa < 1$ be chosen as in Lemma 2.1. Then*

(i) *Let $z, z' \in S^{\ge}$ be arbitrary. If $(Z, Z') \sim \pi^{Z,Z'}((z, z'), \cdot)$, then*

$$Z \sim \mathbb{P}^{Z_t | Z_{t-1} = z} \qquad and \qquad Z' \sim \mathbb{P}^{Z_t | Z_{t-1} = z'}$$

and

$$\widetilde{E}\Delta_{\gamma,\delta}(Z, Z') \le \kappa \, \Delta_{\gamma,\delta}(z, z').$$

(ii) *Let $((\widetilde{Z}_t, \widetilde{Z}_t'))_{t\in\mathbb{Z}}$ be a Markov chain on $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{P})$ with transition kernel $\pi^{Z,Z'}$. Then*

$$\widetilde{E}\Delta_{\gamma,\delta}\left(\widetilde{Z}_t, \widetilde{Z}_t'\right) \le \kappa \, \widetilde{E}\Delta_{\gamma,\delta}\left(\widetilde{Z}_{t-1}, \widetilde{Z}_{t-1}'\right).$$

In order to derive stationarity properties of the process $(Z_t)_{t\in\mathbb{Z}}$, we further translate the contraction result in Proposition 2.2 into a contraction property of the corresponding distributions. For the metric $\Delta_{\gamma,\delta}$ on $S^{\ge}$, we define

$$\mathcal{P}(S^{\ge}) = \big\{Q \colon Q \text{ is a probability distribution on } S^{\ge} \text{ with} \int\Delta_{\gamma,\delta}(z_0, z)dQ(z) < \infty\big\},$$

where $z_0 \in S^{\geq}$ is arbitrary. For two probability measures $Q, Q' \in \mathcal{P}(S^{\geq})$, we define the Kantorovich distance based on the metric $\Delta_{\gamma,\delta}$ (also known as Wasserstein L^1 distance) by

$$\mathcal{K}(Q, Q') := \inf_{Z \sim Q, Z' \sim Q'} \widetilde{E} \Delta(Z, Z'),$$

where the infimum is taken over all random variables Z and Z' defined on a common probability space $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{P})$ with respective laws Q and Q'. We denote the Markov kernel of the processes $(Y_t)_{t \in \mathbb{Z}}$ and $(Z_t)_{t \in \mathbb{Z}}$ by π^Y and π^Z, respectively. The following result follows immediately from Proposition 2.2.

Proposition 2.3. *Suppose that conditions* **(A1)** *and* **(A2)** *are fulfilled. Let $Q, Q' \in \mathcal{P}(S^{\geq})$ be arbitrary distributions. Then, for $\kappa < 1$ given in Lemma 2.1,*

$$\mathcal{K}(Q\pi_\theta^Z, Q'\pi_\theta^Z) \leq \kappa \mathcal{K}(Q, Q').$$

2.3. *Existence and uniqueness of a stationary distribution.* Proposition 2.3 shows that the mapping π^Z is contractive. Therefore, we can conclude by the Banach fixed point theorem that the Markov process $(Z_t)_{t \in \mathbb{Z}}$ has a unique stationary distribution. A simple extra argument shows that that the Markov process $(Y_t)_{t \in \mathbb{Z}}$ has this property as well.

Theorem 2.4. *Suppose that conditions* **(A1)** *and* **(A2)** *are fulfilled.*

(i) *The Markov process $(Z_t)_{t \in \mathbb{Z}}$ with transition kernel π^Z has a unique stationary distribution Q^Z. For $Z_0 = (X_0^2, \ldots, X_{1-p}^2, v_0, \ldots, v_{1-q})$, we have that*

$$E\left[X_0^2 + v_0\right] < \infty. \tag{2.10}$$

(ii) *The Markov process $(Y_t)_{t \in \mathbb{Z}}$ with transition kernel π^Y has a unique stationary distribution Q^Y.*

Remark 2.5. The reader might wonder why we don't derive weak dependence properties introduced by Doukhan and Louhichi (1999). Indeed e.g. Doukhan and Neumann (2008) describe statistical procedures where mixing can be replaced by weak dependence conditions. If $X \sim Q_v = \text{Skellam}(v/2, v/2)$ is symmetric, then $EX = 0$ and $EX^2 = v$. Therefore it is natural to model the volatility process as in (2.1b), where the volatilities appear linearly while the count variables are squared. The properties of Skellam models make also natural the inhomogeneity of (2.1a) and (2.1b) which include both linear and squared factors. In the simplest case of a SkellamARCH(1)-process, with $v_t = f(X_{t-1}^2)$, the function $x \mapsto f(x^2)$ may not be Lipschitz and thus contraction does not hold. Anyway, the process $Y_t = X_t^2$ is again contractive if Lip $f < 1$. Symmetry of the distribution Q_v implies that $X_t = \sigma_t \sqrt{Y_t}$ is a solution of (2.1a) and (2.1b) if $(\sigma_t)_t$ is an iid sequence of symmetric signs ($P(\sigma_t = \pm 1) = 1/2$). Then $\tau-$dependence of the process (Y_t) follows as in Doukhan and Wintenberger (2008). Now, since the 1-Lipschitz function $g(y) = y \wedge \sqrt{y}$ equals $y \mapsto \sqrt{y}$ on $\mathbb{N}_0$, then $X_t = \sigma_t g(Y_t)$; heredity properties of weak dependence imply geometric $\tau-$dependence of $(X_t)_t$; see Dedecker et al. (2007). We proved $\tau-$dependence in this very special symmetric case; in order to work in a more general setting we switch in Subsection 2.4 to the more standard $\beta-$mixing condition to derive asymptotic theory for the statistical analysis.

