Journal of Theoretical Probability (2026) 39:18
https://doi.org/10.1007/s10959-025-01475-0

Nonlinear Marked Poisson Autoregression: Stability and Rate of Convergence to Equilibrium

Matthias Kirchner[1] · Giovanni Luca Torrisi[2]

Received: 15 July 2025 / Revised: 14 November 2025 / Accepted: 16 December 2025

Abstract

We consider nonlinear marked Poisson autoregressions and we prove the existence and uniqueness (in law) of the ergodic version under various assumptions on the transfer function, the coefficient functions, and the mark distribution. Moreover, by means of coupling techniques we discuss the convergence and the rate of convergence to equilibrium of a non-ergodic version of the time series, starting from some "initial condition", to the ergodic version. The theoretical results are illustrated by examples and numerical simulations.

Keywords Nonlinear time series · Nonlinear autoregression · Ergodic processes · Rate of convergence to equilibrium

Mathematics Subject Classification (2020) Primary 62M10 · 60G65 · 60G10

1 Introduction

A time series $(X_n)_{n\in\mathbb{Z}}$, $\mathbb{Z} := \{0, \pm 1, \pm 2, \ldots\}$, is a *nonlinear marked Poisson autoregression* if

$$X_n \mid \mathcal{F}^X_{n-1} \vee \mathcal{F}^Z_{n-1} \sim \mathrm{Pois}(\lambda_n), \quad n \in \mathbb{Z},$$

where

$$\lambda_n = \phi\left(\sum_{k=-\infty}^{n-1}\sum_{i=1}^{X_k}\alpha_{n-k}(Z_{k,i})\right) = \phi\left(\sum_{k=1}^{\infty}\sum_{i=1}^{X_{n-k}}\alpha_k(Z_{n-k,i})\right) \tag{1}$$

✉ Giovanni Luca Torrisi
giovanniluca.torrisi@cnr.it

Matthias Kirchner
matthias.kirchner@phbern.ch

[1] University of Teacher Education Bern, Fabrikstrasse 8, 3011 Bern, Switzerland

[2] Consiglio Nazionale delle Ricerche, via dei Taurini 19, 00185 Rome, Italy

Published online: 05 January 2026

with

- $\mathcal{F}_n^X := \sigma(X_k : k \leq n), n \in \mathbb{Z}$,
- $(Z_{k,i})_{k\in\mathbb{Z}, i\in\mathbb{N}}$, $\mathbb{N} := \{1, 2, \ldots\}$, identically distributed *mark random variables* with values in a measurable space $(\mathrm{Z}, \mathcal{Z})$ such that the mark sequences $(Z_{k,i})_{i\in\mathbb{N}}$ are independent over $k \in \mathbb{Z}$, i.e., the sequences

$$\ldots, (Z_{-1,i})_{i\in\mathbb{N}}, (Z_{0,i})_{i\in\mathbb{N}}, (Z_{1,i})_{i\in\mathbb{N}}, \ldots$$

  are independent. In particular, for each time $k_0 \in \mathbb{Z}$, the mark sequence $(Z_{k_0,i})_{i\in\mathbb{N}}$ is independent of $\mathcal{F}_{k_0}^X$.
- $\alpha_k : \mathrm{Z} \to \mathbb{R}$, $k \in \mathbb{Z}$, measurable *coefficient functions* (setting $\alpha_k(\cdot) := 0$ for $k \in \mathbb{Z}_{\leq 0} := \{0, -1, -2, \ldots\}$) such that $\sum_{k=1}^{\infty} \mathbb{E}\left[|\alpha_k(Z_{0,1})|\right] := \sum_{k=1}^{\infty} \overline{\alpha}_k = 1$,
- $\mathcal{F}_n^Z := \sigma(Z_{k,i} : k \leq n, i \in \mathbb{N}), n \in \mathbb{Z}$,
- $\mathrm{Pois}(\mu)$, $\mu \geq 0$, the Poisson law with mean μ (where a random variable $X \sim \mathrm{Pois}(0)$ is equal to zero almost surely),
- and $\phi : \mathbb{R} \to \mathbb{R}_{\geq 0} := [0, \infty)$, a measurable *transfer function*.

Note that the two representations of the dynamics (1) stem from a change of variable. The second representation is more typical in a time series context, whereas the first one is frequently used in point processes theory. We drew inspiration for the present work from the papers [6, 7, 19] on nonlinear Hawkes point processes. This is why we stick to the "point-process-type" indexing.

First papers on (linear) counting autoregressions are [1] for $p = 1$, and [2] for general $p \in \mathbb{N}$, where p denotes the autoregressive order of the time series. These models go under the name INteger-valued AutoRegressive (INAR) time series. In both articles, binomial thinning is applied to the earlier counts and the innovations are not necessarily Poisson, yielding a non-Poisson conditional distribution. [16] proves the existence of the model for the case $p = \infty$, assuming Poisson reproduction and innovation, and formalizes the analogy to Hawkes point processes. In [17] the authors provide precise deviations and extended central limit theorems for cumulative INAR models with $p = \infty$, and again Poisson reproduction and innovation. The analogy between INAR time series with Poisson conditional distributions and Hawkes point processes makes it unsurprising that some authors use the name "discrete-time Hawkes process" as a synonym for "Poisson autoregression"; see, e.g., [15, 20, 21]. The term "Poisson autoregression" appears for the first time in [12], where the linear case is studied. The choice for the conditional Poisson distribution for counts of events in some interval might be justified by the Poisson limit theorem: we observe (and count) one of many possible but all very unlikely events in one time unit. Furthermore, the Poisson choice makes the conditional distribution tractable and allows the Poisson embedding construction which is very helpful in the proofs.

In the domain of counting time series, we see two main motivations for nonlinear models. First of all, in some applications, there might exist a physical barrier such that the conditional expectations λ_n in (1) cannot exceed a certain relevant threshold. Think of the conditional expectations λ_n depending on some bounded temperature, voltage, or tension. In this case, we typically want to model this situation by a bounded transfer function ϕ, which introduces the nonlinearity. Secondly, there are applications where the occurrence of events at one time point might typically *decrease* the occurrence of events at the next time point. Think of the number of extreme weather events per week, the number of volcanic eruptions per month, the number of fish catches over a certain threshold per year, or the number of social revolutions per decade. To model situations like these, one typically needs to allow for possible negative values of the coefficient functions $\alpha_k(\cdot)$. Therefore, to guarantee a nonnegative conditional expectation λ_n, the transfer function ϕ has to be nonlinear. Here, typical choices would be $\phi : x \mapsto \max\{mx + c, 0\}$ or $x \mapsto \log(1 + \exp(mx + c))$ for some $m, c \in \mathbb{R}$.

Many variants of Poisson autoregression have been considered in literature. [13] investigates the nonlinear Poisson autoregression $(X_n)_{n\in\mathbb{Z}}$ with dynamics

$$\lambda_n = f(\lambda_{n-1}) + b(X_{n-1}),$$

proving the existence of an ergodic version of the process under a Lipschitz condition on f and b. [8] considers the more general dynamics

$$\lambda_n = f(X_{n-1}, \lambda_{n-1}),$$

and again shows the ergodicity of the time series under a Lipschitz condition on f. Another relevant contribution is [9], where a similar model like ours is considered. In this article, at each time point the choice of the transfer function ϕ depends on the value of some Markov chain. If one ignores the Markov modulation in the model of [9] and the marks in our model, then the time series studied in [9] coincides with ours, and our Theorem 2 is a special case of Theorem 2.1 in [9]. We emphasize that we use completely different techniques, which pave the way to prove existence and uniqueness of an ergodic version of the time series even when the transfer function is bounded, Lipschitz but not contractive (see Theorem 3) or sublinear and lower-semi-continuous (see Theorem 4). Furthermore, [8, 9, 13] focus on inference problems, whereas we focus on convergence to equilibrium by using coupling techniques which allow us to derive explicit upper bounds and rates for the total variation distance between the laws of an ergodic and a non-ergodic version of the time series (see Sects. 3 and 4).

As a further simple idea that might be relevant in applications, we introduce in the model identically distributed marks, that are attached to each of the events that are counted. These marks potentially modulate the autoregression coefficients: a "strong" event can have a different impact on the future of the count time series than a "weak" event. As we only assume identical distribution together with independence from the past, the marks corresponding to events at a specific point of time k might be correlated or even identical. Moreover, due to the very general form of the mark space, the marks might model the location of an event.

The paper is organized as follows. In Sect. 2 we study the ergodicity of the model under various assumptions on the parameters of the nonlinear marked Poisson autoregression. In Sects. 3 and 4 we investigate the convergence to equilibrium and the rate of convergence to equilibrium of the model, respectively. In Sect. 5 we provide some examples and numerical illustrations. In Sect. 6 we briefly discuss a multivariate version of the nonlinear marked Poisson autoregression and some open problems which we plan to investigate in the next future.

2 Ergodicity

One of the aims of this paper is to give sufficient conditions on the transfer function ϕ, on the coefficient functions α_k, and on the distribution of the marks $Z_{k,i}$ for the existence of a nontrivial ergodic version of the marked Poisson autoregression. We refer the reader to [4] for an introduction to the concepts (used throughout this article) of flow, flow compatibility, flow invariance, stationarity and ergodicity of stochastic processes.

We start with a negative result.

Theorem 1 *(Non-existence) There does not exist an ergodic nonlinear Poisson autoregression $(X_n)_{n\in\mathbb{Z}}$ with*

- *mean $\lambda := \mathbb{E}\lambda_0 = \mathbb{E}X_0 \in (0,\infty)$,*
- *nonnegative coefficient functions $\alpha_k : \mathrm{Z} \to \mathbb{R}_{\geq 0}$, $k \in \mathbb{N}$,*
- *transfer function ϕ such that $\phi(x) \leq Rx$ for any $x \geq 0$ and some $R > 0$, and*

$$\sum_{k=1}^{\infty} k\overline{\alpha}_k < \infty, \quad \textit{with } \overline{\alpha}_k := \mathbb{E}[|\alpha_k(Z_{0,1})|],\ k \in \mathbb{N}. \tag{2}$$

Proof We first show by induction that

$$\mathbb{P}\left(\sum_{k=1}^{n} X_k = 0 \,|\, \mathcal{F}_0^X \vee \mathcal{F}_0^Z\right) = \exp\left\{-\sum_{k=1}^{n} \phi\left(\sum_{i=-\infty}^{0} \sum_{j=1}^{X_i} \alpha_{k-i}(Z_{i,j})\right)\right\} \tag{3}$$

for all $n \in \mathbb{N}$. For $n = 1$ the claim is true as clearly $\mathbb{P}\left(X_1 = 0 \,|\, \mathcal{F}_0^X \vee \mathcal{F}_0^Z\right) = \mathrm{e}^{-\lambda_1}$. So suppose that the claim is true for $n-1$. We have

$$\begin{aligned}
&\mathbb{P}\left(\sum_{k=1}^{n} X_k = 0 \,|\, \mathcal{F}_{n-1}^X \vee \mathcal{F}_{n-1}^Z\right)\\
&= \mathbb{E}\left[\mathbf{1}\left(\sum_{k=1}^{n-1} X_k = 0\right)\mathbf{1}\left(X_n = 0\right) \,|\, \mathcal{F}_{n-1}^X \vee \mathcal{F}_{n-1}^Z\right]\\
&= \mathbf{1}\left(\sum_{k=1}^{n-1} X_k = 0\right)\mathbb{P}\left(X_n = 0 \,|\, \mathcal{F}_{n-1}^X \vee \mathcal{F}_{n-1}^Z\right)\\
&= \mathbf{1}\left(\sum_{k=1}^{n-1} X_k = 0\right)\mathrm{e}^{-\lambda_n}\\
&= \mathbf{1}\left(\sum_{k=1}^{n-1} X_k = 0\right)\exp\left\{-\phi\left(\sum_{k=-\infty}^{0} \sum_{i=1}^{X_k} \alpha_{n-k}(Z_{k,i})\right)\right\}.
\end{aligned}$$

Taking the conditional expectation with respect to $\mathcal{F}_0^X \vee \mathcal{F}_0^Z$ and applying the induction hypothesis, we obtain

$$\begin{aligned}
&\mathbb{P}\left(\sum_{k=1}^{n} X_k = 0 \,\middle|\, \mathcal{F}_0^X \vee \mathcal{F}_0^Z\right) \\
&= \mathbb{P}\left(\sum_{k=1}^{n-1} X_k = 0 \,\middle|\, \mathcal{F}_0^X \vee \mathcal{F}_0^Z\right) \exp\left\{-\phi\left(\sum_{i=-\infty}^{0} \sum_{j=1}^{X_i} \alpha_{n-i}(Z_{i,j})\right)\right\} \\
&= \exp\left\{-\sum_{k=1}^{n-1} \phi\left(\sum_{i=-\infty}^{0} \sum_{j=1}^{X_i} \alpha_{k-i}(Z_{i,j})\right)\right\} \exp\left\{-\phi\left(\sum_{i=-\infty}^{0} \sum_{j=1}^{X_i} \alpha_{n-i}(Z_{i,j})\right)\right\}.
\end{aligned}$$

It follows that (3) is true for all $n \in \mathbb{N}$. Note that, as the sequence of events

$$\left(\sum_{k=1}^{n} X_k = 0\right)_{n \in \mathbb{N}}$$

is nonincreasing in n and the sequence on the right hand side of (3) is a.s. nonincreasing, taking the limit as $n \to \infty$ in equation (3), we have that this relation remains true when $n = \infty$. Now suppose that there exists an ergodic version of the time series with the properties described in the statement of the theorem. By (3) with $n = \infty$ we have that

$$\begin{aligned}
\mathbb{P}\left(\sum_{k=1}^{\infty} X_k = 0 \,\middle|\, \mathcal{F}_0^X \vee \mathcal{F}_0^Z\right) &= \exp\left\{-\sum_{k=1}^{\infty} \phi\left(\sum_{i=-\infty}^{0} \sum_{j=1}^{X_i} \alpha_{k-i}(Z_{i,j})\right)\right\} \\
&\geq \exp\left\{-R \sum_{k=1}^{\infty} \sum_{i=-\infty}^{0} \sum_{j=1}^{X_i} \alpha_{k-i}(Z_{i,j})\right\}.
\end{aligned}$$

Taking expectations and applying Jensen's inequality, we obtain

$$\mathbb{P}\left(\sum_{k=1}^{\infty} X_k = 0\right) \geq \exp\left\{-R \sum_{k=1}^{\infty} \sum_{i=-\infty}^{0} \mathbb{E}\left[\sum_{j=1}^{X_i} \alpha_{k-i}(Z_{i,j})\right]\right\}. \tag{4}$$

For the expectation, we apply a general form of Wald's identity. Indeed, from our assumptions, we have that for all $i \in \mathbb{Z}$ and for all $k \in \mathbb{N}$, the counting variable X_i and the sequence $(\alpha_{k-i}(Z_{i,j}))_{j \in \mathbb{N}}$ are independent. Furthermore, we have that $\mathbb{E}[X_i] = \lambda < \infty$ and $\mathbb{E}[\alpha_{k-i}(Z_{i,j})] \equiv \overline{\alpha}_{k-i} < \infty$, $j \in \mathbb{N}$. Thus,

$$\mathbb{P}\left(\sum_{k=1}^{\infty} X_k = 0\right) \geq \exp\left\{-\lambda R \sum_{k=1}^{\infty} \sum_{i=-\infty}^{0} \overline{\alpha}_{k-i}\right\} = \exp\left\{-\lambda R \sum_{k=1}^{\infty} \sum_{h=k}^{\infty} \overline{\alpha}_h\right\}$$
$$= \exp\left\{-\lambda R \sum_{k=1}^{\infty} k\overline{\alpha}_k\right\} > 0,$$

where the latter inequality follows by the assumptions (2) and $\lambda < \infty$. Since we are assuming that the time series is ergodic, we then have that

$$\mathbb{P}\left(\sum_{n=1}^{\infty} X_n = 0\right) = 1.$$

Note that this implies $\mathbb{P}(X_n = 0) = 1$ for each $n \in \mathbb{N}$, so that $\lambda = \mathbb{E}X_1 = 0$, which is in contradiction with the original assumption $\lambda \in (0, \infty)$. The proof is completed. □

From Theorem 1 it follows that, if an ergodic marked Poisson autoregression with coefficient functions $\alpha_k : \mathrm{Z} \to \mathbb{R}_{\geq 0}, k \in \mathbb{N}$, transfer function $\phi : x \mapsto x, x \in \mathbb{R}_{\geq 0}$, and mean $\lambda \in (0, \infty)$ exists, then necessarily $\sum_{k=1} k\overline{\alpha}_k = \infty$. Note that, taking expectations on both sides of (1), and dividing by λ, we have $\sum_{k\geq 1} \overline{\alpha}_k = 1$. Therefore, if an ergodic linear marked Poisson autoregression with non-negative coefficient functions and finite and positive mean exists, then $(\overline{\alpha}_k)_{k\geq 1}$ specifies a probability distribution supported on $\mathbb{N}$ with infinite mean.

Next, we establish the existence of an ergodic version of a marked nonlinear Poisson autoregression, under different assumptions on ϕ, the coefficient functions α_k as well as the mark distribution.