2.4. *Absolute regularity.* For the related case of Poisson count processes with a GARCH-type structure, absolute regularity has been first proved for contractive IN-GARCH(1,1) processes in Neumann (2011). This has been generalized in Doukhan and Neumann (2019) to semi-contractive models and in Doukhan et al. (2020) to the case of possibly non-stationary processes. In all of these papers, the mixing properties were derived by an explicit coupling of two versions of the processes which were tailor-made for the respective properties of the processes. In the current work, our approach is slightly different. We derive both stationarity and mixing properties on the basis of a one-step contractivity property given in Proposition 2.2.

Let $(\Omega, \mathcal{A}, P)$ be a probability space and $\mathcal{A}_1$, $\mathcal{A}_2$ be two sub-σ-algebras of $\mathcal{A}$. Then the coefficient of absolute regularity is defined as

$$\beta(\mathcal{A}_1, \mathcal{A}_2) = E\left[\sup\left\{|P(B \mid \mathcal{A}_1) - P(B)|: \ B \in \mathcal{A}_2\right\}\right].$$

For a strictly stationary process $\mathbf{Y} = (Y_t)_t$ on $(\Omega, \mathcal{A}, P)$, the coefficients of absolute regularity are defined as

$$\beta^Y(n) = \beta\left(\sigma(Y_0, Y_{-1}, \ldots), \sigma(Y_n, Y_{n+1}, \ldots)\right).$$

For the count process $(X_t)_t$ on $(\Omega, \mathcal{F}, \mathbb{P})$, we obtain the following estimate of the coefficients of absolute regularity.

$$\beta^X(n)$$
$$= \beta\left(\sigma(X_0, X_{-1}, \ldots), \sigma(X_n, X_{n+1}, \ldots)\right)$$
$$\leq \beta\left(\mathcal{F}_0, \sigma(X_n, X_{n+1}, \ldots)\right)$$
$$= \beta\left(\sigma(Y_0), \sigma(X_n, X_{n+1}, \ldots)\right)$$
$$= \mathbb{E}\left[\sup_{C \in \sigma(\mathcal{Z})}\left\{\left|\mathbb{P}_\theta\left((X_n, X_{n+1}, \ldots) \in C \mid Z_k\right) - \mathbb{P}_\theta\left((X_n, X_{n+1}, \ldots) \in C\right)\right|\right\}\right], \quad (2.11)$$

where $\mathcal{Z} = \{A \times \mathbb{Z} \times \mathbb{Z} \times \cdots \mid A \subseteq \mathbb{Z}^m, m \in \mathbb{N}\}$ is the system of cylinder sets. At this point we employ a coupling argument. Let $((\widetilde{Y}_t, \widetilde{Y}'_t))_{t \in \mathbb{N}_0}$ be a Markov chain on a probability space $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{P})$ with transition kernel $\pi^{Y,Y'}$ and independent variables $\widetilde{Y}_0, \widetilde{Y}'_0 \sim \mathbb{P}_\theta^{Z_k}$. Then

$$\mathbb{E}\left[\sup_{C \in \sigma(\mathcal{Z})}\left\{\left|\mathbb{P}_\theta\left((X_k, X_{n+1}, \ldots) \in C \mid Z_k\right) - \mathbb{P}_\theta\left((X_k, X_{n+1}, \ldots) \in C\right)\right|\right\}\right]$$
$$\leq \widetilde{E}\left[\sup_{C \in \sigma(\mathcal{Z})}\left\{\left|\widetilde{P}\left((\widetilde{X}_n, \widetilde{X}_{n+1}, \ldots) \in C \mid \widetilde{Y}_0\right) - \widetilde{P}\left((\widetilde{X}'_n, \widetilde{X}'_{n+1}, \ldots) \in C \mid \widetilde{Y}'_0\right)\right|\right\}\right]$$
$$\leq \widetilde{P}\left(\widetilde{X}_{n+k} \neq \widetilde{X}'_{n+k} \quad \text{for some } k \geq 0\right)$$
$$\leq \sum_{k=0}^{\infty} \widetilde{P}\left(\widetilde{X}_{n+k} \neq \widetilde{X}'_{n+k}\right). \quad (2.12)$$