Theorem 2 (Existence for Lipschitz transfer) *Suppose that*

$$\phi \textit{ is Lipschitz continuous with Lipschitz constant } \mathrm{Lip}(\phi) < 1. \tag{5}$$

Then there exists an ergodic version of the nonlinear marked Poisson autoregression with finite mean.

Proof Let N be a Poisson process on $\mathbb{R}\times\mathbb{R}_{\geq 0}$ with mean measure equal to the Lebesgue measure. We assume that N is defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which the random variables $Z = \left(Z_{n,i}\right)_{n\in\mathbb{Z}, i\in\mathbb{N}}$ are defined and that N is independent of Z. We assume that the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is provided with a flow $(\theta_n)_{n\in\mathbb{Z}}$ with respect to which $\mathbb{P}$ is ergodic and such that N and $\left(Z_{\cdot,i}\right)_{i\in\mathbb{N}}$ are (θ_n)-compatible (this assumption is not restrictive since we can always consider $\Omega := \Omega_N \times \Omega_Z$, $\mathcal{F} := \mathcal{F}_N \otimes \mathcal{F}_Z$ and $\mathbb{P} := \mathbb{P}_N \otimes \mathbb{P}_Z$, where $(\Omega_N, \mathcal{F}_N, \mathbb{P}_N)$ and $(\Omega_Z, \mathcal{F}_Z, \mathbb{P}_Z)$ are the canonical spaces of N and Z, respectively, and, using an obvious notation, we define $N(\omega) = N((\omega_N, \omega_Z)) := N(\omega_N) = \omega_N$, $Z(\omega) := Z(\omega_Z) = \omega_Z$ and $\theta_n\omega = (\theta_n^N \omega_N, \theta_n^Z \omega_Z)$, where, roughly speaking, θ_n^N is the shift of value n of the measure ω_N along the x-axis and θ_n^Z is the shift of value n of the first index of the double sequence ω_Z).

For $n \in \mathbb{Z}$, set $\lambda_n^{(0)} \equiv 0$. Furthermore, for $n \in \mathbb{Z}$ and $\ell \in \mathbb{N}$, define the recursions:

$$X_n^{(\ell)} := N\left([n-1, n), (0, \lambda_n^{(\ell-1)}]\right) \tag{6}$$

$$\lambda_n^{(\ell)} := \phi\left(\sum_{k=-\infty}^{n-1} \sum_{i=1}^{X_k^{(\ell)}} \alpha_{n-k}(Z_{k,i})\right). \tag{7}$$

By construction, for each $\ell \in \mathbb{N}$, the processes $\left(\lambda_n^{(\ell)}\right)_{n\in\mathbb{Z}}$ and $\left(X_n^{(\ell)}\right)_{n\in\mathbb{Z}}$ are (θ_n)-compatible, $n \in \mathbb{Z}$, and therefore ergodic. Let $n \in \mathbb{Z}$ be fixed. The (θ_n)-compatibility of $\lambda_n^{(\ell)}$ and $X_n^{(\ell)}$ can be proven by induction over $\ell \in \mathbb{N}$. Note that $\lambda_n^{(0)}$ is (θ_n)-compatible, indeed $\lambda_0^{(0)}(\theta_n\omega) = 0 = \lambda_n^{(0)}(\omega)$. We have

$$\begin{aligned} X_0^{(1)}(\theta_n\omega) &= N([-1, 0), (0, \lambda_0^{(0)}(\theta_n\omega)])(\theta_n\omega) = N([n-1, n), (0, \lambda_n^{(0)}(\omega)])(\omega) \\ &= X_n^{(1)}(\omega) \end{aligned}$$

and

$$\begin{aligned} \lambda_0^{(1)}(\theta_n\omega) &= \phi\left(\sum_{k=-\infty}^{-1} \sum_{i=1}^{X_k^{(1)}(\theta_n\omega)} \alpha_{-k}(Z_{k,i}(\theta_n\omega))\right) \\ &= \phi\left(\sum_{k=-\infty}^{-1} \sum_{i=1}^{X_{n+k}^{(1)}(\omega)} \alpha_{-k}(Z_{n+k,i}(\omega))\right) \\ &= \phi\left(\sum_{k=-\infty}^{n-1} \sum_{i=1}^{X_k^{(1)}(\omega)} \alpha_{n-k}(Z_{k,i}(\omega))\right) = \lambda_n^{(1)}(\omega). \end{aligned}$$

Now assume that $\lambda_n^{(\ell-1)}$ and $X_n^{(\ell-1)}$ are (θ_n)-compatible. Then it is obvious that $X_n^{(\ell)}$ is (θ_n)-compatible, and so it immediately follows that $\lambda_n^{(\ell)}$ is (θ_n)-compatible. The proof of the (θ_n)-compatibility of $\left(\lambda_n^{(\ell)}\right)_{n\in\mathbb{Z}}$ and $\left(X_n^{(\ell)}\right)_{n\in\mathbb{Z}}$ is completed.

Let $n \in \mathbb{Z}$ be arbitrarily fixed. We prove by induction that $\left(Z_{n,i}\right)_{i\in\mathbb{N}}$ is independent of $\mathcal{F}_n^{X^{(\ell)}}$ for each $\ell \in \mathbb{N}$. For $\ell = 1$ the claim is obvious as $\mathcal{F}_n^{X^{(1)}} = \{\emptyset, \Omega\}$. Suppose that $\left(Z_{n,i}\right)_{i\in\mathbb{N}}$ is independent of $\mathcal{F}_n^{X^{(\ell-1)}}$, then $\left(Z_{n,i}\right)_{i\in\mathbb{N}}$ is independent of $\mathcal{F}_n^{\lambda^{(\ell-1)}}$, and so $\left(Z_{n,i}\right)_{i\in\mathbb{N}}$ is independent of $\mathcal{F}_n^{X^{(\ell)}}$ (recall that by construction the sequence Z is independent of the Poisson measure N).

For $\ell \in \mathbb{N}$, we have (using the Lipschitz property of ϕ and the independence of $\left(Z_{k,i}\right)_{i\in\mathbb{N}}$ from $\mathcal{F}_k^{X^{(\ell)}}$, for each $\ell \in \mathbb{N}$, and applying Wald's identity as after relation (4))

$$\begin{aligned}
\mathbb{E}\left|\lambda_n^{(\ell+1)} - \lambda_n^{(\ell)}\right| &\leq \operatorname{Lip}(\phi) \sum_{k=-\infty}^{n-1} \overline{\alpha}_{n-k} \mathbb{E}\left|X_k^{(\ell+1)} - X_k^{(\ell)}\right| \\
&\leq \operatorname{Lip}(\phi) \sum_{k=-\infty}^{n-1} \overline{\alpha}_{n-k} \mathbb{E}\left|N([k-1,k),(0,\lambda_k^{(\ell)}]) - N([k-1,k),(0,\lambda_k^{(\ell-1)}])\right| \\
&= \operatorname{Lip}(\phi) \sum_{k=-\infty}^{n-1} \overline{\alpha}_{n-k} \mathbb{E}\left|\lambda_k^{(\ell)} - \lambda_k^{(\ell-1)}\right|.
\end{aligned}$$

And so, by stationarity,

$$\begin{aligned}
\mathbb{E}\left|\lambda_0^{(\ell+1)} - \lambda_0^{(\ell)}\right| &\leq \operatorname{Lip}(\phi) \sum_{k=-\infty}^{n-1} \overline{\alpha}_{n-k} \mathbb{E}\left|\lambda_0^{(\ell)} - \lambda_0^{(\ell-1)}\right| = (\operatorname{Lip}(\phi))\, \mathbb{E}\left|\lambda_0^{(\ell)} - \lambda_0^{(\ell-1)}\right| \\
&\leq (\operatorname{Lip}(\phi))^{\ell}\, \mathbb{E}\left|\lambda_0^{(1)}\right| = \phi(0)\, (\operatorname{Lip}(\phi))^{\ell}. \qquad (8)
\end{aligned}$$

Therefore, by assumption (5), we have that $\left(\lambda_0^{(\ell)}\right)_{\ell\in\mathbb{N}}$ is a Cauchy sequence in $L^1(\mathbb{P})$ and consequently

$$\lambda_n^{(\ell)} \text{ converges to some limit } \lambda_n^{(\mathrm{erg})} \text{ in } L^1(\mathbb{P}) \text{ as } \ell \to \infty. \qquad (9)$$

In particular, note that $\lambda^{(\mathrm{erg})} := \mathbb{E}\lambda_0^{(\mathrm{erg})}$ is finite. By Markov's inequality, for any $\ell \in \mathbb{N}$, we have

$$\begin{aligned}
&\mathbb{P}\left(\left|\lambda_0^{(\ell+1)} - \lambda_0^{(\ell)}\right| > (\operatorname{Lip}(\phi))^{\ell/2}\right) \\
&\qquad \leq (\operatorname{Lip}(\phi))^{-\ell/2}\, \mathbb{E}\left|\lambda_0^{(\ell+1)} - \lambda_0^{(\ell)}\right| \leq \phi(0)\, (\operatorname{Lip}(\phi))^{\ell/2},
\end{aligned}$$

where the latter inequality follows by (8). By assumption (5) we deduce that

$$\sum_{\ell\in\mathbb{N}} \mathbb{P}\left(\left|\lambda_0^{(\ell+1)} - \lambda_0^{(\ell)}\right| > (\operatorname{Lip}(\phi))^{\ell/2}\right) < \infty,$$

and so, by the Borel-Cantelli lemma,

$$\text{for all } \ell \text{ sufficiently large, } \left|\lambda_0^{(\ell+1)} - \lambda_0^{(\ell)}\right| \leq (\operatorname{Lip}(\phi))^{\ell/2} \text{ a.s.}$$

Using Cauchy's criterion, for each $n \in \mathbb{Z}$, $\left(\lambda_n^{(\ell)}\right)_{\ell\in\mathbb{N}}$ converges to $\lambda_n^{(\mathrm{erg})}$ as $\ell \to \infty$, $\mathbb{P}$-a.s. Clearly the process $\left(\lambda_n^{\mathrm{erg}}\right)_{n\in\mathbb{Z}}$ is (θ_n)-compatible, $n \in \mathbb{Z}$, and so ergodic. Similarly,

$$
\begin{aligned}
\sum_{\ell\in\mathbb{N}} \mathbb{P}\left(\left|X_0^{(\ell+1)} - X_0^{(\ell)}\right| > (\mathrm{Lip}(\phi))^{\ell/2}\right)
&\le \sum_{\ell\in\mathbb{N}} (\mathrm{Lip}(\phi))^{-\ell/2}\, \mathbb{E}\left|X_0^{(\ell+1)} - X_0^{(\ell)}\right| \\
&= \sum_{\ell\in\mathbb{N}} (\mathrm{Lip}(\phi))^{-\ell/2}\, \mathbb{E} N([-1,0), (\min\{\lambda_0^{(\ell-1)}, \lambda_0^{(\ell)}\}, \max\{\lambda_0^{(\ell-1)}, \lambda_0^{(\ell)}\}]) \\
&= \sum_{\ell\in\mathbb{N}} (\mathrm{Lip}(\phi))^{-\ell/2}\, \mathbb{E}\left|\lambda_0^{(\ell)} - \lambda_0^{(\ell-1)}\right| \\
&\le \phi(0)\, (\mathrm{Lip}(\phi))^{-1} \sum_{\ell\in\mathbb{N}} (\mathrm{Lip}(\phi))^{\ell/2} < \infty,
\end{aligned}
$$

and so, for each $n \in \mathbb{Z}$, $\left(X_n^{(\ell)}\right)_{\ell\in\mathbb{N}_0}$ converges to some limit $X_n^{(\mathrm{erg})}$ as $\ell \to \infty$, $\mathbb{P}$-a.s. Let $n \in \mathbb{Z}$ be arbitrarily fixed. Since $\left(Z_{n,i}\right)_{i\in\mathbb{N}}$ is independent of $\mathcal{F}_n^{X^{(\ell)}}$, for each $\ell \in \mathbb{N}_0$, then $\left(Z_{n,i}\right)_{i\in\mathbb{N}}$ is independent of $\mathcal{F}_n^{X^{(\mathrm{erg})}}$. Clearly, the sequence $\left(X_n^{(\mathrm{erg})}\right)_{n\in\mathbb{Z}}$ is (θ_n)-compatible, and therefore ergodic. To conclude the proof, we need to show that $\left(X_n^{(\mathrm{erg})}\right)_{n\in\mathbb{Z}}$ solves (1). By Fatou's lemma and (9), for each $k \in \mathbb{Z}$, we have

$$
\begin{aligned}
\mathbb{E}\left|X_n^{(\mathrm{erg})} - N([k-1,k), (0, \lambda_k^{(\mathrm{erg})}])\right| &\le \liminf_{\ell\to\infty} \mathbb{E}\Big|N([k-1,k), (0, \lambda_n^{(\ell)}]) \\
&\qquad - N([k-1,k), (0, \lambda_k^{(\mathrm{erg})}])\Big| \\
&= \liminf_{\ell\to\infty} \mathbb{E}\left|\lambda_n^{(\ell)} - \lambda_k^{(\mathrm{erg})}\right| = 0,
\end{aligned}
$$

and so $\left(X_n^{(\mathrm{erg})}\right)_{n\in\mathbb{Z}}$ is a modification of $\left(N([n-1,n), (0, \lambda_n^{(\mathrm{erg})}])\right)_{n\in\mathbb{Z}}$, i.e.,

$$
\mathbb{P}(X_n^{(\mathrm{erg})} = N([n-1,n), (0, \lambda_n^{(\mathrm{erg})}]), \quad \forall\, n \in \mathbb{Z}) = 1.
$$

The claim then follows if we show that $\left(\lambda_n^{(\mathrm{erg})}\right)_{n\in\mathbb{Z}}$ is a modification of $\left(\phi\left(\sum_{k=-\infty}^{n-1} \sum_{i=1}^{X_k^{(\mathrm{erg})}} \alpha_{n-k}(Z_{k,i})\right)\right)_{n\in\mathbb{Z}}$. Applying again Fatou's lemma, for each $n \in \mathbb{Z}$, we have

$$
\mathbb{E}\left|\lambda_n^{(\mathrm{erg})} - \phi\left(\sum_{k=-\infty}^{n-1} \sum_{i=1}^{X_k^{(\mathrm{erg})}} \alpha_{n-k}(Z_{k,i})\right)\right|
$$

$$
\begin{aligned}
&\leq \liminf_{\ell\to\infty} \mathbb{E}\left|\lambda_n^{(\ell)} - \phi\left(\sum_{k=-\infty}^{n-1}\sum_{i=1}^{X_k^{(\mathrm{erg})}} \alpha_{n-k}(Z_{k,i})\right)\right| \\
&\leq \liminf_{\ell\to\infty} \mathrm{Lip}(\phi)\mathbb{E}\left|\sum_{k=-\infty}^{n-1}\sum_{i=1}^{X_k^{(\ell)}} \alpha_{n-k}(Z_{k,i}) - \sum_{k=-\infty}^{n-1}\sum_{i=1}^{X_k^{(\mathrm{erg})}} \alpha_{n-k}(Z_{k,i})\right| \\
&\leq \mathrm{Lip}(\phi) \liminf_{\ell\to\infty} \mathbb{E}|X_0^{(\ell)} - X_0^{(\mathrm{erg})}| \\
&= \mathrm{Lip}(\phi) \liminf_{\ell\to\infty} \mathbb{E}|\lambda_0^{(\ell)} - \lambda_0^{(\mathrm{erg})}| = 0.
\end{aligned}
$$

The proof is completed. □

If the transfer function is bounded, the contractive condition (5) on the Lipschitz constant is no more necessary, indeed the following theorem holds.

Theorem 3 (Bounded Lipschitz transfer) *Suppose that ϕ is Lipschitz continuous and bounded and assume* (2). *Then there exists an ergodic version of the nonlinear marked Poisson autoregression.*

Proof Let N be a Poisson process on $\mathbb{R} \times \mathbb{R}_{\geq 0}$ with mean measure $\mathrm{d}t\,\mathrm{d}z$. We assume that N is defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which the set of mark random variables $Z = \left(Z_{n,i}\right)_{n\in\mathbb{Z}, i\in\mathbb{N}}$ is defined and that N is independent of Z. We assume that the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is provided with a flow $(\theta_n)_{n\in\mathbb{Z}}$ with respect to which $\mathbb{P}$ is ergodic and N and Z are (θ_n)-compatible (this assumption is not restrictive, see the proof of Theorem 2).

For $n \in \mathbb{Z}$ and $\ell \in \mathbb{N}$, we define recursively $X_n^{(\ell)}$ and $\lambda_n^{(\ell)}$ as in (6) and (7). As checked in the proof of Theorem 2, for each $\ell \in \mathbb{N}$, the processes $\left(\lambda_n^{(\ell)}\right)_{n\in\mathbb{Z}}$ and $\left(X_n^{(\ell)}\right)_{n\in\mathbb{Z}}$ are (θ_n)-compatible, $n \in \mathbb{Z}$, and therefore ergodic. As checked in the proof of Theorem 2, for arbitrarily fixed $n \in \mathbb{Z}$, $\left(Z_{n,i}\right)_{i\in\mathbb{N}}$ is independent of $\mathcal{F}_n^{X^{(\ell)}}$ for each $\ell \in \mathbb{N}$.

For each $n \in \mathbb{Z}$, the $\mathbb{P}$-a.s. convergence of $X_n^{(\ell)}$ as $\ell \to \infty$ holds if we prove that

$$
\begin{aligned}
D_n &:= \limsup_{\ell\to\infty} X_n^{(\ell)} - \liminf_{\ell\to\infty} X_n^{(\ell)} \\
&= N([n-1, n), (\liminf_{\ell\to\infty} \lambda_n^{(\ell)}, \limsup_{\ell\to\infty} \lambda_n^{(\ell)}]) = 0, \quad \mathbb{P}\text{-a.s.}
\end{aligned}
\tag{10}
$$

We shall show later on that, for any integer $K \geq 1$,

$$
\begin{aligned}
&\mathbb{P}\left(\sum_{n=1}^{K} D_n = 0 \,\middle|\, \mathcal{F}_0^D \vee \mathcal{F}_0^Z \vee \sigma\left\{\liminf_{\ell\to\infty} \lambda_1^{(\ell)}, \limsup_{\ell\to\infty} \lambda_1^{(\ell)}\right\}\right) \\
&\quad \geq \exp\left\{-\mathrm{Lip}(\phi)\sum_{n=1}^{K}\sum_{k=-\infty}^{0}\sum_{i=1}^{D_k} |\alpha_{n-k}(Z_{k,i})|\right\}.
\end{aligned}
\tag{11}
$$

Taking the expectation, then using Jensen's inequality, Wald's identity (as after relation (4)), and finally the relation $\mathbb{E}[D_i] \le \mathbb{E}[N([i-1,i),(0,\|\phi\|_\infty])] = \|\phi\|_\infty$, we obtain

$$\mathbb{P}\left(\sum_{n=1}^{K} D_n = 0\right) \ge \exp\left\{-\mathrm{Lip}(\phi)\|\phi\|_\infty \sum_{n=1}^{K}\sum_{k=n}^{\infty} \overline{\alpha}_k\right\}, \quad K \in \mathbb{N}.$$

Letting K tend to infinity, we conclude that

$$\mathbb{P}\left(\sum_{n=1}^{\infty} D_n = 0\right) \ge \exp\left\{-\mathrm{Lip}(\phi)\|\phi\|_\infty \sum_{k=1}^{\infty} k\overline{\alpha}_k\right\} > 0,$$

where the latter inequality is a consequence of assumption (2). Thus, $\mathbb{P}(D_n = 0) > 0$ for each $n \in \mathbb{N}$ and by ergodicity of the sequence $(D_n)_{n\in\mathbb{Z}}$, we have (10), and so $(X_n^{(\ell)})$ converges $\mathbb{P}$-a.s., as $\ell \to \infty$.

Now we prove (11). We proceed by induction over $K \ge 1$. For $K = 1$ we have

$$\begin{aligned}&\mathbb{P}\left(D_1 = 0 \,\Big|\, \mathcal{F}_0^D \vee \mathcal{F}_0^Z \vee \sigma\left\{\liminf_{\ell\to\infty}\lambda_1^{(\ell)}, \limsup_{\ell\to\infty}\lambda_1^{(\ell)}\right\}\right) = \mathrm{e}^{-(\limsup_{\ell\to\infty}\lambda_1^{(\ell)} - \liminf_{\ell\to\infty}\lambda_1^{(\ell)})}\\ &\ge \exp\left\{-\mathrm{Lip}(\phi)\sum_{k=-\infty}^{0}\sum_{i=1}^{D_k}\left|\alpha_{1-k}(Z_{k,i})\right|\right\},\end{aligned}$$

where the latter inequality follows from the following computation:

$$\begin{aligned}\limsup_{\ell\to\infty}\lambda_n^{(\ell)} - \liminf_{\ell\to\infty}\lambda_n^{(\ell)} &= \lim_{\ell\to\infty}\sup_{p,q\ge\ell}(\lambda_n^{(p)} - \lambda_n^{(q)})\\ &\le \mathrm{Lip}(\phi)\lim_{\ell\to\infty}\sup_{p,q\ge\ell}\left|\sum_{k=-\infty}^{n-1}\sum_{i=1}^{X_k^{(p)}}\alpha_{k-1}(Z_{k,i}) - \sum_{k=-\infty}^{n-1}\sum_{i=1}^{X_k^{(q)}}\alpha_{k-1}(Z_{k,i})\right|\\ &\le \mathrm{Lip}(\phi)\lim_{\ell\to\infty}\sup_{p,q\ge\ell}\sum_{k=-\infty}^{n-1}\sum_{i=1}^{|X_k^{(p)}-X_k^{(q)}|}|\alpha_{n-k}(Z_{k,i})|\\ &\le \mathrm{Lip}(\phi)\lim_{\ell\to\infty}\sum_{k=-\infty}^{n-1}\sum_{i=1}^{\sup_{p\ge\ell}X_k^{(p)}-\inf_{q\ge\ell}X_k^{(q)}}|\alpha_{n-k}(Z_{k,i})|\\ &= \mathrm{Lip}(\phi)\sum_{k=-\infty}^{n-1}\sum_{i=1}^{\limsup_{\ell\to\infty}X_k^{(\ell)}-\liminf_{\ell\to\infty}X_k^{(\ell)}}|\alpha_{n-k}(Z_{k,i})| \qquad (12)\\ &= \mathrm{Lip}(\phi)\sum_{k=-\infty}^{n-1}\sum_{i=1}^{D_k}|\alpha_{n-k}(Z_{k,i})|, \qquad (13)\end{aligned}$$

where equality (12) follows by a simple truncation argument. Assume the claim holds for $K-1$. By (13), we have

$$
\begin{aligned}
&\mathbb{P}\left(\sum_{n=1}^{K} D_n = 0 \,\middle|\, \mathcal{F}^D_{K-1} \vee \mathcal{F}^Z_0 \vee \bigvee_{n=1}^{K} \sigma\left\{\liminf_{\ell\to\infty} \lambda_n^{(\ell)}, \limsup_{\ell\to\infty} \lambda_n^{(\ell)}\right\}\right)\\
&=\mathbf{1}\left\{\sum_{n=1}^{K-1} D_n=0\right\} \mathbb{P}\left(D_K=0 \,\middle|\, \mathcal{F}^D_{K-1} \vee \mathcal{F}^Z_0 \vee \bigvee_{n=1}^{K} \sigma\left\{\liminf_{\ell\to\infty} \lambda_n^{(\ell)}, \limsup_{\ell\to\infty} \lambda_n^{(\ell)}\right\}\right)\\
&= \mathbf{1}\left\{\sum_{n=1}^{K-1} D_n = 0\right\} \exp\left\{-(\limsup_{\ell\to\infty} \lambda_K^{(\ell)} - \liminf_{\ell\to\infty} \lambda_K^{(\ell)})\right\}\\
&\geq \mathbf{1}\left\{\sum_{n=1}^{K-1} D_n = 0\right\} \exp\left\{-\mathrm{Lip}(\phi) \sum_{k=-\infty}^{K-1} \sum_{i=1}^{D_k} |\alpha_{K-k}(Z_{k,i})|\right\}\\
&= \mathbf{1}\left\{\sum_{n=1}^{K-1} D_n = 0\right\} \exp\left\{-\mathrm{Lip}(\phi) \sum_{k=-\infty}^{0} \sum_{i=1}^{D_k} |\alpha_{K-k}(Z_{k,i})|\right\}.
\end{aligned}
$$

Taking the conditional expectation with respect to $\mathcal{F}^D_0 \vee \mathcal{F}^Z_0 \vee \sigma\left\{\liminf_{\ell\to\infty} \lambda_1^{(\ell)}, \right.$ $\left.\limsup_{\ell\to\infty} \lambda_1^{(\ell)}\right\}$ and applying the induction hypothesis, we obtain

$$
\begin{aligned}
&\mathbb{P}\left(\sum_{n=1}^{K} D_n = 0 \,\middle|\, \mathcal{F}^D_0 \vee \mathcal{F}^Z_0 \vee \sigma\left\{\liminf_{\ell\to\infty} \lambda_1^{(\ell)}, \limsup_{\ell\to\infty} \lambda_1^{(\ell)}\right\}\right)\\
&\geq \exp\left\{-\mathrm{Lip}(\phi) \sum_{k=-\infty}^{0} \sum_{i=1}^{D_k} |\alpha_{K-k}(Z_{k,i})|\right\}\\
&\exp\left\{-\mathrm{Lip}(\phi) \sum_{n=1}^{K-1} \sum_{k=-\infty}^{0} \sum_{i=1}^{D_k} |\alpha_{n-k}(Z_{k,i})|\right\}\\
&= \exp\left\{-\mathrm{Lip}(\phi) \sum_{n=1}^{K} \sum_{i=-\infty}^{0} \sum_{i=1}^{D_i} |\alpha_{n-k}(Z_{k,i})|\right\}.
\end{aligned}
$$

The proof of (11) is completed.

Finally, we show that the limiting process solves (1). Let $\left(X_n^{(\mathrm{erg})}\right)_{n\in\mathbb{Z}}$ and $\left(\lambda_n^{(\mathrm{erg})}\right)_{n\in\mathbb{Z}}$ be defined by

$$
X_n^{(\mathrm{erg})} := \lim_{\ell\to\infty} X_n^{(\ell)} \quad \text{and} \quad \lambda_n^{(\mathrm{erg})} := \lim_{\ell\to\infty} \lambda_n^{(\ell)}, \quad \mathbb{P}\text{-a.s.}
$$

(note that the almost sure convergence of the sequence $\left(\lambda_n^{(\ell)}\right)_\ell$ is guaranteed by (10) and (13)). Obviously, for $n \in \mathbb{N}$,

$$\lambda_n^{(\text{erg})} = \phi\left(\sum_{k=-\infty}^{n-1} \sum_{i=1}^{X_k^{(\text{erg})}} \alpha_{n-k}(Z_{k,i})\right), \quad \mathbb{P}\text{-a.s.}$$

Clearly these processes are (θ_n)-compatible and consequently ergodic. We have already noticed that, for each $k \in \mathbb{Z}$, $\left(Z_{k,i}\right)_{i\in\mathbb{N}}$ is independent of $\mathcal{F}_k^{X^{(\ell)}}$, for each $\ell \in \mathbb{N}$, and therefore $\left(Z_{k,i}\right)_{i\in\mathbb{N}}$ is independent of $\mathcal{F}_k^{X^{(\text{erg})}}$.

It remains to prove that, for $n \in \mathbb{Z}$, $X_n^{(\text{erg})}$ is a modification of $N([n-1,n),(0,\lambda_n^{(\text{erg})}])$. By Fatou's lemma, we have

$$\begin{aligned}
&\mathbb{E}\left|X_n^{(\text{erg})} - N([n-1,n),(0,\lambda_n^{(\text{erg})}])\right| \\
&\quad\le \liminf_{\ell\to\infty} \mathbb{E}\left|X_n^{(\ell)} - N([n-1,n),(0,\lambda_n^{(\text{erg})}])\right| \\
&\quad= \liminf_{\ell\to\infty} \mathbb{E}\left|N([n-1,n),(0,\lambda_n^{(\ell-1)}]) - N([n-1,n),(0,\lambda_n^{(\text{erg})}])\right| \\
&\quad= \liminf_{\ell\to\infty} \mathbb{E}\left|\lambda_n^{(\ell-1)} - \lambda_n^{(\text{erg})}\right|.
\end{aligned}$$

This latter term is equal to zero by the dominated convergence theorem (indeed the random variables $\lambda_n^{(\ell)}$, $\lambda_n^{(\text{erg})}$ are bounded by $\|\phi\|_\infty$ and $\lambda_n^{(\ell)} \to \lambda_n^{(\text{erg})}$ a.s., as $\ell \to \infty$). Therefore, for each $n \in \mathbb{Z}$, $\mathbb{E}|X_n^{(\text{erg})} - N([n-1,n),(0,\lambda_n^{(\text{erg})}])| = 0$, and so, for each $n \in \mathbb{Z}$, $X_n^{(\text{erg})} = N([n-1,n),(0,\lambda_n^{(\text{erg})}])$ almost surely, which concludes the proof. □

Finally, we present an existence result without continuity assumption.

Theorem 4 (Transfer without continuity assumption) *Suppose that ϕ is nondecreasing, lower semi-continuous, and such that $\phi(x) \le \beta + \gamma x$, $x \in \mathbb{R}_{\ge 0}$, for some $\beta > 0$ and $\gamma \in [0,1)$. Moreover, suppose that $\alpha_k : \mathrm{Z} \to \mathbb{R}_{\ge 0}$, $k \in \mathbb{N}$. Then there exists an ergodic version of the nonlinear marked Poisson autoregression with finite mean.*

Proof Let N be a Poisson process on $\mathbb{R}\times\mathbb{R}_{\ge 0}$ with mean measure equal to the Lebesgue measure. We assume that N is defined on the same probability space $(\Omega,\mathcal{F},\mathbb{P})$ on which the mark random variables $Z = \left(Z_{k,i}\right)_{k\in\mathbb{Z},i\in\mathbb{N}}$ are defined and that N is independent of Z. We assume that the probability space $(\Omega,\mathcal{F},\mathbb{P})$ is provided with a flow $(\theta_n)_{n\in\mathbb{Z}}$ with respect to which $\mathbb{P}$ is ergodic and N and Z are (θ_n)-compatible (this assumption is not restrictive, see the proof of Theorem 2).

For $n \in \mathbb{Z}$ and $\ell \in \mathbb{N}$, we define $X_n^{(\ell)}$ and $\lambda_n^{(\ell)}$ as in (6) and (7). As checked in the proof of Theorem 2, for each $\ell \in \mathbb{N}$, the processes $\left(\lambda_n^{(\ell)}\right)_{n\in\mathbb{Z}}$ and $\left(X_n^{(\ell)}\right)_{n\in\mathbb{Z}}$ are (θ_n)-compatible, $n \in \mathbb{Z}$, and therefore ergodic. As checked in the proof of Theorem 2, for each $i \in \mathbb{Z}$, the mark sequence $\left(Z_{k,i}\right)_{i\in\mathbb{N}}$ is independent of $\mathcal{F}_k^{X^{(\ell)}}$, $\ell \in \mathbb{N}$.

Since the functions $\alpha_k(\cdot)$ are nonnegative and ϕ is nondecreasing, we have that, for each $n \in \mathbb{Z}$, the mappings $\ell \mapsto \lambda_n^{(\ell)}$ and $\ell \mapsto X_n^{(\ell)}$ are nondecreasing. Therefore there exist the limiting processes $\left(\lambda_n^{(\mathrm{erg})}\right)_{n\in\mathbb{Z}}$ and $\left(X_n^{(\mathrm{erg})}\right)_{n\in\mathbb{Z}}$, which are (θ_n)-compatible and therefore ergodic. Note that, in particular,

$$X_n^{(\mathrm{erg})} = N([n-1, n), (0, \lambda_n^{(\mathrm{erg})}]),$$

and so the claim follows if we prove that $\mathbb{E}\lambda_0^{(\mathrm{erg})} < \infty$ and

$$\lambda_n^{(\mathrm{erg})} = \phi\left(\sum_{k=-\infty}^{n-1} \sum_{i=1}^{X_k^{(\mathrm{erg})}} \alpha_{n-k}(Z_{k,i})\right). \tag{14}$$

Stationarity of the processes $\left(\lambda_n^{(\ell)}\right)_{n\in\mathbb{Z}}$ and $\left(X_n^{(\ell)}\right)_{n\in\mathbb{Z}}$, $\ell \in \mathbb{N}$, sublinearity of ϕ, and monotonicity of $\left(\lambda_0^{(\ell)}\right)_{\ell\in\mathbb{N}}$ yield

$$\mathbb{E}\lambda_0^{(\ell)} \le \beta + \gamma\mathbb{E}X_0^{(\ell)} = \beta + \gamma\mathbb{E}\lambda_0^{(\ell-1)} \le \beta + \gamma\mathbb{E}\lambda_0^{(\ell)}, \quad \ell \in \mathbb{N}$$

and thus $\mathbb{E}\lambda_0^{(\ell)} \le \beta/(1-\gamma)$. Passing to the limit as $\ell \to \infty$ and using the monotone convergence theorem we obtain

$$\mathbb{E}\lambda_0^{(\mathrm{erg})} \le \frac{\beta}{1-\gamma} < \infty.$$

For each $n \in \mathbb{Z}$ and $\ell \in \mathbb{N}$, by monotonicity we have

$$\lambda_n^{(\ell)} = \phi\left(\sum_{k=-\infty}^{n-1} \sum_{i=1}^{X_k^{(\ell)}} \alpha_{n-k}(Z_{k,i})\right) \le \phi\left(\sum_{k=-\infty}^{n-1} \sum_{i=1}^{X_k^{(\mathrm{erg})}} \alpha_{n-k}(Z_{k,i})\right).$$

On the other hand

$$\lambda_n^{(\mathrm{erg})} \ge \lambda_n^{(\ell)} = \phi\left(\sum_{k=-\infty}^{n-1} \sum_{i=1}^{X_k^{(\ell)}} \alpha_{n-k}(Z_{k,i})\right).$$