At this point we will more closely examine the remaining part of our approach to derive upper estimates for the mixing coefficients. If the count variables are non-negative, then $\widetilde{X}_{n+k} = \widetilde{X}'_{n+k}$ is equivalent to $|\widetilde{X}_{n+k}| = |\widetilde{X}'_{n+k}|$. Moreover, if the probability mass functions of the Q_v are symmetric about zero, then (2.5a) and (2.5b) ensure that $\widetilde{X}_{n+k}$ and $\widetilde{X}'_{n+k}$ have always the same sign which means again

that $\widetilde{X}_{n+k} = \widetilde{X}'_{n+k}$ is equivalent to $|\widetilde{X}_{n+k}| = |\widetilde{X}'_{n+k}|$. In both cases, we conclude from (2.11) and (2.12) that

$$
\begin{aligned}
\beta^X(n) &\leq \sum_{k=0}^{\infty} \widetilde{P}\left(\widetilde{X}_{n+k} \neq \widetilde{X}'_{n+k}\right) \\
&\leq \frac{1}{\gamma_1} \sum_{k=0}^{\infty} \widetilde{E}\Delta_{\gamma,\delta}(\widetilde{Z}_{n+k}, \widetilde{Z}'_{n+k}) \\
&\leq \frac{1}{\gamma_1} \frac{\kappa^n}{1-\kappa} \widetilde{E}\Delta_{\gamma,\delta}(\widetilde{Z}_0, \widetilde{Z}'_0).
\end{aligned}
\tag{2.13}
$$

Otherwise, we assume that $(Q_v)_{v\in V}$ is such that, for some $K < \infty$,

$$
\frac{1}{2} \sum_{k=-\infty}^{\infty} \left|Q_v(\{k\}) - Q_{v'}(\{k\})\right| \leq K|v-v'| \qquad \forall v, v' \in V.
\tag{2.14}
$$

Then we obtain by (2.7) that

$$
\begin{aligned}
&\widetilde{P}\left(\widetilde{X}_{n+k} \neq \widetilde{X}'_{n+k}\right) \\
&= \widetilde{P}\left(|\widetilde{X}_{n+k}| \neq |\widetilde{X}'_{n+k}|\right) + \widetilde{P}\left(|\widetilde{X}_{n+k}| = |\widetilde{X}'_{n+k}|, \widetilde{X}_{n+k} \neq \widetilde{X}'_{n+k}\right) \\
&\leq \widetilde{P}\left(|\widetilde{X}_{n+k}| \neq |\widetilde{X}'_{n+k}|\right) + K\left|v_{n+k} - v'_{n+k}\right|.
\end{aligned}
$$

In this case, we obtain that

$$
\begin{aligned}
\beta^X(n) &\leq \left(\frac{1}{\gamma_1} + \frac{K}{\delta_1}\right) \sum_{k=0}^{\infty} \Delta_{\gamma,\delta}(\widetilde{Z}_{n+k}, \widetilde{Z}'_{n+k}) \\
&\leq \left(\frac{1}{\gamma_1} + \frac{K}{\delta_1}\right) \frac{\kappa^n}{1-\kappa} \widetilde{E}\Delta_{\gamma,\delta}(\widetilde{Z}_0, \widetilde{Z}'_0).
\end{aligned}
\tag{2.15}
$$

Theorem 2.6. *Suppose that conditions* **(A1)** *and* **(A2)** *are fulfilled and that the process* $(Y_t)_{t\in\mathbb{Z}}$ *is stationary. Furthermore we assume that* $(Q_v)_{v\in V}$ *satisfies one of the following conditions.*

    a) $Q_v(\mathbb{N}_0) = 1 \quad \forall v \in V$,

    b)   *the probability mass functions of* $(Q_v)_{v\in V}$ *are symmetric about zero,*

    c)   *(2.14) is fulfilled for some* $K < \infty$.

Then there exists some $\rho < 1$ *such that*

$$
\beta^X(n) = O(\rho^n).
$$

Remark 2.7. The results of our paper are heavily based on the (fully) contractive condition **(A1)** on the volatility function f. In a related work, Doukhan and Neumann (2019), a weaker so-called semi-contractive condition,

$$
\left|f(x_1, \ldots, x_p, \lambda_1, \ldots, \lambda_q) - f(x_1, \ldots, x_p, \lambda'_1, \ldots, \lambda'_q)\right| \leq \sum_{j=1}^{q} d_j|\lambda_j - \lambda'|,
$$

for some non-negative $d_1, \ldots, d_q$ such that $\sum_{j=1}^{q} d_j < 1$, was imposed which then resulted a a slower subexponential decay of the coefficients of absolute regularity. In our context, it seems also be possible to derive properties such as existence and uniqueness of a stationary distribution and absolute regularity under a semi-contractive condition if some appropriate drift condition is added. Without any kind of contractivity condition, the approach used in this paper fails and there are

counterexamples showing that then our results are non longer valid. Consider the special case of the linear model (2.2), where $\omega > 0$ and $\alpha_1, \ldots, \alpha_p, \beta_1, \ldots, \beta_q$ are non-negative with $\sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j \geq 1$. Then

$$\mathbb{E}v_t = \omega + \sum_{i=1}^{p} \alpha_i \mathbb{E}v_{t-i} + \sum_{j=1}^{q} \beta_j \mathbb{E}v_{t-j},$$

which shows that a stationary process $((X_t, v_t))_{t \in \mathbb{Z}}$ satisfying (2.1a) to (2.1c) does not exist.