The claim (14) follows taking the limit as $\ell \to \infty$ on this inequalities. Indeed, by the first inequality we immediately have

$$\lambda_n^{(\mathrm{erg})} \le \phi\left(\sum_{k=-\infty}^{n-1} \sum_{i=1}^{X_k^{(\mathrm{erg})}} \alpha_{n-k}(Z_{k,i})\right),$$

and by the second inequality we have

$$
\begin{aligned}
\lambda_n^{(\mathrm{erg})} &\geq \liminf_{\ell\to\infty} \phi\left(\sum_{k=-\infty}^{n-1}\sum_{i=1}^{X_k^{(\ell)}} \alpha_{n-k}(Z_{k,i})\right) \\
&\geq \phi\left(\liminf_{\ell\to\infty} \sum_{k=-\infty}^{n-1}\sum_{i=1}^{X_k^{(\ell)}} \alpha_{n-k}(Z_{k,i})\right) \\
&= \phi\left(\lim_{\ell\to\infty}\inf_{q\geq\ell} \sum_{k=-\infty}^{n-1}\sum_{i=1}^{X_k^{(q)}} \alpha_{n-k}(Z_{k,i})\right) \\
&\geq \phi\left(\lim_{\ell\to\infty} \sum_{k=-\infty}^{n-1}\sum_{i=1}^{\inf_{q\geq\ell} X_k^{(q)}} \alpha_{n-k}(Z_{k,i})\right) \\
&= \phi\left(\lim_{\ell\to\infty} \sum_{k=-\infty}^{n-1}\sum_{i=1}^{X_k^{(\ell)}} \alpha_{n-k}(Z_{k,i})\right) \qquad (15)\\
&= \phi\left(\sum_{k=-\infty}^{n-1}\sum_{i=1}^{X_k^{(\mathrm{erg})}} \alpha_{n-k}(Z_{k,i})\right),
\end{aligned}
$$

where the second inequality follows by the lower semi-continuity of ϕ and in (15) we exchanged the limit with the infinite sum by the monotone convegence theorem. □

3 Convergence to Equilibrium and Uniqueness

We say that $\left(X_n'\right)_{n\in\mathbb{Z}}$ is a *nonlinear marked Poisson autoregression with initial condition $\mathcal{C}$ on $\mathbb{Z}_{\leq 0}$* if $\left(X_n'\right)_{k\leq 0}$ satisfies $\mathcal{C}$, and

$$
X_n' \mid \mathcal{F}_{n-1}^{X'} \vee \mathcal{F}_{n-1}^{Z} \sim \mathrm{Pois}(\lambda_n'), \quad n \geq 1,
$$

with

$$
\lambda_n' := \phi\left(\sum_{k=-\infty}^{n-1}\sum_{i=1}^{X_k'} \alpha_{n-k}(Z_{k,i})\right) = \phi\left(\sum_{k=-\infty}^{0}\sum_{i=1}^{X_k'} \alpha_{n-k}(Z_{k,i}) + \sum_{k=1}^{n-1}\sum_{i=1}^{X_k'} \alpha_{n-k}(Z_{k,i})\right). \qquad (16)
$$

Here $\left(Z_{k,i}\right)_{k\in\mathbb{Z},i\in\mathbb{N}}$ is a set of identically distributed mark random variables with values in a measurable space $(\mathrm{Z}, \mathcal{Z})$ such that the mark sequences $\left(Z_{k,i}\right)_{i\in\mathbb{N}}$ are independent over k, and, therefore, $\left(Z_{k,i}\right)_{i\in\mathbb{N}}$ is independent of $\mathcal{F}_k^{X'}$, $k \in \mathbb{N}$. Furthermore, α_k : $\mathrm{Z} \to \mathbb{R}$, $k \in \mathbb{N}$, and $\phi : \mathbb{R} \to \mathbb{R}_{\geq 0}$ are measurable functions, which, as before, we call coefficient and transfer functions, respectively. In this section, we investigate the convergence to equilibrium of such a non-stationary model under various assumptions on ϕ, α_k, $k \in \mathbb{N}$, and the distribution of the marks. Below, we denote by $(X_n)_{n\in\mathbb{Z}}$ (and $(\lambda_n)_{n\in\mathbb{Z}}$) an ergodic version of the nonlinear marked Poisson autoregression with finite mean, as constructed in the previous section. We say that the ergodic law of $(X_n)_{n\in\mathbb{Z}}$ is unique if any other ergodic version of the nonlinear marked Poisson autoregression with finite mean has the same law as $(X_n)_{n\in\mathbb{Z}}$. We refer the reader to [5, 10, 18] for the formal definitions of the convergence concepts for discrete-time and real-valued stochastic processes used throughout the paper. Our use of coupling and convolutions of sequences and discrete distributions is specified by the following definitions.

Definition 1 [Convolution] Given two nonnegative sequences $a = (a_n)_{n\in\mathbb{Z}}$ and $b = (b_n)_{n\in\mathbb{Z}}$, we define their *convolution* as

$$(a * b)_n := \sum_{k\in\mathbb{Z}} a_{n-k} b_k, \quad k \in \mathbb{Z}.$$

We denote by a^{*m} the m-fold convolution of a with itself and we set $a_k^{*0} = \mathbf{1}_{(0)}(k)$. For A a distribution function supported on $\mathbb{N}$ with $A(n) := \sum_{k=1}^n a_n$ for some nonnegative sequence (a_n) such that $\sum_{n\geq 1} a_n = 1$, we set

$$A^{*0} \equiv 1 \quad \text{and} \quad A^{*i}(n) := \sum_{k=1}^{n} A^{*(i-1)}(k) a_{n-k}, \quad 1 \leq i \leq n.$$

Definition 2 (Coupling [18]) Let $(X_n)_{n\in\mathbb{Z}}$ and $(Y_n)_{n\in\mathbb{Z}}$ be two stochastic processes defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and with values on $\mathbb{N}$. They are said to *couple* if there exists a $\mathbb{P}$-a.s. finite random variable T such that

$$\mathbb{P}(X_n = Y_n \quad \forall\, n \geq T) = 1.$$

T is called coupling time.

First, we consider nonlinear marked Poisson autoregressions with contractive Lipschitz transfer functions.

Theorem 5 (Convergence to equilibrium and uniqueness with Lipschitz transfer) *Suppose that the transfer function ϕ is Lipschitz continuous with* $\mathrm{Lip}(\phi) < 1$ *and let $(X_n)_{n\in\mathbb{Z}}$ be an ergodic version of the nonlinear marked Poisson autoregression with finite mean, as constructed in Theorem 2. Then:*

(i) If $\left(X_n^{'}\right)_{n\in\mathbb{Z}}$ is a nonlinear marked Poisson autoregression with initial condition

$$\mathcal{C}_1 : \quad \lim_{n\to\infty} \sum_{k=-\infty}^{0} \overline{\alpha}_{n-k} \mathbb{E}|X_k^{'}| = 0,$$

then $\left(X'_{K+n}\right)_{n\in\mathbb{Z}}$ *converges in total variation (and therefore in law) to* $(X_n)_{n\in\mathbb{Z}}$ *as* $K \to \infty$.

(ii) The ergodic law of $(X_n)_{n\in\mathbb{Z}}$ *is unique.*

(iii) If $\left(X'_n\right)_{n\in\mathbb{Z}}$ *is a nonlinear marked Poisson autoregression with initial condition*

$$\widetilde{\mathcal{C}}_1 : \quad \sum_{n=1}^{\infty}\sum_{k=-\infty}^{0} \overline{\alpha}_{n-k}\mathbb{E}|X'_k| < \infty$$

and (2) *holds, then the processes* $\left(X'_n\right)_{n\in\mathbb{Z}}$ *and* $(X_n)_{n\in\mathbb{Z}}$ *can be constructed in such a way that they couple.*

Proof (i). We start constructing a nonlinear marked Poisson autoregression $\left(X'_n\right)_{n\in\mathbb{Z}}$ with initial condition $\mathcal{C}_1$ on $\mathbb{Z}_{\le 0}$. Let $\left(X'_n\right)_{k\le 0}$ satisfy $\mathcal{C}_1$ on $\mathbb{Z}_{\le 0}$. Let N be a Poisson process on $\mathbb{R}\times\mathbb{R}_{\ge 0}$ with mean measure $\mathrm{d}t\mathrm{d}z$ independent of $\mathcal{F}_0^{X'}$. For $n \in \mathbb{N}$, we set

$$\begin{aligned}
X'_n &:= N([n-1,n),(0,\lambda'_n]), \quad \text{and} \\
\lambda'_n &:= \phi\left(\sum_{k=-\infty}^{n-1}\sum_{i=1}^{X'_k}\alpha_{n-k}(Z_{k,i})\right) \\
&= \phi\left(\sum_{k=-\infty}^{0}\sum_{i=1}^{X'_k}\alpha_{n-k}(Z_{k,i}) + \sum_{k=1}^{n-1}\sum_{i=1}^{X'_k}\alpha_{n-k}(Z_{k,i})\right),
\end{aligned}$$

where we use the convention $\sum_{i=1}^{0}\dots := 0$. Here, $\left(Z_{k,i}\right)_{k\in\mathbb{Z},i\in\mathbb{N}}$ are identically distributed Z-valued mark random variables, independent of N, such that the mark sequences $\left(Z_{k,i}\right)_{i\in\mathbb{N}}$ are independent over $k\in\mathbb{Z}$, and, thus, for each $k\in\mathbb{N}$, the mark sequence $\left(Z_{k,i}\right)_{i\in\mathbb{N}}$ is independent of $\mathcal{F}_k^{X'}$. Obviously, $\left(X'_n\right)_{n\in\mathbb{Z}}$ follows the dynamics (16) on $\mathbb{N}$. Note that, for $n\in\mathbb{N}$, $\lambda'_n < \infty$ a.s. Indeed, by the initial condition $\mathcal{C}_1$ we have that

$$\forall\, n\in\mathbb{N}, \sum_{k=-\infty}^{0}\sum_{i=1}^{X'_k}\left|\alpha_{n-k}(Z_{k,i})\right| < \infty, \quad \mathbb{P}\text{-a.s.}$$

Let $(X_n)_{n\in\mathbb{Z}}$ be an ergodic version of the marked Poisson autoregression with finite mean λ, as constructed in Theorem 2. We construct such a version using the same N and the same marks $(Z_{k,i})$ used to construct $(X'_n)_{n\in\mathbb{Z}}$. By Theorem 2 $(X_n)_{n\in\mathbb{Z}}$ is a modification of $(N([n-1,n),(0,\lambda_n]))_{n\in\mathbb{Z}}$, where $(\lambda_n)_{n\in\mathbb{Z}}$ is a modification of $\left(\phi\left(\sum_{k=-\infty}^{n-1}\sum_{i=1}^{X_k}\alpha_{n-k}(Z_{k,i})\right)\right)_{n\in\mathbb{Z}}$.

Set

$$a_k := \mathbb{E}\left[\left|\lambda_k - \lambda_k^{'}\right|\right] \mathbf{1}_{(k\in\mathbb{N})}, \quad k \in \mathbb{Z}.$$

We aim to show that $\lim_{k\to\infty} a_k = 0$. For $n \in \mathbb{N}$, we have

$$\begin{aligned}
a_n &\leq \mathrm{Lip}(\phi)\mathbb{E}\left|\sum_{k=-\infty}^{n-1}\sum_{i=1}^{X_k}\alpha_{n-k}(Z_{k,i}) - \sum_{k=-\infty}^{n-1}\sum_{i=1}^{X_k^{'}}\alpha_{n-k}(Z_{k,i})\right| \\
&\leq \mathrm{Lip}(\phi)\sum_{k=-\infty}^{n-1}\mathbb{E}\left|\sum_{i=1}^{X_k}\alpha_{n-k}(Z_{k,i}) - \sum_{i=1}^{X_k^{'}}\alpha_{n-k}(Z_{k,i})\right| \\
&\leq \mathrm{Lip}(\phi)\sum_{k=-\infty}^{n-1}\mathbb{E}\left[|\alpha_{n-k}(Z_{0,1})|\right]\mathbb{E}\left[|X_k - X_k^{'}|\right] \\
&\leq \mathrm{Lip}(\phi)\sum_{k=-\infty}^{0}\overline{\alpha}_{n-k}(\lambda + \mathbb{E}X_k^{\prime}) + \mathrm{Lip}(\phi)\sum_{k=1}^{n-1}\overline{\alpha}_{n-k}a_k,
\end{aligned}$$

where we have used Wald's identity as after relation (4). Setting

$$b_n := \mathrm{Lip}(\phi)\overline{\alpha}_n \quad \text{and} \quad c_n := \mathrm{Lip}(\phi)\sum_{k=-\infty}^{0}\overline{\alpha}_{n-k}(\lambda + \mathbb{E}X_k^{\prime}), \quad n \in \mathbb{N},$$

we have

$$a_n \leq c_n + \sum_{k=1}^{n-1} a_k b_{n-k}, \quad n \in \mathbb{N} \tag{17}$$

$$\lim_{n\to\infty} c_n = 0, \tag{18}$$

$$B := \sum_{n=1}^{\infty} b_n < 1. \tag{19}$$

By (18) we have $\sup_{n\geq 1} c_n < C < \infty$ for some constant $C \geq 0$. Consequently, also the sequence (a_n) is bounded. Indeed, together with (17) we find

$$a_n \leq C + \sum_{k=1}^{n-1} b_{n-k} \max_{1\leq \ell \leq \leq n-1} a_\ell = C + \sum_{k=1}^{n-1} b_k \max_{1\leq \ell\leq n-1} a_\ell \leq C + B \max_{1\leq\ell\leq n-1} a_\ell.$$

Taking the maximum over $n \in (1, 2, \dots, K)$ on both sides and iterating the inequality, we obtain for all $K \in \mathbb{N}$

$$\max_{1\leq n\leq K} a_n \leq C + B \max_{1\leq n\leq K-1} a_n \leq C + B(C + B \max_{1\leq n\leq K-2} a_n)$$

$$\leq C(1 + B + B^2 + \dots) = \frac{C}{1 - B} =: A,$$

and thus $\sup_{n \geq 1} a_n \leq A < \infty$ by (19).

Next, we iterate the inequality (17) to obtain

$$a_n \leq c_n + \sum_{j_1=1}^{n-1} b_{n-j_1} c_{j_1} + \sum_{j_1=1}^{n-1} b_{n-j_1} \sum_{j_2=1}^{j_1-1} b_{j_1-j_2} a_{j_2} \leq s_n^{(1)} + s_n^{(2)},$$

with

$$s_n^{(1)} := c_n + \sum_{j_1=1}^{n-1} b_{n-j_1} c_{j_1} + \dots + \sum_{j_1=1}^{n-1} b_{n-j_1} \sum_{j_2=1}^{j_1-1} b_{j_1-j_2} \dots \sum_{j_n=1}^{j_{n-1}-1} b_{j_{n-1}-j_n} c_{j_n}$$

and

$$s_n^{(2)} := \sum_{j_1=1}^{n-1} b_{n-j_1} \sum_{j_2=1}^{j_1-1} b_{j_1-j_2} \dots \sum_{j_{n+1}=1}^{j_n-1} b_{j_n-j_{n+1}} a_{j_{n+1}}.$$

We have

$$s_n^{(2)} \leq \left(\sum_{i=1}^{\infty} b_i \right)^{n+1} \max_{1 \leq j \leq n-1} a_j \leq B^{n+1} A$$

and consequently $s_n^{(2)} \to 0$, as $n \to \infty$. We rewrite $s_n^{(1)}$ as

$$s_n^{(1)} = \sum_{j=0}^{n} (b^{*j} * c)_n,$$

where the symbol $*$ denotes the convolution between sequences; here we set $b_j, c_j :=$ 0 for $j \leq 0$. We have

$$s_n^{(1)} \leq C \sum_{j=0}^{\infty} (b^{*j})_n.$$

We aim to show that $\sum_{n=1}^{\infty} \sum_{j=0}^{\infty} (b^{*j})_n < \infty$, which implies $\lim_{n \to \infty} s_n^{(1)} = 0$ and then $\lim_{n \to \infty} a_n = 0$. Reasoning by induction over $j \geq 0$, we obtain that

$$\sum_{n \geq 1} (b^{*j})_n \leq B^j, \quad \text{for each } j \geq 0, n \in \mathbb{N}.$$

Therefore $\sum_{n=1}^{\infty} \sum_{j=0}^{\infty} (b^{*j})_n \leq \sum_{j \geq 0} B^j = (1 - B)^{-1} < \infty$. Let $K, \kappa_1, \kappa_2 \in \mathbb{Z}$ with $\kappa_1 \leq \kappa_2$. We have

$$\mathbb{P}\Big((X'_{K+\kappa_1}, X'_{K+\kappa_1+1}, \dots, X'_{K+\kappa_2}) \neq (X_{K+\kappa_1}, X_{K+\kappa_1+1}, \dots, X_{K+\kappa_2}) \Big)$$

$$
\begin{aligned}
&= \mathbb{P}\left(\bigcup_{n=K+\kappa_1}^{K+\kappa_2} \left\{ |X_n - X_n^{'}| > 0 \right\} \right) \\
&\leq \sum_{n=K+\kappa_1}^{K+\kappa_2} \mathbb{P}\left(|X_n - X_n^{'}| > 0 \right) \\
&\leq \sum_{n=K+\kappa_1}^{K+\kappa_2} \mathbb{E}\left[|X_n - X_n^{'}| \right] \\
&= \sum_{n=K+\kappa_1}^{K+\kappa_2} a_n \leq (\kappa_2 - \kappa_1 + 1) \max_{k \in (K+\kappa_1, K+\kappa_1+1, \ldots, K+\kappa_2)} a_k .
\end{aligned}
$$

Since $\lim_{K\to\infty} a_{K+\kappa} = 0$ for each $\kappa \in \mathbb{Z}$, we have that,

$$
\mathbb{P}\left((X^{'}_{K+\kappa_1}, X^{'}_{K+\kappa_1+1}, \ldots, X^{'}_{K+\kappa_2}) \neq (X_{K+\kappa_1}, X_{K+\kappa_1+1}, \ldots, X_{K+\kappa_2}) \right) \to 0, \quad \text{as } K \to \infty \tag{20}
$$

for each $\kappa_1, \kappa_2 \in \mathbb{Z}$.

In particular, letting d_{TV} denote the total variation distance (see e.g. [10, 18]), for $n, x_1, \ldots, x_n \in \mathbb{N}_0$, and $\kappa_1, \ldots, \kappa_n \in \mathbb{Z}$, $\kappa_1 < \kappa_2 < \ldots < \kappa_n$, we have

$$
\begin{aligned}
& d_{\mathrm{TV}}\left((X^{'}_{K+\kappa_1}, X^{'}_{K+\kappa_2}, \ldots, X^{'}_{K+\kappa_n}), (X_{\kappa_1}, X_{\kappa_2}, \ldots, X_{\kappa_n}) \right) \\
&= \frac{1}{2} \sum_{x_1,\ldots,x_n \in \mathbb{N}_0} \left| \mathbb{P}(X^{'}_{K+\kappa_1} = x_1, \ldots, X^{'}_{K+\kappa_n} = x_n) - \mathbb{P}(X_{\kappa_1} = x_1, \ldots, X_{\kappa_n} = x_n) \right| \\
&= \frac{1}{2} \sum_{x_1,\ldots,x_n \in \mathbb{N}_0} \left| \mathbb{P}(X^{'}_{K+\kappa_1} = x_1, \ldots, X^{'}_{K+\kappa_n} = x_n) - \mathbb{P}(X_{K+\kappa_1} = x_1, \ldots, X_{K+\kappa_n} = x_n) \right| \\
&\leq \frac{1}{2} \mathbb{P}\left((X^{'}_{K+\kappa_1}, \ldots, X^{'}_{K+\kappa_n}) \neq (X_{K+\kappa_1}, \ldots, X_{K+\kappa_n}) \right) \qquad (21) \\
&\leq \frac{1}{2} \mathbb{P}\left((X^{'}_{K+\kappa_1}, X^{'}_{K+\kappa_1+1}, \ldots, X^{'}_{K+\kappa_n}) \neq (X_{K+\kappa_1}, X_{K+\kappa_1+1}, \ldots, X_{K+\kappa_n}) \right).
\end{aligned}
$$

In (21), we used the (basic) coupling inequality, see e.g. [18, eq. (2.6)]. Therefore, by (20), $\left(X^{'}_{K+n} \right)_{n\in\mathbb{Z}}$ converges in total variation to $(X_n)_{n\in\mathbb{Z}}$, and the proof of (i) is completed.

(ii). Let $\left(\widetilde{X}_n\right)_{n\in\mathbb{Z}}$ be an ergodic version of the nonlinear marked Poisson autoregression with finite mean, say $\widetilde{\lambda}$, different from $(X_n)_{n\in\mathbb{Z}}$. We will check later on that $\left(\widetilde{X}_n\right)_{n\in\mathbb{Z}}$ satisfies the initial condition $\mathcal{C}_1$ on $\mathbb{Z}_{\leq 0}$. Therefore, by (i) we have that $\left(\widetilde{X}_{K+n}\right)_{n\in\mathbb{Z}}$ converges in law to $(X_n)_{n\in\mathbb{Z}}$ as $K \to \infty$. Since $\left(\widetilde{X}_n\right)_{n\in\mathbb{Z}}$ is stationary, for each $K \in \mathbb{N}$, $\left(\widetilde{X}_{K+n}\right)_{n\in\mathbb{Z}}$ has the same law as $\left(\widetilde{X}_n\right)_{n\in\mathbb{Z}}$, which implies that the processes $\left(\widetilde{X}_n\right)_{n\in\mathbb{Z}}$ and $(X_n)_{n\in\mathbb{Z}}$ have the same law, i.e., the claimed uniqueness in law. It remains to check that $\left(\widetilde{X}_n\right)_{n\in\mathbb{Z}}$ satisfies the initial condition $\mathcal{C}_1$ on $\mathbb{Z}_{\leq 0}$. We have

$$\sum_{k=-\infty}^{0} \overline{\alpha}_{n-k}\mathbb{E}|\widetilde{X}_k| = \widetilde{\lambda}\sum_{k=n}^{\infty}\overline{\alpha}_k \to 0, \quad \text{as } n \to \infty.$$

(iii). By (17), for all $K \in \mathbb{N}$, we have

$$\begin{aligned}\sum_{n=1}^{K} a_n &\leq \mathrm{Lip}(\phi)\sum_{n=1}^{K}\sum_{k=-\infty}^{0}\overline{\alpha}_{n-k}(\lambda + \mathbb{E}X'_n) + \mathrm{Lip}(\phi)\sum_{n=1}^{K}\sum_{k=1}^{n-1}\overline{\alpha}_{n-k}a_k \\ &\leq c + \mathrm{Lip}(\phi)\sum_{n=1}^{K} a_n, \end{aligned} \tag{22}$$

where

$$c := \lambda\mathrm{Lip}(\phi)\sum_{n=1}^{\infty} n\overline{\alpha}_n + \mathrm{Lip}(\phi)\sum_{n=1}^{\infty}\sum_{k=-\infty}^{0}\overline{\alpha}_{n-k}\mathbb{E}X'_n.$$

By the assumption $\sum_{k=1}^{\infty} k\overline{\alpha}_k < \infty$ and by the sharper initial condition $\widetilde{\mathcal{C}}_1$ we have $c < \infty$. From (22), we obtain

$$\sum_{n=1}^{K} a_n \leq \frac{c}{1-\mathrm{Lip}(\phi)},$$

and letting K tend to infinity we finally have

$$\sum_{n=1}^{\infty} a_n \leq \frac{c}{1-\mathrm{Lip}(\phi)} < \infty.$$

Therefore, by the definition of a_n, we have

$$\mathbb{E}\left[\sum_{n=1}^{\infty}\left|X_n - X'_n\right|\right] < \infty.$$

So the processes $\left(X'_n\right)_{n\in\mathbb{Z}}$ and $(X_n)_{n\in\mathbb{Z}}$ (taking values on $\mathbb{N}_0$) couple. □

Next, we consider nonlinear marked Poisson autoregressions with bounded and Lipschitz transfer functions, not necessarily contractive.

Theorem 6 (Convergence to equilibrium and uniqueness with bounded transfer) *Assume that ϕ is Lipschitz continuous and bounded,* (2) *and let $(X_n)_{n\in\mathbb{Z}}$ be a version of a nonlinear marked Poisson autoregression, as constructed in the proof of*

Theorem 3. Then:
(i) If $\left(X_n^{'}\right)_{n\in\mathbb{Z}}$ is a nonlinear marked Poisson autoregression with initial condition

$$\mathcal{C}_2\text{:}\quad \lim_{m\to\infty}\sum_{n=m+1}^{\infty}\sum_{k=-\infty}^{0}\sum_{i=1}^{X_k'}|\alpha_{n-k}(Z_{k,i})| = 0 \quad \mathbb{P}\text{-}a.s.,$$

then the processes $\left(X_n^{'}\right)_{n\in\mathbb{Z}}$ and $(X_n)_{n\in\mathbb{Z}}$ couple. In particular, $\left(X_{K+n}^{'}\right)_{n\in\mathbb{Z}}$ converges in total variation (and therefore in distribution) to $(X_n)_{n\in\mathbb{Z}}$ as $K\to\infty$.
(ii) The ergodic law of $(X_n)_{n\in\mathbb{Z}}$ is unique.

Proof (i). We construct a nonlinear marked Poisson autoregression $\left(X_n'\right)_{n\in\mathbb{Z}}$ with initial condition $\mathcal{C}_2$ on $\mathbb{Z}_{\le 0}$ exactly as in the proof of Theorem 5. Using the same Poisson process N and the same mark random variables $\left(Z_{k,i}\right)_{k\in\mathbb{Z},i\in\mathbb{N}}$, we construct $(X_n)_{n\in\mathbb{Z}}$ (and $(\lambda_n)_{n\in\mathbb{Z}}$). By the Lipschitz property of ϕ, for each $n\in\mathbb{N}$, we have

$$\begin{aligned}
|\lambda_n'-\lambda_n| &\le \mathrm{Lip}(\phi)\sum_{k=-\infty}^{n-1}\sum_{i=\min\{X_k^{'},X_k\}+1}^{\max\{X_k^{'},X_k\}}|\alpha_{n-k}(Z_{k,i})| \\
&= \mathrm{Lip}(\phi)\sum_{k=-\infty}^{0}\sum_{i=\min\{X_k^{'},X_k\}+1}^{\max\{X_k^{'},X_k\}}|\alpha_{n-k}(Z_{k,i})| \\
&\quad + \mathrm{Lip}(\phi)\sum_{k=1}^{n-1}\sum_{i=\min\{X_k^{'},X_k\}+1}^{\max\{X_k^{'},X_k\}}|\alpha_{n-k}(Z_{k,i})| \\
&\le \mathrm{Lip}(\phi)\sum_{k=-\infty}^{0}\sum_{i=1}^{X_k^{'}}|\alpha_{n-k}(Z_{k,i})| + \mathrm{Lip}(\phi)\sum_{k=-\infty}^{0}\sum_{i=1}^{X_k}|\alpha_{n-k}(Z_{k,i})| \\
&\qquad + \mathrm{Lip}(\phi)\sum_{k=1}^{n-1}\sum_{i=\min\{X_k^{'},X_k\}+1}^{\max\{X_k^{'},X_k\}}|\alpha_{n-k}(Z_{k,i})| \\
&\le \mathrm{Lip}(\phi)\sum_{k=-\infty}^{0}\sum_{i=1}^{X_k^{'}}|\alpha_{n-k}(Z_{k,i})| \\
&\quad + \mathrm{Lip}(\phi)\sum_{k=-\infty}^{0}\sum_{i=1}^{N([k-1,k),(0,\|\phi\|_\infty])}|\alpha_{n-k}(Z_{k,i})|, \\
&\qquad + \mathrm{Lip}(\phi)\sum_{k=1}^{n-1}\sum_{i=\min\{X_k^{'},X_k\}+1}^{\max\{X_k^{'},X_k\}}|\alpha_{n-k}(Z_{k,i})|, \qquad \mathbb{P}\text{-a.s.}. \qquad (23)
\end{aligned}$$

Let $\left(\mathcal{F}_t^N\right)_{t\in\mathbb{R}}$ be the natural filtration of N, i.e.,

$$\mathcal{F}_t^N := \sigma\left(N(A,B):\ A \in \mathcal{B}((-\infty,t]),\ B \in \mathcal{B}(\mathbb{R}_+)\right).$$

Set $\mathcal{G}_k := \mathcal{F}_k^N \vee \mathcal{F}_k^Z$. For $m \geq 0$ and $K \geq m+1$, we have

$$\begin{aligned}
&\mathbb{P}\left(\sum_{n=m+1}^{K} |X_n - X_n'| = 0 \,\middle|\, \mathcal{G}_{K-1}\right)\\
&= \mathbb{E}\left[\mathbf{1}\left(\sum_{n=m+1}^{K-1} |X_n - X_n'| = 0\right)\mathbf{1}\left(|X_K - X_K'| = 0\right) \,\middle|\, \mathcal{G}_{K-1}\right]\\
&= \mathbf{1}\left(\sum_{n=m+1}^{K-1} |X_n - X_n'| = 0\right)\mathbb{P}\left(|X_K - X_K'| = 0 \,\middle|\, \mathcal{G}_{K-1}\right)\\
&= \mathbf{1}\left(\sum_{n=m+1}^{K-1} |X_n - X_n'| = 0\right)\mathrm{e}^{-|\lambda_K - \lambda_K'|}\\
&\geq \mathbf{1}\left(\sum_{n=m+1}^{K-1} |X_n - X_n'| = 0\right)\exp\left\{-\left(c_K + \mathrm{Lip}(\phi)\sum_{k=1}^{K-1}\sum_{i=\min\{X_k',X_k\}+1}^{\max\{X_k',X_k\}} |\alpha_{K-k}(Z_{k,i})|\right)\right\} \qquad (24)\\
&= \mathbf{1}\left(\sum_{n=m+1}^{K-1} |X_n - X_n'| = 0\right)\exp\left\{-\left(c_K + \mathrm{Lip}(\phi)\sum_{k=1}^{m}\sum_{i=\min\{X_k',X_k\}+1}^{\max\{X_k',X_k\}} |\alpha_{K-k}(Z_{k,i})|\right)\right\}, \qquad (25)
\end{aligned}$$

where

$$c_K := \mathrm{Lip}(\phi)\sum_{k=-\infty}^{0}\sum_{i=1}^{X_k'} |\alpha_{K-k}(Z_{k,i})| + \mathrm{Lip}(\phi)\sum_{k=-\infty}^{0}\sum_{i=1}^{N([k-1,k),(0,\|\phi\|_\infty])} |\alpha_{K-k}(Z_{k,i})|$$

and in (24) we used (23). Note that $\mathbb{P}$-a.s. we have

$$\begin{aligned}
&c_K + \mathrm{Lip}(\phi)\sum_{k=1}^{m}\sum_{i=\min\{X_k',X_k\}+1}^{\max\{X_k',X_k\}} |\alpha_{K-k}(Z_{k,i})|\\
&\quad\leq c_K + \mathrm{Lip}(\phi)\sum_{k=1}^{m}\sum_{i=1}^{\max\{X_k',X_k\}} |\alpha_{K-k}(Z_{k,i})|\\
&\quad\leq c_K + \mathrm{Lip}(\phi)\sum_{k=1}^{m}\sum_{i=1}^{N([k-1,k),(0,\|\phi\|_\infty])} |\alpha_{K-k}(Z_{k,i})|\\
&\quad\leq \mathrm{Lip}(\phi)\sum_{k=-\infty}^{0}\sum_{i=1}^{X_k'} |\alpha_{K-k}(Z_{k,i})| + \mathrm{Lip}(\phi)\sum_{k=-\infty}^{m}\sum_{i=1}^{N([k-1,k),(0,\|\phi\|_\infty])} |\alpha_{K-k}(Z_{k,i})|
\end{aligned}$$

$$=: d_{m,K}.$$

Combining this inequality with relation (25), and then taking the conditional expectation with respect to $\mathcal{G}_m$, we have

$$\mathbb{P}\left(\sum_{n=m+1}^{K} |X_n - X'_n| = 0 \,|\, \mathcal{G}_m\right) \geq \mathrm{e}^{-d_{m,K}}\, \mathbb{P}\left(\sum_{n=m+1}^{K-1} |X_n - X'_n| = 0 \,|\, \mathcal{G}_m\right).$$

Iterating this inequality we clearly have

$$\mathbb{P}\left(\sum_{n=m+1}^{K} |X_n - X'_n| = 0 \,|\, \mathcal{G}_m\right) \geq \exp\left\{-\sum_{n=m+1}^{K} d_{m,n}\right\}.$$

Note that

$$\left\{\sum_{n=m+1}^{K} |X_n - X'_n| = 0\right\}_K \quad \text{is a decreasing sequence of events,}$$

therefore letting K tend to infinity we obtain

$$\forall m \in \mathbb{N}, \quad \mathbb{P}\left(\sum_{n=m+1}^{\infty} |X_n - X'_n| = 0 \,|\, \mathcal{G}_m\right) \geq \exp\left\{-\sum_{n=m+1}^{\infty} d_{m,n}\right\}, \quad \mathbb{P}\text{-a.s.},$$

which we rewrite as

$$\forall m \in \mathbb{N}, \quad \mathbb{P}\left(X_h = X'_h, \quad \forall \quad h \geq m+1 \,|\, \mathcal{G}_m\right) \geq F_m - \varepsilon_m \quad \mathbb{P}\text{-a.s.}, \tag{26}$$

where

$$F_m := \exp\left\{-\mathrm{Lip}(\phi) \sum_{n=m+1}^{\infty} \sum_{k=-\infty}^{m} \sum_{i=1}^{N([k-1,k),(0,\|\phi\|_\infty])} |\alpha_{n-k}(Z_{k,i})|\right\} \quad (\leq 1),$$

and

$$\varepsilon_m := F_m - \exp\left\{-\sum_{n=m+1}^{\infty} d_{m,n}\right\}.$$

Note that the processes $(X_n)_{n\in\mathbb{Z}}$ and $\left(X'_n\right)_{n\in\mathbb{Z}}$ are adapted with respect to the filtration $(\mathcal{G}_k)$, and that $(F_n)_{n\in\mathbb{Z}}$ is ergodic. Furthermore, note that from Wald's equality we obtain

$$\mathbb{E}\sum_{n=1}^{\infty} \sum_{k=-\infty}^{0} \sum_{i=1}^{N([k-1,k),(0,\|\phi\|_\infty])} |\alpha_{n-k}(Z_{k,i})| = \|\phi\|_\infty \sum_{n=1}^{\infty} \sum_{k=n}^{\infty} \overline{\alpha}_k = \|\phi\|_\infty \sum_{k=1}^{\infty} k\overline{\alpha}_k < \infty,$$

where the latter inequality follows by (2). Therefore $\mathbb{P}(F_0 > 0) = 1$. By the initial condition $\mathcal{C}_2$ it follows

$$
\begin{aligned}
0 \le \varepsilon_m &= F_m \left(1 - \mathrm{e}\left\{-\mathrm{Lip}(\phi) \sum_{n=m+1}^{\infty} \sum_{k=-\infty}^{0} \sum_{i=1}^{X_k^{'}} |\alpha_{n-k}(Z_{k,i})|\right\}\right) \\
&\le F_m \mathrm{Lip}(\phi) \sum_{n=m+1}^{\infty} \sum_{k=-\infty}^{0} \sum_{i=1}^{X_k^{'}} |\alpha_{n-k}(Z_{k,i})| \\
&\le \mathrm{Lip}(\phi) \sum_{n=m+1}^{\infty} \sum_{k=-\infty}^{0} \sum_{i=1}^{X_k^{'}} |\alpha_{n-k}(Z_{k,i})| \to 0, \quad \text{as } m \to \infty.
\end{aligned}
$$

Then by Lemma 5 in [6] we have that (26) implies that the processes $(X_n)_{n\in\mathbb{Z}}$ and $\left(X_n'\right)_{n\in\mathbb{Z}}$ couple. Convergence in variation (and therefore in law) is a consequence of coupling, and it can be shown similarly as in the proof of Theorem 5.