3. Applications

We choose to develop the asymptotic theory for the OLSE of Skellam models (3.1) as the most standard application of the above results. Much more may be done including tests of goodness-of-fit as in Doukhan et al. (2020). Prediction or model selection issues are also important and should be developed theoretically. Additional research work will make use of the bound of absolute regularity for many other questions such a more quantitative study of prediction, qualitative tests of goodness-of-fit such as model choice problems, or more nonparametric based statistics or resampling or subsampling procedures.

3.1. *OLSE of a Skellam-ARCH model.* We consider the special case of an Skellam-ARCH(p) model, where (2.1b) reduces to

$$v_t = \omega + \sum_{i=1}^{p} \alpha_i X_{t-i}^2. \tag{3.1}$$

We assume that $\omega > 0$, and that $\alpha_1, \ldots, \alpha_p$ are non-negative with $\alpha = \sum_{i=1}^{p} \alpha_i < 1/\sqrt{3}$. We further assume that the process $((X_t, v_t))_{t \in \mathbb{Z}}$ is in its unique stationary regime. On the basis of observations $X_1, \ldots, X_n$, we intend to estimate the vector of unknown parameters $\theta = (\omega, \alpha_1, \ldots, \alpha_p)^T$. We embed the observed random variables into a linear regression model,

$$X_t^2 = \omega + \sum_{i=1}^{p} X_{t-i}^2 \alpha_i + \varepsilon_t, \qquad t = p+1, \ldots, n,$$

where $\varepsilon_t = X_t^2 - v_t$ satisfies $E(\varepsilon_t \mid \mathcal{F}_{t-1}) = 0$ a.s. Then the ordinary least squares estimator is given by

$$\begin{aligned}
\widehat{\theta}_n &\in \arg\min_{\theta} \sum_{t=p+1}^{n} \left(X_t^2 - (\omega + \alpha_1 X_{t-1}^2 + \cdots + \alpha_p X_{t-p}^2) \right)^2 \\
&= \arg\min_{\theta} \left\| Y_{(n)} - X_{(n)}\theta \right\|^2,
\end{aligned}$$

where

$$Y_{(n)} = \begin{pmatrix} X_{p+1}^2 \\ \vdots \\ X_n^2 \end{pmatrix}, \qquad X_{(n)} = \begin{pmatrix} 1 & X_p^2 & \cdots & X_1^2 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n-1}^2 & \cdots & X_{n-p}^2 \end{pmatrix}.$$

If the matrix $X_{(n)}^T X_{(n)}$ is regular, then $\widehat{\theta}_n$ is uniquely defined and

$$\widehat{\theta}_n = \left(X_{(n)}^T X_{(n)} \right)^{-1} X_{(n)}^T Y_{(n)}, \tag{3.2}$$

which implies that

$$\sqrt{n} \left(\widehat{\theta}_n - \theta \right) = \left(\frac{1}{n} X_{(n)}^T X_{(n)} \right)^{-1} \frac{1}{\sqrt{n}} X_{(n)}^T \varepsilon_{(n)}, \tag{3.3}$$

where $\varepsilon_{(n)} = (\varepsilon_{p+1}, \ldots, \varepsilon_n)^T$.

The condition $\sum_{i=1}^p \alpha_i < 1/\sqrt{3}$ ensures by Lemma 4.1 that $EX_0^4 < \infty$. Hence, we obtain from the ergodic theorem that

$$\frac{1}{n} X_{(n)}^T X_{(n)} \xrightarrow{a.s.} \Sigma = \begin{pmatrix} 1 & EX_1^2 & \ldots & EX_p^2 \\ EX_1^2 & EX_1^2 X_1^2 & \ldots & EX_1^2 X_p^2 \\ \vdots & \vdots & \ddots & \vdots \\ EX_p^2 & EX_p^2 X_1^2 & \ldots & EX_p^2 X_p^2 \end{pmatrix}. \tag{3.4}$$

Lemma 4.2 below shows that Σ is a regular matrix which means that equation (3.3) holds true with a probability tending to 1. Furthermore, it follows from a central limit theorem for sums of martingale differences (Corollary 3.1 in Hall and Heyde, 1980 page 58) and the Cramér-Wold device that

$$W_n := \frac{1}{\sqrt{n}} X_{(n)}^T \varepsilon_{(n)} \xrightarrow{d} Z_0, \tag{3.5}$$

where $Z_0 \sim \mathcal{N}(0_{p+1}, \eta^2 \Sigma)$ and $\eta^2 = E\varepsilon_t^2$.

From (3.3) to (3.5) we conclude that

$$\sqrt{n} \left(\widehat{\theta}_n - \theta \right) \xrightarrow{d} Z \sim \mathcal{N}(0_{p+1}, \eta^2 \Sigma^{-1}). \tag{3.6}$$

3.2. *Simulation study.* We simulate a process $((X_t, v_t))_t$, where v_t obeys (3.1) and $X_t \mid \mathcal{F}_t \sim Q_{v_t} = \text{Skellam}(v_t/2, v_t/2)$. The parameters in (3.1) are chosen such that $\omega > 0$ and $\sum \alpha_i < 1/\sqrt{3}$, which ensures finiteness of fourth moments of the count variables; see Lemma 4.1 below. We assume a suitable set of values for the different order p, $\alpha_1 = 0.26$, $\alpha_2 = 0.16$, $\alpha_3 = 0.11$, $\alpha_4 = 0.02$, for sample sizes $T = 30, 80, 100, 500$, and 1000. 1000 replication are made for each sample size and the simulated mean estimates and their corresponding standard errors as deduced from the result (3.6).

The estimates reflect that for increased sample size, the values of the different parameters become more consistent with the standard errors that are seen to be constantly decreasing; see Table 3.1. It is worth reporting that for some simulation processes, the standard OLS equation (3.3) does not yield relevant output estimates since the constraints on the $\widehat{\alpha}_i$ and $\widehat{\omega}$ were not conformed. To overcome this shortcomings, we apply the `QP.solve` routine with the appropriate constraint matrix to obtain reliable results. Furthermore, for $p = 4$, some simulations initially failed since the Hessian matrix was near to ill-conditioned and the evaluation of the inverse was then suitably handled by the `ginv` function. Broadly, after accommodating these computational amendments, the average number of convergent simulations turn around 92 %, 87 %, 85 % and 75 % for $p = 1, 2, 3, 4$ respectively. As noticed from these percentages, as we increase the order p, we expect some number of failed simulations.

p	T	$\omega = 1.50$	$\alpha_1 = 0.26$	$\alpha_2 = 0.16$	$\alpha_3 = 0.11$	$\alpha_4 = 0.02$
1	30	1.251 (0.321)	0.178 (0.211)			
	80	1.355 (0.151)	0.225 (0.188)			
	100	1.751 (0.101)	0.231 (0.124)			
	500	1.442 (0.087)	0.244 (0.091)			
	1000	1.542 (0.075)	0.136 (0.088)			
2	30	1.389 (0.278)	0.232 (0.209)	0.152 (0.327)		
	80	1.477 (0.150)	0.246 (0.123)	0.166 (0.111)		
	100	1.511 (0.081)	0.271 (0.099)	0.148 (0.098)		
	500	1.552 (0.032)	0.276 (0.042)	0.152 (0.038)		
	1000	1.467 (0.022)	0.255 (0.031)	0.169 (0.021)		
3	30	1.481 (0.455)	0.244 (0.303)	0.152 (0.276)	0.104 (0.152)	
	80	1.551 (0.210)	0.232 (0.155)	0.166 (0.101)	0.119 (0.110)	
	100	1.462 (0.111)	0.255 (0.101)	0.147 (0.089)	0.114 (0.088)	
	500	1. 541 (0.088)	0.275 (0.076)	0.158 (0.042)	0.121 (0.034)	
	1000	1.4601 (0.061)	0.266 (0.045)	0.166 (0.034)	0.111 (0.026)	
4	30	1.495 (0.323)	0.255 (0.212)	0.152 (0.318)	0.112 (0.176)	0.018 (0.272)
	80	1.510 (0.188)	0.276 (0.124)	0.164 (0.232)	0.114 (0.123)	0.015 (0.103)
	100	1.498 (0.092)	0.237 (0.110)	0.166 (0.101)	0.110 (0.075)	0.019 (0.064)
	500	1.502 (0.088)	0.242 (0.064)	0.167 (0.054)	0.113 (0.033)	0.017 (0.042)
	1000	1.489 (0.052)	0.262 (0.043)	0.159 (0.038)	0.109 (0.018)	0.018 (0.028)

TABLE 3.1. Simulated Mean estimates and standard errors/100 Simulations

4. Proofs and some auxiliary results

4.1. *Proofs of the main results.*

Proof of Lemma 2.1: A comparison of coefficients in (2.9) reveals that it suffices to find strictly positive constants $\gamma_1, \ldots, \gamma_p, \delta_1, \ldots, \delta_q$ such that the following inequalities are satisfied.

$$\left(\gamma_1 + \delta_1\right) c_1 + \gamma_2 \leq \kappa \gamma_1$$

$$\vdots$$

$$\left(\gamma_1 + \delta_1\right) c_{p-1} + \gamma_p \leq \kappa \gamma_{p-1}$$

$$\left(\gamma_1 + \delta_1\right) c_p \leq \kappa \gamma_p$$

$$\left(\gamma_1 + \delta_1\right) d_1 + \delta_2 \leq \kappa \delta_1$$

$$\vdots$$

$$\left(\gamma_1 + \delta_1\right) d_{q-1} + \delta_q \leq \kappa \delta_{q-1}$$

$$\left(\gamma_1 + \delta_1\right) d_q \leq \kappa \delta_q. \tag{4.1}$$

We set, w.l.o.g., $\gamma_1 + \delta_1 = 1$. Let $\epsilon = (1 - L)/(p + q)$. We consider the following system of equations.