(ii). The proof is similar to the proof of Theorem 5(ii). Indeed, applying Wald's identity as after relation (4), we find that any ergodic version $(\widetilde{X}_n)_{n\in\mathbb{Z}}$ of the nonlinear marked Poisson autoregression satisfies

$$
\mathbb{E}\left[\sum_{n=1}^{\infty} \sum_{k=-\infty}^{0} \sum_{i=1}^{\widetilde{X}_k} |\alpha_{n-k}(Z_{k,i})|\right] = \widetilde{\lambda} \sum_{n=1}^{\infty} \sum_{k=-\infty}^{0} \overline{\alpha}_{n-k} = \widetilde{\lambda} \sum_{n\ge 1} n\overline{\alpha}_n
$$

and this latter quantity is finite by (2). Consequently, $(\widetilde{X}_n)_{n\in\mathbb{Z}}$ satisfies the initial condition $\mathcal{C}_2$. ◻

4 Rate of Convergence to Equilibrium

The following theorem provides a bound on the coupling time between an ergodic and a non-ergodic version of the nonlinear marked Poisson autoregression.

Theorem 7 (Coupling time with Lipschitz transfer) *Assume that the transfer function ϕ is Lipschitz continuous with $\mathrm{Lip}(\phi) < 1$ and (2). Let $\left(X_n^{'}\right)_{n\in\mathbb{Z}}$ be a nonlinear marked Poisson autoregression with initial condition $\widetilde{\mathcal{C}}_1$ on $\mathbb{Z}_{\le 0}$ and let $(X_n)_{n\in\mathbb{Z}}$ be a version of an ergodic nonlinear marked Poisson autoregression with finite mean. These processes are constructed in such a way that they couple (see Theorem 5(iii)).*

Letting T denote the coupling time between $\left(X_n^{'}\right)_{n\in\mathbb{Z}}$ and $(X_n)_{n\in\mathbb{Z}}$, we have

$$
\mathbb{P}(T > m) \le 1 - \exp\left\{-\sum_{n=m+1}^{\infty} z_n\right\}, \quad m \in \mathbb{N}, \tag{27}
$$

where

$$z_n := \sum_{\ell=0}^{\infty} \mathrm{Lip}(\phi)^{\ell} (\overline{\alpha}^{*\ell} * \beta)_n, \quad n \in \mathbb{N}$$

and

$$\beta_k := \lambda \mathrm{Lip}(\phi) \sum_{\ell=k}^{\infty} \overline{\alpha}_{\ell} + \mathrm{Lip}(\phi) \sum_{\ell=-\infty}^{0} \overline{\alpha}_{k-\ell} \mathbb{E} X'_{\ell}, \quad k \in \mathbb{N}. \tag{28}$$

Remark 1 Note that the initial condition $\widetilde{\mathcal{C}}_1$ and the assumption (2) guarantee that the upper bound for the tail probability in (27) is nontrivial. For later purposes, we also note that $(z_n)_{n\in\mathbb{Z}}$ satisfies the renewal equation

$$z_n = \beta_n + \sum_{\ell=1}^{\infty} \mathrm{Lip}(\phi)^{\ell} (\overline{\alpha}^{*\ell} * \beta)_n = \beta_n + \mathrm{Lip}(\phi)(\overline{\alpha} * z)_n, \quad n \in \mathbb{N}. \tag{29}$$

Finally, we remark that, under the assumptions of Theorem 7, for the mean λ of $(X_n)_{n\in\mathbb{Z}}$, we have

$$\lambda \leq \phi(0)/(1 - \mathrm{Lip}(\phi)). \tag{30}$$

Indeed

$$\begin{aligned}
\lambda &= \mathbb{E}\phi\left(\sum_{k=-\infty}^{-1} \sum_{i=1}^{X_k} \alpha_{-k}\right) \\
&\leq \phi(0) + \mathrm{Lip}(\phi)\mathbb{E}\left|\sum_{k=-\infty}^{-1} \sum_{i=1}^{X_k} \alpha_{-k}(Z_{k,i})\right| \\
&\leq \phi(0) + \mathrm{Lip}(\phi) \sum_{k=-\infty}^{-1} \mathbb{E}X_k \mathbb{E}|\alpha_{-k}(Z_{k,1})| \\
&= \phi(0) + \mathrm{Lip}(\phi)\lambda
\end{aligned}$$

which, since $\lambda < \infty$, yields the claim. So in the case when λ is unknown, we upper bound β_k by

$$\tilde{\beta}_k := \frac{\phi(0)\mathrm{Lip}(\phi)}{1 - \mathrm{Lip}(\phi)} \sum_{\ell=k}^{\infty} \overline{\alpha}_{\ell} + \mathrm{Lip}(\phi) \sum_{\ell=-\infty}^{0} \overline{\alpha}_{k-\ell} \mathbb{E} X'_{\ell}, \quad k \in \mathbb{N}, \tag{31}$$

and we obtain an explicit upper bound for the tail probability of the coupling time substituting β by $\tilde{\beta}$ in (27).

We proceed with the proof of Theorem 7.

Proof Construct $\left(X_n^{'}\right)_{n\in\mathbb{Z}}$ and $(X_n)_{n\in\mathbb{Z}}$ by embedding with respect to the same Poisson measure as in the proof of Theorem 5. For any $K \geq m+1$, $k \in \mathbb{N}$, setting $\mathcal{G}_j := \mathcal{F}_j^{X'} \vee \mathcal{F}_j^{X} \vee \mathcal{F}_j^{Z}$, $j \in \mathbb{N}$, by the Lipschitz property of ϕ we have

$$\begin{aligned}
&\mathbb{P}\left(\sum_{n=m+1}^{K} |X_n - X_n'| = 0 \,|\, \mathcal{G}_{K-1}\right)\\
&= \mathbb{E}\left[\mathbf{1}\left(\sum_{n=m+1}^{K-1} |X_n - X_n'| = 0\right) \mathbf{1}\left(|X_K - X_K'| = 0\right) \,|\, \mathcal{G}_{K-1}\right]\\
&= \mathbf{1}\left(\sum_{n=m+1}^{K-1} |X_n - X_n'| = 0\right) \mathbb{P}\left(|X_K - X_K'| = 0 \,|\, \mathcal{G}_{K-1}\right)\\
&= \mathbf{1}\left(\sum_{n=m+1}^{K-1} |X_n - X_n'| = 0\right) \mathrm{e}^{-|\lambda_K - \lambda_K'|}\\
&= \mathbf{1}\left(\sum_{n=m+1}^{K-1} |X_n - X_n'| = 0\right) \exp\left\{-\left|\phi\left(\sum_{k=-\infty}^{K-1}\sum_{i=1}^{X_k} \alpha_{K-k}(Z_{k,i})\right)\right.\right.\\
&\quad \left.\left. -\phi\left(\sum_{k=-\infty}^{K-1}\sum_{i=1}^{X_k'} \alpha_{K-k}(Z_{k,i})\right)\right|\right\}\\
&\geq \mathbf{1}\left(\sum_{n=m+1}^{K-1} |X_n - X_n'| = 0\right) \exp\left\{-\mathrm{Lip}(\phi)\sum_{k=-\infty}^{K-1}\sum_{i=\min\{X_k, X_k^{'}\}+1}^{\max\{X_k, X_k^{'}\}} |\alpha_{K-k}(Z_{k,i})|\right\}\\
&= \mathbf{1}\left(\sum_{n=m+1}^{K-1} |X_n - X_n'| = 0\right)\\
&\quad \exp\left\{-\mathrm{Lip}(\phi)\sum_{k=-\infty}^{0}\sum_{i=\min\{X_k, X_k^{'}\}+1}^{\max\{X_k, X_k^{'}\}} |\alpha_{K-k}(Z_{k,i})|\right.\\
&\quad \left. -\mathrm{Lip}(\phi)\sum_{k=1}^{m}\sum_{i=\min\{X_k, X_k^{'}\}+1}^{\max\{X_k, X_k^{'}\}} |\alpha_{K-k}(Z_{k,i})|\right\}.
\end{aligned}$$

Taking the conditional expectation with respect to $\mathcal{G}_m$, we obtain

$$\mathbb{P}\left(\sum_{n=m+1}^{K} |X_n - X_n'| = 0 \,|\, \mathcal{G}_m\right)$$

$$\geq \mathbb{P}\left(\sum_{n=m+1}^{K-1} |X_n - X_n'| = 0 \mid \mathcal{G}_m\right)$$
$$\exp\left\{-\mathrm{Lip}(\phi) \sum_{k=-\infty}^{0} \sum_{i=\min\{X_k, X_k'\}+1}^{\max\{X_k, X_k'\}} |\alpha_{K-k}(Z_{k,i})| - \mathrm{Lip}(\phi) \sum_{k=1}^{m} \sum_{i=\min\{X_k, X_k'\}+1}^{\max\{X_k, X_k'\}} |\alpha_{K-k}(Z_{k,i})|\right\}.$$

Iterating this inequality we then have

$$\mathbb{P}\left(\sum_{n=m+1}^{K} |X_n - X_n'| = 0 \mid \mathcal{G}_m\right)$$
$$\geq \exp\left\{-\mathrm{Lip}(\phi) \sum_{n=m+1}^{K} \left(\sum_{k=-\infty}^{0} \sum_{i=\min\{X_k, X_k'\}+1}^{\max\{X_k, X_k'\}} |\alpha_{n-k}(Z_{k,i})| + \sum_{k=1}^{m} \sum_{i=\min\{X_k, X_k'\}+1}^{\max\{X_k, X_k'\}} |\alpha_{n-k}(Z_{k,i})|\right)\right\}.$$

Letting $K \to \infty$, taking the expectation, using Jensen's inequality, and then Wald's identity as after (4), we obtain for all $m \in \mathbb{N}$

$$\mathbb{P}(T \leq m) = \mathbb{P}\left(\sum_{n=m+1}^{\infty} |X_n - X_n'| = 0\right)$$
$$\geq \exp\left\{-\mathrm{Lip}(\phi) \sum_{n=m+1}^{\infty} \left(\sum_{k=-\infty}^{0} \mathbb{E}|\alpha_{n-k}(Z_{k,1})|\mathbb{E}|X_k - X_k'| + \sum_{k=1}^{m} \mathbb{E}|\alpha_{n-k}(Z_{k,1})|\mathbb{E}|X_k - X_k'|\right)\right\}$$
$$\geq \exp\left\{-\lambda \mathrm{Lip}(\phi) \sum_{n=m+1}^{\infty} \sum_{k=n}^{\infty} \overline{\alpha}_k - \mathrm{Lip}(\phi) \sum_{n=m+1}^{\infty} \sum_{k=-\infty}^{0} \overline{\alpha}_{n-k} \mathbb{E} X_k' \right.$$
$$\left. - \mathrm{Lip}(\phi) \sum_{n=m+1}^{\infty} \sum_{k=1}^{m} \overline{\alpha}_{n-k} \mathbb{E}|\lambda_k - \lambda_k'|\right\}$$
$$= \exp\left\{- \sum_{n=m+1}^{\infty} \beta_n - \mathrm{Lip}(\phi) \sum_{n=m+1}^{\infty} \sum_{k=1}^{m} \overline{\alpha}_{n-k} \mathbb{E}|\lambda_k - \lambda_k'|\right\}, \tag{32}$$

where (β_k) is defined as in (28). To proceed with the chain of inequalities, we calculate a suitable upper bound for $\mathbb{E}|\lambda_k - \lambda_k'|$. For $k \in \mathbb{N}$, by the Lipschitz property of ϕ we have almost surely

$$|\lambda_k - \lambda_k'| \leq \mathrm{Lip}(\phi) \sum_{\ell=-\infty}^{k-1} \sum_{i=\min\{X_\ell, X_\ell'\}+1}^{\max\{X_\ell, X_\ell'\}} \left|\alpha_{k-\ell}(Z_{\ell,i})\right|.$$

Taking the expectation and applying Wald's identity as above, we obtain

$$
\begin{aligned}
\mathbb{E}[|\lambda_k-\lambda'_k|] &\le \mathrm{Lip}(\phi)\sum_{\ell=-\infty}^{0}\overline{\alpha}_{k-\ell}\mathbb{E}[|X_\ell-X'_\ell|]+\mathrm{Lip}(\phi)\sum_{\ell=1}^{k-1}\overline{\alpha}_{k-\ell}\mathbb{E}[|X_\ell-X'_\ell|]\\
&\le \lambda\mathrm{Lip}(\phi)\sum_{\ell=k}^{\infty}\overline{\alpha}_\ell+\mathrm{Lip}(\phi)\sum_{\ell=-\infty}^{0}\overline{\alpha}_{k-\ell}\mathbb{E}X'_\ell+\mathrm{Lip}(\phi)\sum_{\ell=1}^{k-1}\overline{\alpha}_{k-\ell}\mathbb{E}[|\lambda_\ell-\lambda'_\ell]\\
&=\beta_k+\mathrm{Lip}(\phi)\sum_{\ell=1}^{k-1}\overline{\alpha}_{k-\ell}\mathbb{E}[|\lambda_\ell-\lambda'_\ell|]. \qquad (33)
\end{aligned}
$$

Finally, setting

$$\gamma_k := \mathbb{E}[|\lambda_k-\lambda'_k|]\mathbf{1}\,(k\ge 1)\,,\quad k\in\mathbb{Z},$$

we obtain from (33) the renewal inequality

$$\gamma_k \le \beta_k+\mathrm{Lip}(\phi)(\overline{\alpha}*\gamma)_k.$$

Iterating this inequality, we have

$$\gamma_k \le \sum_{i=0}^{n}\mathrm{Lip}(\phi)^i(\overline{\alpha}^{*i}*\beta)_k+\mathrm{Lip}(\phi)^{n+1}\left(\gamma*\overline{\alpha}^{*(n+1)}\right)_k,\quad k\in\mathbb{N}.$$

We note that $\sup_{k\in\mathbb{Z}}\gamma_k<\infty$. Indeed

$$\sum_{k\in\mathbb{Z}}\gamma_k=\sum_{k\in\mathbb{N}}\mathbb{E}|\lambda_k-\lambda'_k|=\sum_{k\in\mathbb{N}}\mathbb{E}|X_k-X'_k|<\infty,$$

where the finiteness of the latter infinite sum follows as in the proof of Theorem 5(iii). Therefore $\mathrm{Lip}(\phi)^{n+1}\left(\gamma*\overline{\alpha}^{*(n+1)}\right)_k\to 0$, as $n\to\infty$, and so

$$\gamma_k\le\sum_{\ell=0}^{\infty}\mathrm{Lip}(\phi)^\ell(\overline{\alpha}^{*\ell}*\beta)_k.$$

Plugging this upper bound into (32), we conclude that

$$
\begin{aligned}
\mathbb{P}(T>m) &\le 1-\exp\left\{-\sum_{n=m+1}^{\infty}\beta_n-\mathrm{Lip}(\phi)\sum_{n=m+1}^{\infty}\sum_{k=1}^{m}\overline{\alpha}_{n-k}\sum_{\ell=0}^{\infty}\mathrm{Lip}(\phi)^\ell(\overline{\alpha}^{*\ell}*\beta)_k\right\}\\
&\le 1-\exp\left\{-\sum_{n=m+1}^{\infty}\beta_n-\sum_{n=m+1}^{\infty}\sum_{\ell=0}^{\infty}\mathrm{Lip}(\phi)^{\ell+1}(\overline{\alpha}^{*(\ell+1)}*\beta)_n\right\}\\
&=1-\exp\left\{-\sum_{n=m+1}^{\infty}z_n\right\},\quad m\in\mathbb{N},
\end{aligned}
$$

and the proof is completed. □

The following two theorems provide the rate of convergence to zero of the tail of the coupling time respectively under a light-tail and a heavy-tail assumption on the probability distribution $(\overline{\alpha}_k)$.

Theorem 8 *[Rate of convergence in the light-tailed case] Let the assumptions and notation of Theorem 7 prevail. Assume that there exists a positive constant $\eta > 0$ such that*

$$\mathrm{Lip}(\phi) \sum_{k \in \mathbb{N}} \mathrm{e}^{\eta k} \overline{\alpha}_k = 1, \qquad \sum_{k \in \mathbb{N}} \mathrm{e}^{\eta k} \beta_k < \infty \quad \textit{and} \quad \sum_{k \in \mathbb{N}} k \mathrm{e}^{\eta k} \overline{\alpha}_k < \infty. \tag{34}$$

Then, for any $\varepsilon > 0$ there exists $m_\varepsilon \geq 1$ such that for all $m \geq m_\varepsilon$,

$$\mathbb{P}(T > m) \leq 1 - \exp\left\{-(C_\eta + \varepsilon) \frac{\mathrm{e}^{-(m+1)\eta}}{1 - \mathrm{e}^{-\eta}}\right\}, \quad \textit{where} \quad C_\eta := \frac{\sum_{k \in \mathbb{N}} \mathrm{e}^{\eta k} \beta_k}{\mathrm{Lip}(\phi) \sum_{k \in \mathbb{N}} k \mathrm{e}^{\eta k} \overline{\alpha}_k}.$$

And so $\mathbb{P}(T > m) = O(\mathrm{e}^{-(m+1)\eta})$, as $m \to \infty$.

Proof By (29) for any $k \in \mathbb{N}$ we have

$$\mathrm{e}^{\eta k} z_k = \mathrm{e}^{\eta k} \beta_k + \mathrm{e}^{\eta k} \mathrm{Lip}(\phi) \sum_{j \in \mathbb{Z}} z_{k-j} \overline{\alpha}_j = \mathrm{e}^{\eta k} \beta_k + \mathrm{Lip}(\phi) \sum_{j=1}^{k} \mathrm{e}^{\eta(k-j)} z_{k-j} \mathrm{e}^{\eta j} \overline{\alpha}_j,$$

i.e., the sequence $(\mathrm{e}^{\eta k} z_k)_{k \in \mathbb{N}}$ solves a proper (due to the light-tail assumption, i.e., the first relation in (34)) renewal equation. The second and the third relation in (34) allow us to apply the key renewal theorem (see, e.g., Theorem 2.2 in [3]), which yields

$$\lim_{k \to \infty} \mathrm{e}^{\eta k} z_k = C_\eta.$$

So, for any $\varepsilon > 0$ there exists $k_\varepsilon \geq 1$ such that for all $k \geq k_\varepsilon$ we have $z_k < (C_\eta + \varepsilon)\mathrm{e}^{-\eta k}$. Combining this with (27) we finally have that for any $\varepsilon > 0$ there exists $m_\varepsilon \geq 1$ such that for all $m \geq m_\varepsilon$

$$\mathbb{P}(T > m) \leq 1 - \exp\left\{-(C_\eta + \varepsilon) \sum_{n=m+1}^{\infty} \mathrm{e}^{-\eta n}\right\} = 1 - \exp\left\{-(C_\eta + \varepsilon) \frac{\mathrm{e}^{-(m+1)\eta}}{1 - \mathrm{e}^{-\eta}}\right\},$$

and the proof is completed. □

Next, we consider the case when the distribution function $A(n) := \sum_{k=1}^{n} \overline{\alpha}_j, n \in \mathbb{N}$, is subexponential, i.e., for all $n \geq 2$

$$\lim_{x \to \infty} \frac{\overline{A^{*n}}(x)}{\overline{A}(x)} = n,$$

where $\overline{A} := 1 - A$; we refer the reader to [11, 14] for more insight into subexponential distributions.

Theorem 9 (Rate of convergence in the heavy-tailed case) *Let the assumptions and notation of Theorem 7 prevail, and assume that* $\left(X_n^{'}\right)_{n\in\mathbb{Z}}$ *satisfies the initial condition*

$$\sup_{n\le 0} \mathbb{E}|X_n'| < \infty.$$

If A is subexponential, then, for any $\varepsilon > 0$ there exists $m_\varepsilon \ge 1$ such that for all $m \ge m_\varepsilon$,

$$\mathbb{P}(T > m) \le 1 - \exp\left\{-\left[\mathrm{Lip}(\phi)\left(\frac{\lambda}{1-\mathrm{Lip}(\phi)} + \sup_{n\le 0}\mathbb{E}|X_n'|\right) + \varepsilon\right]\sum_{n\ge m+1}\overline{A}(n)\right\}.$$

And so $\mathbb{P}(T > m) = O\left(\sum_{n\ge m+1}\overline{A}(n)\right)$, as $m \to \infty$.

Proof Setting $R_k := \mathrm{Lip}(\phi)\sum_{\ell=-\infty}^{0}\overline{\alpha}_{k-\ell}\mathbb{E}X_\ell'$, the sequence β from (28) reads

$$\beta_k = R_k + \lambda\mathrm{Lip}(\phi)\sum_{\ell=k}^{\infty}\overline{\alpha}_\ell = R_k + \lambda\mathrm{Lip}(\phi)\overline{A}(k-1), \quad k \in \mathbb{N}.$$

We recall (see Definition 1) that, being A a distribution function, we have the relations

$$A^{*0} \equiv 1 \quad \text{and} \quad A^{*i}(n) = \sum_{k=1}^{n}\overline{\alpha}_{n-k}A^{*(i-1)}(k) = \sum_{k=1}^{n}\overline{\alpha}_{n-k}^{*(i-1)}A(k), \quad n, i \ge 1.$$

Therefore

$$\begin{aligned}
z_n &= R_n + \lambda\mathrm{Lip}(\phi)\sum_{k=1}^{n}\overline{A}(k-1)\sum_{i=0}^{\infty}\mathrm{Lip}(\phi)^i\overline{\alpha}_{n-k}^{*i} \\
&= R_n + \lambda\mathrm{Lip}(\phi)\sum_{k=1}^{n}(1 - A(k-1))\sum_{i=0}^{\infty}\mathrm{Lip}(\phi)^i\overline{\alpha}_{n-k}^{*i} \\
&= R_n + \lambda\mathrm{Lip}(\phi)\sum_{k=1}^{n}\sum_{i=0}^{\infty}\mathrm{Lip}(\phi)^i\overline{\alpha}_{n-k}^{*i} - \lambda\mathrm{Lip}(\phi)\sum_{i=0}^{\infty}\mathrm{Lip}(\phi)^i\sum_{k=1}^{n}A(k-1)\overline{\alpha}_{n-k}^{*i} \\
&= R_n + \lambda\mathrm{Lip}(\phi)\sum_{i=0}^{\infty}\mathrm{Lip}(\phi)^i\sum_{k=1}^{n}\overline{\alpha}_{n-k}^{*i} - \lambda\mathrm{Lip}(\phi)\sum_{i=0}^{\infty}\mathrm{Lip}(\phi)^i\sum_{k=1}^{n-1}A(k)\overline{\alpha}_{n-1-k}^{*i} \\
&= R_n + \lambda\sum_{i=0}^{\infty}\mathrm{Lip}(\phi)^{i+1}A^{*i}(n) - \lambda\sum_{i=0}^{\infty}\mathrm{Lip}(\phi)^{i+1}A^{*(i+1)}(n-1)
\end{aligned}$$

$$= R_n + \lambda \sum_{i=0}^{\infty} \mathrm{Lip}(\phi)^{i+1} \left(A^{*i}(n) - A^{*(i+1)}(n-1) \right)$$
$$= R_n + \lambda \sum_{i=0}^{\infty} \mathrm{Lip}(\phi)^{i+1} \left(\overline{A^{*(i+1)}}(n-1) - \overline{A^{*i}}(n) \right). \tag{35}$$

Note that by the subexponentiality of the distribution function A, we have

$$\lim_{n\to\infty} \frac{\overline{A^{*(i+1)}}(n-1) - \overline{A^{*i}}(n)}{\overline{A}(n)} = (i+1) - i = 1, \quad \text{for any } i. \tag{36}$$

Let $\delta > 0$ be so small that $\mathrm{Lip}(\phi)(1+\delta) < 1$. By the subexponentiality of A and Kesten's bound, see e.g. [11, Lemma 1.3.5], there exists a constant K_δ such that for any $i \geq 0$ and $n \geq 1$

$$\frac{\overline{A^{*(i+1)}}(n-1) - \overline{A^{*i}}(n)}{\overline{A}(n)} \leq K_\delta (2+\delta)(1+\delta)^i.$$

By the choice of δ we have

$$\sum_{i=0}^{\infty} \mathrm{Lip}(\phi)^{i+1} (1+\delta)^i < \infty.$$

Thus, by (35), (36) and the dominated convergence theorem, we have

$$\limsup_{n\to\infty} \frac{z_n}{\overline{A}(n)} \leq \mathrm{Lip}(\phi) \sup_{n\leq 0} \mathbb{E}|X_n'| + \lambda \frac{\mathrm{Lip}(\phi)}{1-\mathrm{Lip}(\phi)}.$$

So, for any $\varepsilon > 0$ there exists $m_\varepsilon \geq 1$ such that for all $n \geq m_\varepsilon$ we have

$$z_n < [\lambda \mathrm{Lip}(\phi)/(1-\mathrm{Lip}(\phi)) + \mathrm{Lip}(\phi) \sup_{n\leq 0} \mathbb{E}|X_n'| + \varepsilon] \overline{A}(n).$$

Plugging this upper bound into (27) yields the claim. □

5 Examples

In this section, we illustrate the theory developed in the previous sections in the case of a nonlinear marked Poisson autoregression with a "softplus" transfer function

$$\phi(x) := \log\left(1 + \mathrm{e}^{c+\mu x}\right), \quad x \in \mathbb{R},\ c \in \mathbb{R},\ \mu \in [0,1), \tag{37}$$

and with marks $\left(Z_{n,i}\right)_{n\in\mathbb{Z}, i\in\mathbb{N}}$ such that $Z_{n,i} = Z_{n,1}$, $n \in \mathbb{Z}$, $i \in \mathbb{N}$, where $Z_{n,1}$, $n \in \mathbb{Z}$, are independent exponential random variables with mean 1. Note that ϕ is increasing, convex and unbounded; moreover it is Lipschitz continuous with $\mathrm{Lip}(\phi) = \mu$, indeed

$$\sup_{x\in\mathbb{R}} \phi'(x) = \sup_{x\in\mathbb{R}} \frac{\mu e^{c+\mu x}}{1+e^{c+\mu x}} = \sup_{x\in\mathbb{R}} \frac{\mu}{1+e^{-c-\mu x}} = \mu < 1.$$

The softplus transfer function seems a very natural choice. First of all, it seems natural to look for a non-decreasing function ϕ with $\lim_{x\to-\infty} \phi(x) = 0$ and $\lim_{x\to\infty} \phi(x) = \infty$. Secondly, note that exponential functions are excluded, because they are not Lipschitz, and also linear functions are not possible, as linear functions are not nonnegative on the real line, and we want to make negative coefficient functions, and thus "inhibition" possible. These restrictions leave us with the choice $(mx + c)^+$ (which is not everywhere differentiable) or "something like" the softplus function. The smoothness of the softplus function makes numerical optimization procedures for estimation unproblematic.

Light-tailed case

Suppose that the time series has coefficient functions

$$\alpha_k^{(\text{light})} : \mathbb{R}_{>0} \to \mathbb{R}, \quad z \mapsto (-1)^k \frac{1-p}{p} p^k z, \quad k \in \mathbb{N}, \tag{38}$$

for some $p \in (0, 1)$. We have

$$\sum_{k=1}^{\infty} \overline{\alpha}_k^{(\text{light})} = \sum_{k=1}^{\infty} \mathbb{E}\left[\left|\alpha_k^{(\text{light})}(Z_{0,1})\right|\right] = \sum_{k=1}^{\infty} \frac{1-p}{p} p^k \mathbb{E}[Z_{0,1}] = 1.$$

Let $\left(X_n'\right)_{n\in\mathbb{Z}}$ be a nonlinear marked Poisson autoregression with initial condition $\mathbb{E}X_n' \equiv \lambda^{(\text{initial})}$, $n \in \mathbb{Z}_{\leq 0}$, for some constant $\lambda^{(\text{initial})} \geq 0$, and let $(X_n)_{n\in\mathbb{Z}}$ be a version of an ergodic nonlinear marked Poisson autoregression with finite mean, as constructed in Theorem 2. We have

$$\sum_{n=1}^{\infty} \sum_{k=-\infty}^{0} \overline{\alpha}_{n-k}^{(\text{light})} = \sum_{n=1}^{\infty} n\overline{\alpha}_n^{(\text{light})} = \frac{1-p}{p} \sum_{n=1}^{\infty} np^n < \infty,$$

so the initial condition $\widetilde{\mathcal{C}}_1$ is satisfied. Thus, we obtain from Theorem 5 that the transfer function ϕ together with the coefficient functions $\alpha_k^{(\text{light})}$ and the exponential mark distribution specify the unique ergodic distribution of the time series.

Furthermore, Theorem 8 can be applied to this example since the first and the third relation in (34) are fulfilled with $\eta = -\log\{\mu(1-p) + p\} > 0$. Indeed,

$$\mathrm{Lip}(\phi) \sum_{k=1}^{\infty} e^{\eta k} \overline{\alpha}_k^{(\text{light})} = \mu \frac{1-p}{p} \sum_{k=1}^{\infty} (e^{\eta} p)^k \mathbb{E}[Z_{0,1}] = \mu \frac{1-p}{p} \frac{e^{\eta} p}{1-e^{\eta} p} = \mu \frac{1-p}{e^{-\eta} - p} = 1$$

and

$$\sum_{k=1}^{\infty} k\mathrm{e}^{\eta k}\overline{\alpha}_k^{(\text{light})} = \frac{1-p}{p}\sum_{k=1}^{\infty} k(p\mathrm{e}^{\eta})^k = \frac{1-p}{p}\sum_{k=1}^{\infty} k\left(\frac{p}{p+\mu(1-p)}\right)^k$$
$$= \frac{p+\mu(1-p)}{\mu^2(1-p)}.$$

Also the second relation in (34) is satisfied, as the following argument shows. Hereafter, we work with the sequence $(\tilde{\beta}_k)_{k\in\mathbb{N}}$ defined in (31), which is an upper bound for the sequence $(\beta_k)_{k\in\mathbb{N}}$. We have

$$\begin{aligned}\tilde{\beta}_k &= \phi(0)\frac{\operatorname{Lip}(\phi)}{1-\operatorname{Lip}(\phi)}\sum_{\ell=k}^{\infty}\overline{\alpha}_\ell^{(\text{light})} + \operatorname{Lip}(\phi)\sum_{\ell=-\infty}^{0}\overline{\alpha}_{k-\ell}^{(\text{light})}\mathbb{E}X'_\ell \\ &= \phi(0)\frac{\mu}{1-\mu}\frac{1-p}{p}\sum_{\ell=k}^{\infty}p^\ell + \mu\lambda^{(\text{initial})}\frac{1-p}{p}\sum_{\ell=-\infty}^{0}p^{k-\ell} \\ &= \phi(0)\frac{\mu}{1-\mu}\frac{1-p}{p}\frac{p^k}{1-p} + \mu\lambda^{(\text{initial})}\frac{1-p}{p}\frac{p^k}{1-p} \\ &= \mu\left(\frac{\log(1+\mathrm{e}^c)}{1-\mu}+\lambda^{(\text{initial})}\right)p^{k-1},\end{aligned}$$

and so

$$\sum_{k\in\mathbb{N}}\mathrm{e}^{\eta k}\tilde{\beta}_k = \mu\left(\frac{\log(1+\mathrm{e}^c)}{1-\mu}+\lambda^{(\text{initial})}\right)\sum_{k\in\mathbb{N}}\mathrm{e}^{\eta k}p^{k-1} = \frac{1}{1-p}\left(\frac{\log(1+\mathrm{e}^c)}{1-\mu}+\lambda^{(\text{initial})}\right).$$