$$c_p + \epsilon = \gamma_p$$

$$c_{p-1} + \gamma_p + \epsilon = \gamma_{p-1}$$

$$\vdots$$

$$
\begin{aligned}
c_1 + \gamma_2 + \epsilon &= \gamma_1 \\
d_q + \epsilon &= \delta_q \\
d_{q-1} + \delta_q + \epsilon &= \delta_{q-1}
\end{aligned}
$$

$$\vdots$$

$$d_1 + \delta_2 + \epsilon = \delta_1.$$

It is obvious that this system of equations has a unique solution with strictly positive $\gamma_1, \ldots, \gamma_p, \delta_1, \ldots, \delta_q$. Moreover, it follows from

$$\sum_{i=1}^{p} c_i + \sum_{j=1}^{q} d_j + \sum_{i=2}^{p} \gamma_i + \sum_{j=2}^{q} \delta_j + (p+q)\,\epsilon = \sum_{i=1}^{p} \gamma_i + \sum_{j=1}^{q} \delta_j$$

that $\gamma_1 + \delta_1 = 1$, as required. Therefore, we see that, with such a choice of $\gamma_1, \ldots, \gamma_p, \delta_1, \ldots, \delta_q$, the following strict inequalities are fulfilled.

$$
\begin{aligned}
c_1 + \gamma_2 &< \gamma_1
\end{aligned}
$$

$$\vdots$$

$$
\begin{aligned}
c_{p-1} + \gamma_p &< \gamma_{p-1} \\
c_p &< \gamma_p \\
d_1 + \delta_2 &< \delta_1
\end{aligned}
$$

$$\vdots$$

$$
\begin{aligned}
d_{q-1} + \delta_q &< \delta_{q-1} \\
d_q &< \delta_q.
\end{aligned}
$$

Choosing $\kappa = \max\{(c_1+\gamma_2)/\gamma_1, \ldots, (c_{p-1}+\gamma_p)/\gamma_{p-1}, c_p/\gamma_p, (d_1+\delta_2)/\delta_1, \ldots, (d_{q-1}+\delta_q)/\delta_{q-1}, d_q/\delta_q\}$ we obtain that the system of inequalities (4.1) is satisfied. □

Proof of Proposition 2.2: (i) follows from (2.8) and (2.9), and (ii) is an immediate consequence of (i). □

Proof of Proposition 2.3: Let Q and Q' be arbitrary probability measures supported in $S^{\geq}$ and let ξ be the optimal coupling of Q and Q' w.r.t. the Kantorovich distance, that is,

$$\mathcal{K}(Q, Q') = \int_{S^{\geq} \times S^{\geq}} \Delta_{\gamma, \delta}(z, z')\, \xi(dz, dz').$$

Then $\xi\pi^{Z,Z'}$ is a coupling of $Q\pi^Z$ and $Q'\pi^Z$ and it follows from Proposition 2.2(i) that

$$
\begin{aligned}
\mathcal{K}(Q\pi, Q'\pi) &\leq \int \Delta_{\gamma, \delta}(u, u')\, \xi\pi^{Z,Z'}(du, du') \\
&= \int \left[\int \Delta_{\gamma, \delta}(u, u')\, \pi^{Z,Z'}((z, z'), du\, du') \right] \xi(dz, dz')
\end{aligned}
$$

$$\leq \quad \kappa \int \Delta_{\gamma,\delta}(z, z') \, \xi(dz, dz') \;=\; \kappa \, \mathcal{K}(Q, Q').$$

$\square$

Proof of Theorem 2.4: We consider first the Markov process $(Z_t)_{t\in\mathbb{Z}}$. Let

$$\mathcal{P} = \left\{ Q\colon\; Q \text{ is a probability distribution based in } S^{\geq}, \int_S \sum_{i=1}^{2p} |x_i| \, Q(dx) < \infty \right\}.$$

It is well known that the space $\mathcal{P}$ equipped with the Kantorovich metric $\mathcal{K}$ is complete. Since by Proposition 2.3 the mapping π^Z is contractive it follows by the Banach fixed point theorem that the Markov kernel π^Z admits a unique fixed point Q^Z, i.e. $Q^Z \pi^Z = Q^Z$. In other words, Q^Z is the unique stationary distribution of the process $(Z_t)_{t\in\mathbb{Z}}$.

Now we consider the process $(Y_t)_{t\in\mathbb{Z}}$. If the X_t are non-negative random variables, then we have a one-to-one relationship between Z_t and Y_t and, for

$$Z_0 = \big(X_0^2, \ldots, X_{1-p}^2, v_0, \ldots, v_{1-q}\big) \sim Q^Z,$$

the distribution of the vector $Y_0 = \big(X_0, \ldots, X_{1-p}, v_0, \ldots, v_{1-q}\big)$ is the unique stationary distribution of $(Y_t)_{t\in\mathbb{Z}}$.

If X_t attains both positive and negative values, we need a simple extra argument. Suppose that $Z_0 = \big(X_0^2, \ldots, X_{1-p}^2, v_0, \ldots, v_{1-q}\big) \sim Q^Z$. Now we can recursively generate suitable $(v_1, X_1), \ldots, (v_p, X_p)$ as follows.
We set $v_1 = f\big(X_0^2, \ldots, X_{1-p}^2, v_0, \ldots, v_{1-q}\big)$ and generate $X_1 \sim Q_{v_1}$. Then we set $v_2 = f\big(X_1^2, \ldots, X_{-p}^2, v_1, \ldots, v_{-q}\big)$ and choose $X_1 \sim Q_{v_1}$, and so on. After p such steps we have collected enough X_ts with suitable signs and the random vector $Y_p := \big(X_p, \ldots, X_1, v_p, \ldots, v_{p-q+1}\big)$ has the unique stationary distribution, say Q^Y, of $(Y_t)_{t\in\mathbb{Z}}$. $\square$

Proof of Theorem 2.6: Since by (2.10) $\widetilde{E}\Delta_{\gamma,\delta}(\widetilde{Z}_0, \widetilde{Z}_0') < \infty$, this theorem is an immediate consequence of (2.13) and (2.15). $\square$

4.2. *Some auxiliary results.*

Lemma 4.1. *Let $((X_t, v_t))_{t\in\mathbb{Z}}$ be a stationary process satisfying (3.1), where ω, $\alpha_1, \ldots, \alpha_p$ are non-negative constants and let $X_t \mid \mathcal{F}_{t-1} \sim \mathrm{Skellam}(v_t/2, v_t/2)$.*

   (i) *If $\alpha = \sum_{i=1}^{n} \alpha_i < 1$, then $EX_0^2 < \infty$.*

   (ii) *If $\alpha = \sum_{i=1}^{n} \alpha_i < \dfrac{1}{\sqrt{3}}$, then $EX_0^4 < \infty$.*

Proof of Lemma 4.1: Let $((\widetilde{X}_t, \widetilde{v}_t))_{t\in\mathbb{N}}$ be Skellam-ARCH process satisfying (3.1), but with initial values $\widetilde{X}_1 = \cdots = \widetilde{X}_p = \sqrt{\omega}$. (The latter condition is imposed to ensure that $E\widetilde{X}_1^4, \ldots, E\widetilde{X}_p^4$ are guaranteed to be finite.) Since $\widetilde{X}_n \xrightarrow{d} X_0$ it follows from Theorem III.6.31 in Pollard (1984, page 58) that we can construct a coupling of these random variables where we have almost sure convergence rather than convergence in probability. Hence, we obtain by Fatou's lemma that

$$EX_0^k \leq \liminf_{n\to\infty} E\widetilde{X}_n^k, \quad \text{for } k = 2, 4. \tag{4.2}$$

(i) It follows from (3.1) that, for $t > p$,

$$E\widetilde{X}_t^2 = \widetilde{v}_t \leq \omega + \alpha \max\{E\widetilde{X}_{t-1}^2, \ldots, E\widetilde{X}_{t-p}^2\}.$$

Let $Z_t = \max\{E\widetilde{X}_t^2, \ldots, E\widetilde{X}_{t-p+1}^2\}$. We obtain from the previous display the recursion

$$Z_t \leq \max\{\omega + \alpha Z_{t-1}, Z_{t-1}\}.$$

Therefore,

$$E\widetilde{X}_t^2 \leq \frac{\omega}{1-\alpha},$$

which yields in conjunction with (4.2) that (i) holds true.

(ii) If $X \sim \text{Skellam}(v/2, v/2)$, then $EX^4 = v + 3v^2$. Hence, for $t > p$

$$
\begin{aligned}
E\widetilde{X}_t^4 &= \omega + \sum_{i=1}^p \alpha_i E\widetilde{X}_{t-i}^2 + 3 E\left[\left(\omega + \sum_{i=1}^p \alpha_i \widetilde{X}_{t-i}^2\right)^2\right] \\
&\leq \omega + 3\omega^2 + (1 + 6\omega)\alpha \max\{E\widetilde{X}_{t-1}^2, \ldots, E\widetilde{X}_{t-p}^2\} \\
&\quad + 3\alpha^2 \max\{E\widetilde{X}_{t-1}^4, \ldots, E\widetilde{X}_{t-p}^4\}.
\end{aligned}
$$

With $\bar{Z}_t = \max\{E\widetilde{X}_t^4, \ldots, E\widetilde{X}_{t-p+1}^4\}$ and $\bar{\omega} = \omega + 3\omega^2 + (1+6\omega)\alpha \cdot \frac{\omega}{1-\alpha}$, we obtain the recursion

$$\bar{Z}_t \leq \max\{\bar{\omega} + 3\alpha^2 \bar{Z}_{t-1}, \bar{Z}_{t-1}\},$$

which leads to

$$E\widetilde{X}_t^4 \leq \frac{\bar{\omega}}{1 - 3\alpha^2}.$$

(ii) follows now from (4.2).

$\square$

Lemma 4.2. *Let $((\widetilde{X}_t, \widetilde{\lambda}_t))_{t\in\mathbb{Z}}$ be a stationary Skellam-ARCH process satisfying (2.1a) and (3.1) and with $\omega > 0$. Then the matrix Σ defined in (3.4) is regular.*

Proof of Lemma 4.2: We have that

$$\Sigma = E[ZZ^T],$$

where $Z = (1, X_p^2, \ldots, X_1^2)^T$.