We have verified the three assumptions in (34), and thus we obtain from Theorem 8 that, for any $\varepsilon > 0$, there exists $m_\varepsilon \geq 1$ such that for all $m \geq m_\varepsilon$,

$$\mathbb{P}(T > m) \leq 1 - \exp\left\{-(\tilde{C}_\eta + \varepsilon)\frac{\mathrm{e}^{-(m+1)\eta}}{1-\mathrm{e}^{-\eta}}\right\}, \tag{39}$$

where

$$\begin{aligned}\tilde{C}_\eta &:= \frac{\sum_{k\in\mathbb{N}}\mathrm{e}^{\eta k}\tilde{\beta}_k}{\operatorname{Lip}(\phi)\sum_{k\in\mathbb{N}}k\mathrm{e}^{\eta k}\overline{\alpha}_k^{(\text{light})}} = \frac{\frac{1}{1-p}\left(\frac{\log(1+\mathrm{e}^c)}{1-\mu}+\lambda^{(\text{initial})}\right)}{\frac{p+\mu(1-p)}{\mu(1-p)}} \\ &= \frac{\mu\left(\frac{\log(1+\mathrm{e}^c)}{1-\mu}+\lambda^{(\text{initial})}\right)}{\mu(1-p)+p}.\end{aligned}$$

Plugging this expression into (39), we obtain for the light-tailed case the asymptotic bound

$$\mathbb{P}(T > m) \leq 1 - \exp\left\{ - \left(\frac{\mu \left(\frac{\log(1+e^c)}{1-\mu} + \lambda^{(\text{initial})} \right)}{\mu(1-p)+p} + \varepsilon \right) \frac{(\mu(1-p)+p)^{m+1}}{1-(\mu(1-p)+p)} \right\}. \tag{40}$$

Subexponential case

Suppose that the time series has coefficient functions

$$\alpha_k^{(\text{subexp})} : \mathbb{R}_{>0} \to \mathbb{R}, \quad z \mapsto (-1)^k \zeta(\gamma)^{-1} k^{-\gamma} z, \tag{41}$$

for some $\gamma > 2$, where $\zeta(\cdot)$ denotes the Riemann zeta function. Note that $\sum_{k=1}^{\infty} \overline{\alpha}_k^{(\text{subexp})} = 1$. Let $\left(X_n'\right)_{n \in \mathbb{Z}}$ be a nonlinear marked Poisson autoregression with initial condition $\mathbb{E}X_n' = 0$ for each $n \in \mathbb{Z}_{\leq 0}$ (we assume this particular initial condition only for ease of notation; clearly, with minor modifications, one could consider the more general initial condition $\sup_{n \leq 0} \mathbb{E}X_n' < \infty$) and let $(X_n)_{n \in \mathbb{Z}}$ be a version of an ergodic nonlinear marked Poisson autoregression with finite mean, as constructed in Theorem 2. Clearly the initial condition $\widetilde{\mathcal{C}}_1$ is satisfied and by Theorem 5 we know that the transfer function ϕ together with the coefficient functions $\alpha_k^{(\text{subexp})}$ and the exponential mark distribution specify the unique ergodic distribution of the time series. Since the law $\left(\overline{\alpha}_k^{(\text{subexp})}\right)_{k \geq 1}$ is subexponential (see, e.g., [14, Theorem 4.14]) all the assumptions of Theorem 9 are satisfied and, exploiting the inequality (30), we have that, for any $\varepsilon > 0$ there exists $m_\varepsilon \geq 1$ such that for all $m \geq m_\varepsilon$,

$$\begin{aligned}
\mathbb{P}(T > m) &\leq 1 - \exp\left\{ - \left(\phi(0) \frac{\text{Lip}(\phi)}{(1-\text{Lip}(\phi))^2} + \varepsilon \right) \sum_{n=m+1}^{\infty} \overline{A}(n) \right\} \\
&= 1 - \exp\left\{ - \left(\phi(0) \frac{\mu}{(1-\mu)^2} + \varepsilon \right) \zeta(\gamma)^{-1} \sum_{n=m+1}^{\infty} \sum_{k=n+1}^{\infty} k^{-\gamma} \right\} \\
&= 1 - \exp\left\{ - \left(\phi(0) \frac{\mu}{(1-\mu)^2} + \varepsilon \right) \zeta(\gamma)^{-1} \sum_{k=m+2}^{\infty} \sum_{n=m+1}^{k-1} k^{-\gamma} \right\} \\
&= 1 - \exp\left\{ - \left(\phi(0) \frac{\mu}{(1-\mu)^2} + \varepsilon \right) \zeta(\gamma)^{-1} \sum_{k=m+2}^{\infty} (k-m-1) k^{-\gamma} \right\} \\
&= 1 - \exp\left\{ - \left(\phi(0) \frac{\mu}{(1-\mu)^2} + \varepsilon \right) \zeta(\gamma)^{-1} \left(\sum_{k=m+2}^{\infty} k^{1-\gamma} - (m+1) \sum_{k=m+2}^{\infty} k^{-\gamma} \right) \right\}.
\end{aligned} \tag{42}$$

Note that, for a non-increasing function $f : \mathbb{R} \to \mathbb{R}_{\geq 0}$, for all $m \in \mathbb{N}$, we have

$$\int_m^\infty f(x)\,dx \leq \sum_{k=m}^\infty f(k) \leq \int_{m-1}^\infty f(x)\,dx.$$

Therefore, since $\gamma > 2$, we have

$$\sum_{k=m+2}^\infty k^{1-\gamma} \leq \int_{m+1}^\infty x^{1-\gamma}\,dx = \frac{(m+1)^{2-\gamma}}{\gamma-2},$$

and

$$\sum_{k=m+2}^\infty k^{-\gamma} \geq \frac{(m+2)^{1-\gamma}}{\gamma-1}.$$

Combining these inequalities we find

$$\sum_{k=m+2}^\infty k^{1-\gamma} - (m+1)\sum_{k=m+2}^\infty k^{-\gamma} \leq \frac{(m+1)^{2-\gamma}}{\gamma-2} - (m+1)\frac{(m+2)^{1-\gamma}}{\gamma-1}.$$

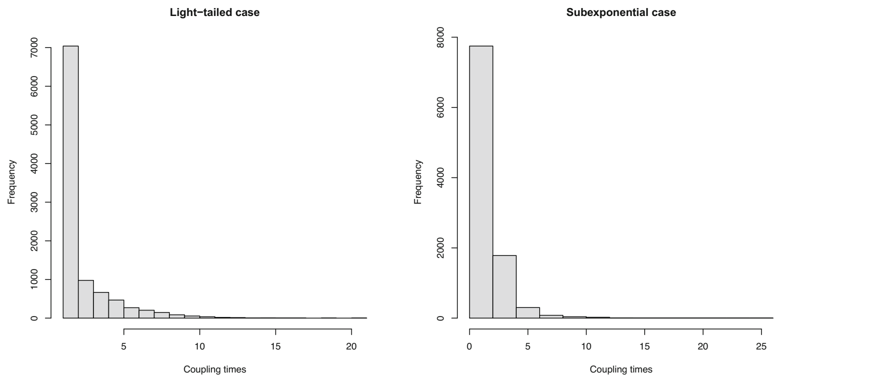
Plugging this expression into (42), we finally obtain for the subexponential case the asymptotic bound

$$\mathbb{P}(T > m) \leq 1 - \exp\left\{-\left(\frac{\phi(0)\mu}{(1-\mu)^2} + \varepsilon\right)\zeta(\gamma)^{-1}\left(\frac{(m+1)^{2-\gamma}}{\gamma-2} - (m+1)\frac{(m+2)^{1-\gamma}}{\gamma-1}\right)\right\}. \quad (43)$$

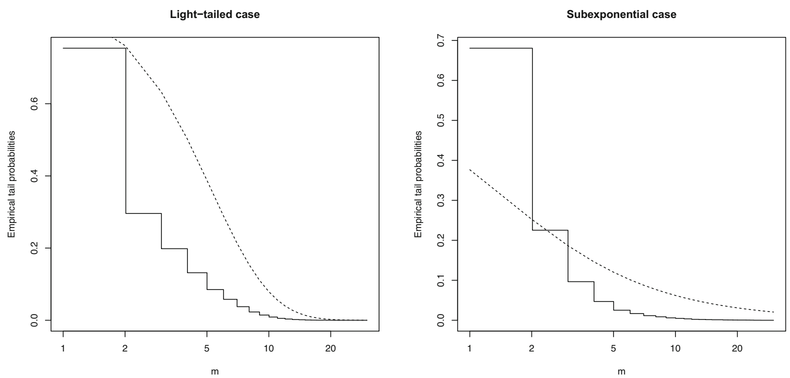
See Fig. 1 for a numerical illustration. In particular, we have chosen $c = 1$ and $\mu = 0.4$ for the transfer function ϕ from (37), $p = 0.5$ for the light-tailed case (see (38)) and $\gamma = 3$ for the subexponential case (see (41)). For each model, we simulated pairs, once with a void initial condition, and once with a burn-in of 100 time steps to mimic an ergodic version. From time 0 on, both simulations are constructed with respect to the same realization of the driving Poisson process, and also the same realization of mark variables. For both models, we simulated 10'000 pairs giving us 10'000 realizations of coupling times each.

6 Outlook

One could generalize nonlinear Poisson autoregressions in various directions. For instance, one could introduce *multivariate* nonlinear Poisson autoregressions. Such



(a) Histograms of the coupling times. Note that in both models, over 70% of the simulations couple after only one time step.



(b) The solid lines give the empirical tail function of the coupling times. The dashed lines give the theoretical upper bound for the tail function from equations (**??**) (light-tailed case) and (**??**) (subexponential case), respectively.

Fig. 1 Numerical illustration of the examples from Sect. 5. The parameters are $c = 1$ and $\mu = 0.4$ for the transfer function ϕ from (37), and $p = 0.5$ for the light-tailed case (see (38)) and $\gamma = 3$ for the subexponential case (see (41)). From each model, we simulate pairs—once with a void initial condition, and once with a burn-in of 100 time steps to mimic an ergodic version. From time 0 on, both simulations are constructed with respect to the same realization of the driving Poisson process and to the same realization of mark variables. For both models, we simulate 10'000 such path pairs, giving us thus 10'000 realizations of coupling times each. We give histograms **a** and empirical tail functions **b** of the coupling times

time series are defined by considering $d \in \mathbb{N}$ different type of event counts $X = (X^{(1)}, \dots, X^{(d)})$, with $X^{(i)} = (X_n^{(i)})_{n\in\mathbb{Z}}$, $i = 1, \dots, d$,

$$X_n^{(i)} \mid \mathcal{F}_{n-1}^X \vee \mathcal{F}_{n-1}^Z \sim \mathrm{Pois}(\lambda_n^{(i)}), \quad \text{and} \quad \lambda_n^{(i)} := \sum_{j=1}^{d} \phi_{i,j} \left(\sum_{k=-\infty}^{n-1} \sum_{l=1}^{X_k^{(j)}} \alpha_{n-k}^{(i,j)}(Z_{k,l}^{(j)}) \right), \quad n \in \mathbb{Z},$$

where

- $\mathcal{F}_n^X := \sigma\left(X_k^{(j)} : \ k \le n, \, j = 1, \dots, d\right), n \in \mathbb{Z},$

- for any fixed $j = 1, \ldots, d$, $\left(Z_{k,l}^{(j)}\right)_{k\in\mathbb{Z},l\in\mathbb{N}}$ are identically distributed mark random variables with values in a measurable space $(\mathrm{Z}, \mathcal{Z})$ such that the mark sequences $\left(Z_{k,l}^{(j)}\right)_{l\in\mathbb{N}}$ are independent over $k \in \mathbb{Z}$, i.e., the sequences

$$\ldots, \left(Z_{-1,l}^{(j)}\right)_{l\in\mathbb{N}}, \left(Z_{0,l}^{(j)}\right)_{l\in\mathbb{N}}, \left(Z_{1,l}^{(j)}\right)_{l\in\mathbb{N}}, \ldots$$

  are independent. In particular, for each time $k_0 \in \mathbb{Z}$, the mark sequence $(Z_{k_0,l}^{(j)})_{l\in\mathbb{N}}$ is independent of $\mathcal{F}_{k_0}^X$. Note that the mark distribution might depend on j, and, for a fixed point in time k_0, the marks $(Z_{k_0,l}^{(j)})$ might be correlated over $l \in \mathbb{N}$ and/or over the event types $j = 1, \ldots, d$.
- $\alpha_k^{(i,j)} : \mathrm{Z} \to \mathbb{R}$, $k \in \mathbb{Z}$, $i, j = 1, 2 \ldots, d$, are measurable coefficient functions such that $\sum_{k=1}^{\infty} \mathbb{E}\left[|\alpha_k^{(i,j)}(Z_{0,1}^{(j)})|\right] = 1$,
- $\mathcal{F}_n^Z := \sigma\left(Z_{k,l}^{(j)} : \ k \le n, \ l \in \mathbb{N}, j = 1, \ldots, d\right), n \in \mathbb{Z}$,
- and $\phi_{i,j} : \mathbb{R} \to \mathbb{R}_{\ge 0}$, $i, j = 1, 2 \ldots, d$, are measurable transfer functions.

Most of our results transfer to this multivariate model with quite obvious modifications of the proofs. For instance, our earlier Theorem 2 can be generalized as follows. Assume that the transfer functions $\phi_{i,j}$ are Lipschitz continuous with Lipschitz constants $L_{i,j}$, $i, j = 1, \ldots, d$, such that the $d \times d$ matrix $(L_{i,j})$ has spectral radius strictly less than 1. Then there exists an ergodic version of the multivariate Poisson autoregression with finite mean. One could even think of more general monlinear multivariate models, where the nonlinearity transfer appears on the level of the types. For the main part of the paper, we have chosen to stick with the univariate case to avoid heavy notation and further technicalities.
As a second way of generalization, one could consider the case when the transfer function is not sublinear. For still obtaining ergodicity, such a transfer function (e.g., an exponential transfer function) presumably has to be balanced by very or mostly negative coefficient functions α_k. In any case, this model would ask for very different lines of argumentation than the proofs in this paper. Finally, we mention that one could analyze the case when the Lipschitz constant of the transfer function equals 1. In this case, one would first have to give ergodicity conditions for the critical linear case of the model, and then use this critical linear model as an upper bound for the nonlinear model with Lipschitz constant equal to 1.

Acknowledgements We thank an anonymous referee for feedbacks. In particular, we thank the referee for proposing to relax our original conditions on the marks.

Author Contributions The authors contributed equally to this work.

Data Availability No datasets were generated or analysed during the current study.

Declarations

Conflict of interest The authors declare no conflict of interest.

References

1. Al-Osh, M.A., Alzaid, A.A.: First-order integer-valued autoregressive (INAR(1)) process. J. Time Ser. Anal. **8**(3), 261–275 (1987)
2. Alzaid, A.A., Al-Osh, M.A.: An integer-valued pth-order autoregressive structure (INAR(p)) process. J. Appl. Probab. **27**(2), 314–324 (1990)
3. Asmussen, S.: Applied probability and queues, volume 51 of Applications of Mathematics. Springer, Berlin, 2nd edition, (2003)
4. Baccelli, F., Brémaud, P.: Elements of queuing theory. Springer, Berlin (1994)
5. Billingsley, P.: Convergence of probability measures. Wiley, New York (1999)
6. Brémaud, P., Massoulié, L.: Stability of nonlinear Hawkes processes. Ann. Probab. **24**(3), 1563–1588 (1996)
7. Brémaud, P., Nappo, G., Torrisi, G.L.: Rate of convergence to equilibrium of marked Hawkes processes. J. Appl. Probab. **39**(1), 123–136 (2002)
8. Christou, V., Fokianos, K.: Estimation and testing linearity for non-linear mixed Poisson autoregressions. Electron. J. Stat. **9**(1), 1357–1377 (2015)
9. Doukhan, P., Fokianos, K., Rynkiewicz, J.: Mixtures of nonlinear Poisson autoregressions. J. Time Ser. Anal. **42**(1), 107–135 (2021)
10. Dudley, R.: Real analysis and probability. Cambridge University Press, Cambridge (2003)
11. Embrechts, P., Klüppelberg, C., Mikosch, T.: Modeling extremal events. Springer, Berlin (2003)
12. Fokianos, K., Rahbek, A., Tjøstheim, D.: Poisson autoregression. J. Am. Stat. Assoc. **104**(488), 1430–1439 (2009)
13. Fokianos, K., Tjøstheim, D.: Nonlinear Poisson autoregression. Ann. Inst. Stat. Math. **64**(6), 1205–1225 (2012)
14. Foss, S., Korshunov, D., Zachary, S.: An introduction to heavy-tailed and subexponential distributions. Springer, Berlin, 2nd edition (2013)
15. Huang, L., Khabou, M.: Nonlinear Poisson autoregression and nonlinear Hawkes processes. Stoch. Process. Appl. **161**, 201–241 (2023)
16. Kirchner, M.: Hawkes and INAR(∞) processes. Stoch. Process. Appl. **126**(8), 2494–2525 (2016)
17. Kirchner, M., Torrisi, G.L.: Fluctuations and precise deviations of cumulative INAR time series. Stoch. Process. Appl. **164**, 1–32 (2023)
18. Lindvall, T.: Lectures on the coupling method. Wiley, New York (1992)
19. Torrisi, G.L.: A class of interacting marked point processes: rate of convergence to equilibrium. J. Appl. Probab. **39**, 137–160 (2002)
20. Wang, H.: Limit theorems for a discrete-time marked Hawkes process. Stat. Prob. Lett. **184**, 109368 (2022)
21. Wang, H.: Large and moderate deviations for a discrete-time marked Hawkes process. Commun. Stat. Theory Methods **52**(17), 6037–6042 (2023)