Assume that Σ is singular: then there exists some $\gamma = (\gamma_0, \ldots, \gamma_p)^T \neq 0_{p+1}$ such that

$$0 = \gamma^T \Sigma \gamma = E[(Z^T\gamma)^2],$$

which implies that

$$P\left(Z^T\gamma = 0\right) = 1.$$

This means that

$$\gamma_0 + \sum_{i=1}^p X_{t-i+1}^2 \gamma_i = 0$$

holds with probability 1. Since $\gamma_1 = \cdots = \gamma_p = 0$ would then imply that $\gamma = 0_{p+1}$, there exists some $i_0 \geq 1$ such that $\gamma_1 = \cdots = \gamma_{i_0-1} = 0$ and $\gamma_{i_0} \neq 0$. Then

$$X_{t-i_0}^2 = \frac{1}{\gamma_{i_0}}\left\{\gamma_0 + \sum_{i=i_0+1}^p \gamma_i X_{t-i+1}^2\right\}$$

that is, $X_{t-i_0}^2$ is fully determined by the past values of the count process. This, however, leads to a contradiction since

$$X_{t-i_0} \mid \mathcal{F}_{t-i_0-1} \sim \text{Skellam}(v_{t-i_0}/2, v_{t-i_0}/2)$$

with $v_{t-i_0} \geq \omega > 0$. Hence, Σ is a regular matrix. $\square$

Acknowledgements

References

Alomani, G. A., Alzaid, A. A., and Omair, M. A. A Skellam GARCH model. *Braz. J. Probab. Stat.*, **32** (1), 200–214 (2018). MR3770870.

Alzaid, A. A. and Omair, M. A. On the Poisson difference distribution inference and applications. *Bull. Malays. Math. Sci. Soc. (2)*, **33** (1), 17–45 (2010). MR2603338.

Andersson, J. and Karlis, D. A parametric time series model with covariates for integers in Z. *Statistical Modelling*, **14** (2), 135–156 (2014). DOI: 10.1177/1471082X13504719.

Bollerslev, T. Generalized autoregressive conditional heteroskedasticity. *J. Econometrics*, **31** (3), 307–327 (1986). MR853051.

Dedecker, J., Doukhan, P., Lang, G., León R., J. R., Louhichi, S., and Prieur, C. *Weak dependence: with examples and applications*, volume 190 of *Lecture Notes in Statistics*. Springer, New York (2007). ISBN 978-0-387-69951-6. MR2338725.

Douc, R., Moulines, E., Priouret, P., and Soulier, P. *Markov chains*. Springer Series in Operations Research and Financial Engineering. Springer, Cham (2018). ISBN 978-3-319-97703-4; 978-3-319-97704-1. MR3889011.

Doukhan, P., Leucht, A., and Neumann, M. H. Mixing properties of non-stationary INGARCH(1, 1) processes (2020). Manuscript.

Doukhan, P. and Louhichi, S. A new weak dependence condition and applications to moment inequalities. *Stochastic Process. Appl.*, **84** (2), 313–342 (1999). MR1719345.

Doukhan, P. and Neumann, M. H. The notion of ψ-weak dependence and its applications to bootstrapping time series. *Probab. Surv.*, **5**, 146–168 (2008). MR2426177.

Doukhan, P. and Neumann, M. H. Absolute regularity of semi-contractive GARCH-type processes. *J. Appl. Probab.*, **56** (1), 91–115 (2019). MR3981148.

Doukhan, P. and Wintenberger, O. Weakly dependent chains with infinite memory. *Stochastic Process. Appl.*, **118** (11), 1997–2013 (2008). MR2462284.

Eberle, A. Markov processes, Lecture Notes at University of Bonn (2019). https://wt.iam.uni-bonn.de/fileadmin/WT/Inhalt/people/Andreas_Eberle/MarkovProcesses1920/MarkovProcesses1920.pdf.

Fokianos, K. and Tjø stheim, D. Nonlinear Poisson autoregression. *Ann. Inst. Statist. Math.*, **64** (6), 1205–1225 (2012). MR2981620.

Fokianos, K. and Tjøstheim, D. Log-linear Poisson autoregression. *Journal of Multivariate Analysis*, **102** (3), 563–578 (2011). DOI: 10.1016/j.jmva.2010.11.002.

Hall, P. and Heyde, C. C. *Martingale limit theory and its application.* Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York-London (1980). ISBN 0-12-319350-8. MR624435.

Irwin, J. O. The frequency distribution of the difference between two independent variates following the same Poisson distribution. *Journal of the Royal Statistical Society, Series A*, **100**, 415–416 (1937).

Karlis, D. and Ntzoufras, I. Bayesian analysis of the differences of count data. *Stat. Med.*, **25** (11), 1885–1905 (2006). MR2227435.

Neumann, M. H. Absolute regularity and ergodicity of Poisson count processes. *Bernoulli*, **17** (4), 1268–1284 (2011). MR2854772.

Neumann, M. H. Bootstrap for integer-valued GARCH(p,q) processes (2020). To appear in *Statistica Neerlandica*.

Pollard, D. *Convergence of stochastic processes.* Springer Series in Statistics. Springer-Verlag, New York (1984). ISBN 0-387-90990-7. MR762984.

Skellam, J. G. The frequency distribution of the difference between two Poisson variates belonging to different populations. *J. Roy. Statist. Soc. (N.S.)*, **109**, 296 (1946). MR20750.